A study of selected endocrine disrupting chemicals and their binding to host molecules with molecular modelling.

Anna Helena Mazurek

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2024IPPAX046

Thèse de doctorat

A study of selected endocrine disrupting chemicals and their binding to host molecules with molecular modelling

Doctorat en cotutelle

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à Département de Biologie

Thèse de doctorat de Medical University of Warsaw, Poland
préparée à Department of Organic and Physical Chemistry

Ecole Doctorale de l'Institut Polytechnique de Paris (ED IP Paris)
n°626
Spécialité de doctorat: Biologie

Doctoral School of Medical University of Warsaw, Poland
Spécialité de doctorat: sciences médicales et de la santé

Thèse présentée et soutenue à Palaiseau, 17.09.2024, par

Anna Helena Mazurek

Composition du Jury :

Tap Ha-Duong Prof., Univeristé Paris-Saclay, BIOCIS, France	Président
Matthieu Montes Prof., GBCM Lab, CNAM, Paris, France	Rapporteur
Andrzej Leś Prof., Quantum Chemistry Laboratory, University of Warsaw, Poland	Rapporteur
Beata Brzozowska Ass. Prof.,Biomedical Physics Division, University of Warsaw, Poland	Rapporteur
Maciej Dawidowski Prof., Dep. of Drug Technology and Pharmaceutical Biotechnology, Medical University of Warsaw, Poland	Examinateur
Thomas Simonson Prof., BIOC, École polytechnique, Palaiseau, France	Directeur de thèse
Łukasz Szeleszczuk Ass. Prof., Dep. of Organic and Physical Chemistry, Faculty of Pharmacy, Medical University of Warsaw, Poland	Directeur de thèse

Les perturbateurs chimiques endocriniens (EDC) sont des substances qui présentent des effets néfastes en raison d'un mode d'action endocrinien. Les EDC se lient aux récepteurs en raison de la similitude de leur structure chimique avec les hormones naturelles. Les produits pharmaceutiques constituent une autre source importante de perturbateurs endocriniens. Lorsqu'elles sont administrées sous forme de médicaments, les hormones sont décrites comme des perturbateurs endocriniens car elles modifient l'homéostasie hormonale naturelle du corps humain.

Le 17-β-estradiol (EST) est la forme la plus puissante d'œstrogènes naturels. De plus, de nombreux cocristaux d'EST ont été conçus pour résoudre l'un des problèmes majeurs associés à l'application de l'EST : sa faible biodisponibilité orale causée par une très faible solubilité dans l'eau. Ce problème pourrait être potentiellement résolu par la complexation de l'EST avec la β-cyclodextrine (βCD). Les cyclodextrines (CD) sont des oligosaccharides cycliques constitués de sous-unités de glucose (α-D-glucopyranoside).

Le cas de l'EST comme exemple d'EDC est un fait bien connu, décrit et expliqué, également au niveau moléculaire, car l'EST est un ligand naturel qui se lie au récepteur des œstrogènes. Il existe cependant de nombreux perturbateurs endocriniens dont le mode d'action n'est pas connu ou qui n'ont même pas encore été définis comme perturbateurs endocriniens. L'approche de modélisation moléculaire permet de comprendre les interactions entre la substance chimique donnée et le récepteur impacté.

Les FF additifs décrivent les interactions électrostatiques utilisant des charges atomiques à point fixe et traitent les interactions de Van der Waals via les potentiels de Lennard-Jones ou d'autres fonctions simples. Cela signifie que l'influence de la polarisation est moyennée, et que la transférabilité d'une telle charge fixe est donc faible. De plus, l'absence de multipôles atomiques d'ordre supérieur empêche une description précise du potentiel électrostatique anisotrope autour des molécules. De plus, comme les champs de force additifs n'incluent pas de représentation explicite de l'induction, ils peuvent mal représenter l'électrostatique des molécules qui jouent souvent un rôle crucial dans les interactions intermoléculaires.

Au contraire, les champs de force polarisables, c'est-à-dire ceux qui traitent explicitement la polarisation électronique, permettent à la structure électronique d'une molécule de changer en fonction des altérations du champ électrique local. En d'autres termes, dans de tels modèles, les contributions de plusieurs corps sont incluses dans les interactions électrostatiques.

Les tâches définies au début de ce travail ont été réalisées avec succès. Pour la première fois, le complexe estradiol + β-cyclodextrine a été déterminé à l'aide de méthodes expérimentales et informatiques. Le complexe a également été analysé dans la solution aqueuse en utilisant à la fois des méthodes expérimentales et une bonne variété d'approches informatiques qui ont été comparées les unes aux autres. Pour la première fois, le rapport molaire du complexe a été déterminé de manière définitive.

L'estradiol, la progestérone, le bisphénol A et la cyclodextrine ont été paramétrés avec succès dans l'AMOEBA FF polarisable. Grâce à AMOEBA FF, le système récepteur estradiol + œstrogène a été analysé et la simulation obtenue était stable.

Ce travail constitue un prélude à une analyse complexe des systèmes EDC-récepteurs et EDC-cyclodextrines qui serait suivie par la formation de lignes directrices générales sur la modélisation moléculaire concernant de tels systèmes.

Les résultats de la recherche ont été décrits dans les 3 ouvrages joints à cette thèse.

Tytuł: A study of selected endocrine disrupting chemicals and their binding to host molecules with molecular modelling

Słowa kluczowe: estradiol, cyclodextrin, AMOEBA force field, force field parametrization, DFT, endocrine disrupting chemicals

Streszczenie : Substancje zaburzające funkcjonowanie układu hormonalnego (tzw. Endocrine Disrupting Chemicals, EDC) to substancje, które wykazują niekorzystny wpływ na funkcjonowanie układu hormonalnego. Często spowodowane jest to interakcją EDC z receptorami w taki sam sposób, w jaki wiążą się z nim naturalne ligandy receptora. Wśród EDC znajdują się aktywne substancje farmaceutyczne (Active Pharmaceutical Ingredients, API), takie jak hormony steroidowe.
Cyklodekstryny (CD) to cykliczne oligosacharydy stosowane jako nośniki dla API o niskiej rozpuszczalności w wodzie oraz jako substancje usuwające toksyny. Celem tego badania było opracowanie różnych technik modelowania molekularnego w celu analizy interakcji pomiędzy wybranymi EDC a receptorem estrogenowym lub CD.
Zastosowano następujące metody badawcze: parametryzację wybranych EDC (estradiol, progesteron, bisfenol A) i CD w polaryzowalnym polu siłowym AMOEBA, a następnie symulację dynamiki molekularnej układu Receptor Estrogenu + EDC; testy porównawcze różnych podejść obliczeniowych opartych na mechanice kwantowej (DFT, podejścia półempiryczne) i mechanice molekularnej (MD/MMGBSA) jak i testowanie wybranych parametrów obliczeń, na przykładzie układu estradiol + βCD.

This international PhD project, realized in cooperation between École Polytechnique and Medical University of Warsaw, was possible thanks to the French Government Scholarship (Bourse du gouvernement français, BGF) issued by Campus France.

In Poland the research was financed by two Young Researcher grants issued by the Medical University of Warsaw in 2022 (F/MB/02/22) and in 2023 (FW7/1/F/MB/N/23).

The cooperation with the Agricultural University of Athens in Greece was possible thanks to the STER grant issued by the Polish National Agency for Academic Exchange.

I would like to thank my both supervisors: **Prof. Łukasz Szeleszczuk** and **Prof. Thomas Simonson** for their constant support and, to paraphrase J.W. Goethe, for being this good teacher who treated me as if I was what I needed to be, to bring me where I needed to get.

Moreover, I would like to express my great thanks to **Carine Clavaguéra** from Institut de Chimie Physique, Université Paris-Saclay, for her enduring support in the topic of polarizable force field parametrization.

Wholeheartedly, I would like to thank **Kostas Bethanis** from Physics Lab, Biotechnology Department, Agricultural University of Athens who has accepted me for a project in the topic of cyclodextrin complexes crystallization what happened to change the course of my PhD project and later started to be a stable scientific collaboration and friendship.

Last but not least, I would like to leave here my great thanks to **Prof. Dariusz Maciej Pisklak** who believed in me from the same beginning of my pharmacy studies and who already then has seen me as a possible young scientist. I am thankful for endless scientific discussions, his willingness to help, his full engagement and a very human side of this person.

‘’Science is a very human form of knowledge.

We are always at the brink of the Known,

We always feel forward for what is hoped.

Every judgement in science stands on the edge of error, and is personal.

Science is a tribute to what we can know although we are fallible.’’

Jan Bronowski, Polish-British mathematician and philosopher

Contents

Introduction

According to the European Commission Regulation from 2018 [1], Endocrine Disrupting Chemicals (EDCs) are substances that exhibit adverse effects as a consequence of an endocrine mode of action. EDCs bind to receptors due to the similarity of their chemical structure shared with natural hormones. Examples and sources of EDCs are presented in Fig. 1.

Those are among others: dichlorodiphenyltrichloroethane (DDT) present in pesticides, parabens from cosmetics, phthalates which are products originating from plasticizers' depolymerization, bisphenol A which is a depolymerization product of polycarbonates and epoxyd resins present e.g. in water bottles, dioxins from paper industry, pharmaceuticals, plant/mushroom derivatives etc. [2].

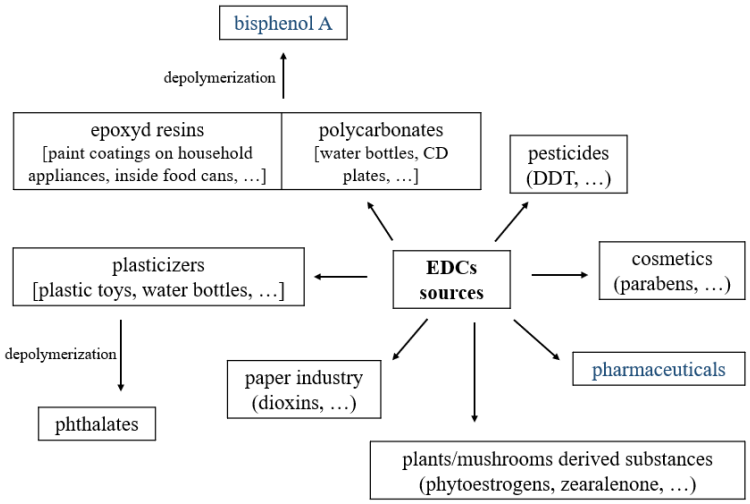


Fig. 1 Examples of EDCs sources.

A separate important source of EDCs are pharmaceuticals. When delivered as medication, hormones are described as EDCs as they alter the natural hormonal homeostasis in a human body. An important group of hormonal drugs are estrogens. In case of such external application of the naturally occurring estrogens, they are sometimes described as 'xenoestrogens' [3]. The same name is applied to other substances mimicking estrogens, like plant or mushroom derived substances (phytoestrogens, zearalenone etc.) [4,5].

This particular type of EDCs, xenoestrogens, poses a specific problem especially in the densely inhabited regions. Xenoestrogens can be found in wastewater and drinking water. Although the issue of removing xenoestrogens from water is not new, it is nevertheless a recurrent one and the subject of conflicting views. In 2009 in a broadly cited publication Daniel J. Caldwell *et al.* reported that the level of estrogens present in drinking waters in the United States does not exceed the margins of safety [6]. However, from today's perspective, two aspects must be taken into account. Firstly, already 15 years has passed since this study was performed. Secondly, even if this study is concerning a huge country, these results cannot be extrapolated for other regions in the world, for instance less-developed countries where the wastewater purification methods are less technically developed. A much more recent publication from 2020 about the occurrence of EDCs in Malaysian drinking water serves as a good example [7]. According to this report, the reproductive system connected hormones like testosterone, progesterone, estrone, 17β-estradiol and 17α-ethynylestradiol were observed to reach mean concentrations from 0.03 to 0.83 ng/L and 0.20 to 1.59 ng/L in river and tap

water, respectively. However, it has also been demonstrated that certain substances, such as 17α-ethynylestradiol can exert a triggering effect towards the endocrinal disfunction already at concentrations below 1ng/L [8]. Higher EDCs levels in the tap water then in river water are explained in the study by the water supply chain and purification methods malfunction.

As it was mentioned above, the scientific results are not unambiguous and coincident depending on the region of the measurements. The one thing is the level of xenoestrogens in waste and tap water, the other is their influence on human health if delivered in such quantities as detected. In 2020 a comprehensive review has been published on the influence of the present in water EDCs on the reproductive system [9]. With regards to estrogens, the authors state that ‘’estrogens that contaminate surface waters worldwide can negatively influence the fertility and reproductive capacity of humans’’ but at the same time they claim that ‘’data are limited on the levels and types of estrogens in the environment’’. This explains why the water contamination with EDCs, and especially with estrogens, is still a current scientific topic.

As the toxicological studies have defined the predicted no-effect concentration for estradiol to be ranging from 1 to 5 ng L/1 [10,11] Caldwell et al. 2012; Laurenson et al. 2014) and for ethinylestradiol (contraception) from 0.035 to 0.35 ng L/1 [10,11], in 2011 the European Commission proposed environmental quality standards for estradiol and ethinylestradiol as 0.4 and 0.035 ng L/1, respectively [12,13]. A drinking water quality standard of 1 ng L/1 was proposed for estradiol [14], as advised by the World Health Organization. This means that right now there are well-defined levels which are acceptable at least in the EU. In the recent years, numerous water purification systems targeted at steroidal hormones have been developed as reviewed in 2023 [15]. It seems that, thanks to quite a few adjustments and technological progress over the last few years, we have arrived at the systems which are able to eliminate hormones like estrone, 17β-estradiol and 17α-ethinylestradiol to almost non-detectable levels [16,17]. As previously mentioned, there are still areas where those techniques are not used, and research is still being done to find better, more affordable, and more efficient technology.

After pharmaceuticals, another substantial EDCs group are pesticides, defined by the European Union (EU) as Endocrine Disrupting Pesticides (EDPs). After almost 15-year-long procedure, the first EDP was banned in EU only in 2023 [18]. Even though EDPs are similarly well-described and regulated, the removal of EDPs from water is a much more complex topic because, unlike estrogens, EDPs frequently exhibit significant structural differences from one another, making it more difficult to develop one method applicable to all molecules.

One of the toxin removing agents are cyclodextrins (CDs). Those are non-toxic cyclic oligosaccharides which can form inclusion complexes [19]. Moreover, complexation between a CD and a molecule characterized by a low solubility in water, enhances the bioavailability of the molecule

[20]. This fact is widely known and used in the pharmaceutical industry. It will be explained in details in the further part of the thesis.

When it comes to the objects, the main concern in this work has been put on 17-β-estradiol, also known as estradiol. The goal of the project was to obtain and analyze the structure of the estradiol+β-cyclodextrin complex. If successful, this would be the first time when a steroidal hormone encapsulated in a cyclodextrin has been described. This could be also a beginning for further analysis of steroidal hormones and cyclodextrin complexes for potential both pharmaceutical and toxicological uses. This, as it will be explained later, requires examination both in the water solution and in the solid state. The same concept could be applied to other, non-pharmaceutical EDCs.

17-β-estradiol (EST) is the most potent form of naturally occurring estrogens [21]. Therefore, it has found wide application in hormonal contraception, hormone replacement therapy (HRT), and treatment of menopausal and postmenopausal symptoms [22]. Oral administration of EST in a solid dosage form is the most favourable form of HRT [23]. While in the European Pharmacopoeia only the hemihydrate form of EST is described, recently its anhydrous form was successfully obtained [24]. Moreover, numerous cocrystals of EST have been designed [25,26] to solve one of the major problems associated with the application of EST: its poor oral bioavailability caused by very low water solubility (0.2–5 μg mL) [27]. This issue could be potentially solved by EST complexation with β-cyclodextrin (βCD).

The case of EST being an example of EDC is a well-known, described and explained fact, also at the molecular level because EST is a natural ligand binding to the Estrogen Receptor. However, there are numerous EDCs whose mode of action is not known or which have not even been defined as EDCs yet. And this all in the situation when more and more potential EDCs are being put on the market yearly. For so numerous cases, the molecular modelling approach is probably the best choice: it will help to understand the interactions between the given chemical substance and the impacted receptor. Moreover, computational approach, if properly constructed, could be used before the experimental examination as a first screening method for detection of possible EDCs. In order to create such computational verification model, firstly the best theoretical approaches [Fig. 2] and technics for such analysis must be chosen and developed.

In such studies there are always two general areas of interest: the structure (geometry and intermolecular interactions) and the energy of the system. Interaction absolute energy between host and guest is described by the following basic equation (eq. 1 [28]), where a host can be e.g. a protein, DNA, CD and a guest is a ligand, e.g. drug or toxin molecule:

$$E_{system} = E_{complex} - \left(E_{host} + E_{guest}\right) \quad (1)$$

The thermodynamic properties like enthalpy (ΔH) and entropy (ΔS) can be derived from the computation and they sum up to the Gibbs free energy of binding (ΔG), eq. 2 [28]:

$$\Delta G = \Delta H - T\Delta S \quad (2)$$

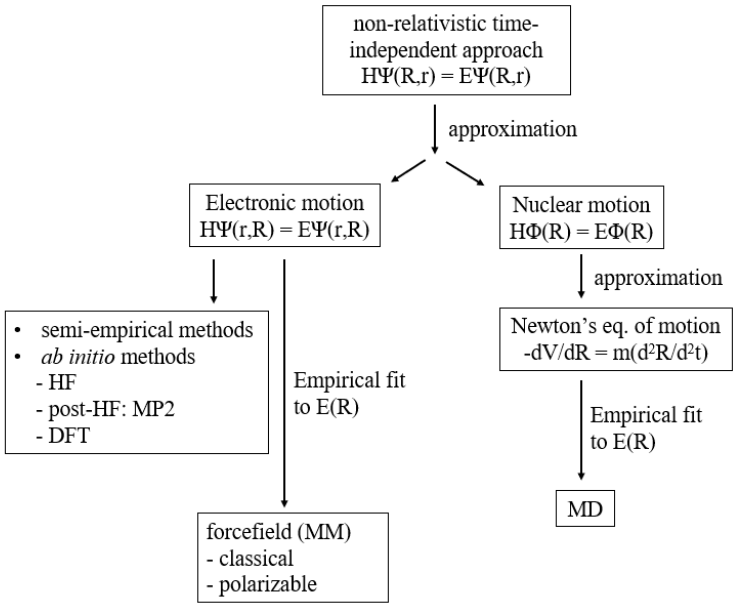


Fig. 2 A general scheme of non-relativistic time-dependent molecular modelling approaches.
Ψ and Φ - wave function in, respectively: position and momentum representations, H - Hamiltonian, E - potential energy surface, R - coordinates of the nuclei, r - coordinates of the electrons, V - velocity.

The first and already signalized topic is interaction between an EDC and a CD. In silico methods are widely applied to different aspects regarding CDs, xenoestrogens and toxins in general. Good example of the variety of objects and wide spectrum of methods used for this purpose are the following recent works: β-CD complexation with methyldrostanolone [29] which is both a toxin and estradiol derivative (conformational analysis, 2022), encapsulation of sarin by heptakis(2,3,6-tri-O-methyl)-β-CD (MD simulation and QM structural analysis, 2021) [30]. For years, CD complexes have been analyzed using the Molecular Mechanics (MM) approach as the computational power available at the time was not sufficient to apply the Quantum Mechanical (QM) methods. This has begun to change in the course of the last few years. CD complexes are started to be examined using QM-based methods, however there is no consistency in the techniques and parameters applied. Therefore, there was a need to perform benchmarking tests on the chosen example of CD-including complex. One of the aims of this work was to analyze the structure and thermodynamic properties of EST-βCD complex in water solution and in solid state using different computational approaches (semi-empirical, Moller-Plesset, DFT) and testing various computation parameters. The results were compared to the experimental data which was obtained in the first step of this work.

Nevertheless, to make this study complete, a previously standard approach used for the analysis of CD complexes, MD-MMGBSA calculations, has been applied, as well.

The second already mentioned aspect is interaction between EDC and a receptor. For the analysis of such systems MM-based approach must be introduced. Here, the forcefield (FF) term is used. FF is a set of mathematical potentials and parameters extracted from ab initio and/or experimental data and is used to calculate the energy of the inter- and intramolecular interactions between atoms [31]. There are two main types of FFs: classical or additive and polarizable.

Additive FFs describe electrostatic interactions using fixed point atomic charges and treat van der Waals interactions via Lennard-Jones potentials or other simple functions [32]. This means that the influence of polarization is averaged, hence the transferability of such fixed-charge is low. Moreover, lack of higher order atomic multipoles prevents an accurate description of the anisotropic electrostatic potential around molecules [33]. What is more, as additive force fields do not include explicit representation of induction, they may poorly represent the electrostatics of molecules which often play a crucial role in the intermolecular interactions [34].

On the contrary, polarizable force fields, that is those which treat electronic polarization explicitly, allow the electronic structure of a molecule to change with regards to alterations of the local electric field. In other words, in such models multi-body contributions are included in the electrostatic interactions. As J. A. Lemkuhl has described [35]: ‘’if a molecule is removed from the system, the dipoles of the other species will be aligned differently and will have different magnitudes, leading to different interaction energies among the remaining molecules’’. This is an answer for the non-transferability characteristic for the additive force fields. Superiority of polarizable force fields over the classical ones has been depicted on a great variety of objects [36-67].

In polarizable force fields the many-body interaction energy is explicitly treated through the introduction of electronic polarization. This can be implemented through application of [68]:

- fluctuating charge models: fluctuating charge represents the response of the system to the electrostatic potential [69]
- Drude oscillator models: Drude particles on polarizable sites describe the response of the system to the surrounding [70]
- atomic induced dipole models: induced dipoles respond to the surrounding electrostatic field [40,47].

A force field which uses the third approach is AMOEBA FF. It is being developed since 1990s and currently there is available a full set of parameters for proteins, nucleobases, organic molecules [71-73]. The parametrization process has been automatized and for this purpose the Tinker software is frequently used [74]. However, still there have been published only few studies applying this approach to big systems like a receptor-ligand complex. More research is needed in this direction.
Therefore, one of the purposes of this work was to, in the first place, parametrize selected EDCs and secondly, perform a receptor-EDC simulation using AMOEBA FF. The chosen molecules are: estradiol, progesterone and bisphenol A. Both estradiol and bisphenol A are model representatives of EDCs, with estradiol being a natural hormone whose receptor binding is mimicked by EDCs. Progesterone has been chosen for the two reasons. Firstly, it is another example of a potent pharmaceutical EDC. But even though, we know significantly less about progesterone’s binding to the progesterone receptor than about the estradiol + estrogen receptor interaction. This makes the

‘progesterone+progesterone receptor’ an interesting system to analyze. Secondly, so far there were no parameters available for the steroid fused rings core which is a basis for multiples molecules including hormones. Therefore, the first challenging element of this part of the work was parametrization of the three molecules. In the next step, the assumption was to use at least one of those molecules (preferably estradiol as a model molecule) to perform Molecular Dynamics simulation with receptor, in this case estrogen receptor, using AMOEBA FF.

As it is explained in the further part of this work, when Molecular Dynamics calculations including cyclodextrins are performed, a carbohydrates-targeted additive GLYCAM force field is used [75]. The limitations of the classical approach are well-known and already when the latest version of GLYCAM was published, the authors mentioned works on the polarizable version of this force field. However, having already a well-functioning polarizable AMOEBA FF, it has been decided to include a cyclodextrin molecule in the parametrization process. This would allow to perform MD simulations on the cyclodextrin-EDC complexes using the polarizable FF approach. This would be also complementary to other previously mentioned computational approaches used in the benchmark analysis of CD-including systems.

Endocrine Disrupting Chemicals – mechanisms of action

A highly cited article from 2020 [76] points out that one of the issues regarding EDCs is lack of well-defined characteristics of such hazardous substances. This is especially crucial as the regulatory agencies use various approaches to evaluate the hazard coming from potentially endocrine disrupting chemicals. Michele A. La Merrill et. al. propose 10 EDCs key characteristics based on the end points of their acting. According to this research, as an EDC can be defined a substance which:

- interacts with or activates hormone receptors
- antagonizes hormone receptors
- alters hormone receptor expression
- alters signal transduction in hormone-responsive cells
- induces epigenetic modifications in hormone-producing or hormone-responsive cells
- alters hormone synthesis
- alters hormone transport across cell membranes
- alters hormone distribution or circulating hormone levels
- alters hormone metabolism or clearance
- alters fate of hormone-producing or hormone-responsive cells

It is worth mentioning that two of the most well-known and described EDCs: bisphenol A and already withdrawn diethylstilbesterol, fulfil 9 out of 10 above mentioned key characteristics.

A thorough discussion of the EDCs and the disease endpoints, including reproductive, metabolic, neurologic and cardiovascular disorders, can be found in a recent review on the topic [77].
EDCs can be absorbed by a human body via digestive system, skin and inhalation or even via placenta to the foetus. An example for the latter, is a perinatal exposure to bisphenol A which causes physiological and functional underdevelopment of genitalia, tracts and glands that may result in reduced fertility, aspermia, immature reproductive systems and the growth of several cancers such as breast, ovary and prostate cancer [78].

In a human body EDCs target primarily 6 receptors: estrogen, androgen, progesterone, thyroid hormone, glucocorticoid, peroxisome proliferator-activated receptors gamma and aryl hydrocarbon receptors [79,80]. There are two main mechanisms of interaction between natural activators and receptors: direct (known also as 'genomic') and indirect (or 'non-genomic') [78,79]. The same mechanisms are used by EDCs. In the direct mechanism, the ligand binds directly to a receptor and therefore affects the transcription of target genes in the nucleus. In the indirect mechanism, the ligand interacts with the components of the hormone signalling pathways, for instance with G protein-coupled receptor (GPR30) located in the cytoplasmic membrane. Activation of GPR30 by a ligand leads to downstream cellular signalling like protein kinase activation and phosphorylation what in turn may affect the transcription of target genes. In fact, what is observed, is the pleiotropic effect induced by a ligand via different pathways (nuclear and extracellular) and by interactions with different receptors, like ERα, ERβ, GPR30, depending on the location within the cell and the body [81]. The same differentiation in used mechanisms is observed for EDCs [81,82]. Regardless of the mechanism, EDCs alter the endogenous synthesis of hormones. This leads to toxic effects like hormonal imbalance, decrease of fertility, alterations in sperm quality and fertility, abnormalities in sex organs, endometriosis, early puberty, altered nervous system function and immunity, sex organ cancers etc. [83].

R.K. Gupta et al. underlines that the reproductive hormones, such as progestins, androgens, and estrogens are the primary targets of EDCs such as: pesticides (e.g. dichlorodiphenyltrichloroethane (DDT)), methoxychlor, vinclozolin, atrazine), detergents and surfactants (e.g. octyphenol, nonylphenol, bisphenol A (BPA)), plasticizers (e.g. phthalates), industrial compounds (e.g. polychlorinated biphenyls (PCBs), 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD)), natural plant derivatives (e.g. genistein, coumesterol) [84].

Estrogen Receptor

There are two subtypes of the estrogen receptor: ERα and ERβ, each of them characterized by a tissue-specific expression [85]. Despite being encoded by different genes, both estrogen receptors show high homology, and in both of them the E domain contains the ligand-binding domain (LBD)

and C domain, which is the DNA binding domain (DBD) [86]. In ER structures, two transactivation functions are present, called AF-1 (located in N-terminal domain) and AF-2 (located in LBD) [87]. They contain the nuclear location signals. After proper exposure of their surface, happening as the result of the ligand binding, they are responsible for incorporation of the co-activators, which is a necessary step to induce activation of the intercellular signalling pathways [Fig. 3].

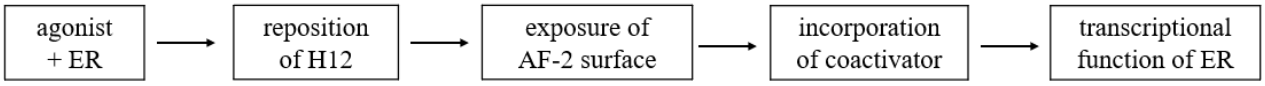


Fig. 3 A general scheme of ER activation.

Binding of an agonist to the ER requires creating a hydrogen bond with His524 (in ERα) [88] or His475 (in ERβ) [89]. This leads to a unique agonist-bound conformation of the receptor's LBD, characterized by a specific repositioning of the H12 helix, which is the most C-terminal helix of the LBD (molecular switch) [90]. On the contrary, selective ER modulators (SERMs) such as raloxifene or tamoxifen induce relocation of H12 into the co-activator binding cleft, which blocks AF-2 activity. Finally, pure antagonists completely destabilize H12 [91].

There are four natural steroid hormones which act as activators towards ER. Those are: estrone (E1), estradiol (E2, EST), estriol (E3), and estretrol (E4) [92]. The last one is produced only during pregnancy by the foetus liver [93]. Among these four compounds, E2 plays the most important role in the human organism and, therefore, is of high importance in breast or ovarian cancer progression. With regard to their relative binding affinity (RBA) to ER, with the exclusion of estretrol, estrogens are ranked in the following order: estradiol > estrone > estriol. In comparison to estradiol, the activity and potency of estrone and estriol are, respectively, 10 and 100 times lower [94].

E1 functions mainly as estradiol's metabolite and serves as its precursor (the estrone-to-estradiol transformation is reversible) [95] [Fig. 4]. E3 is a metabolite of E1 [Fig. 4], however, in non-pregnant women, estriol levels in the blood are hardly detectable. During pregnancy its amount distinctively grows because it it is produced by the placenta as well [96]. All estrogens are used as medication in menopausal hormone therapy, although estradiol is the most applied [97,98].

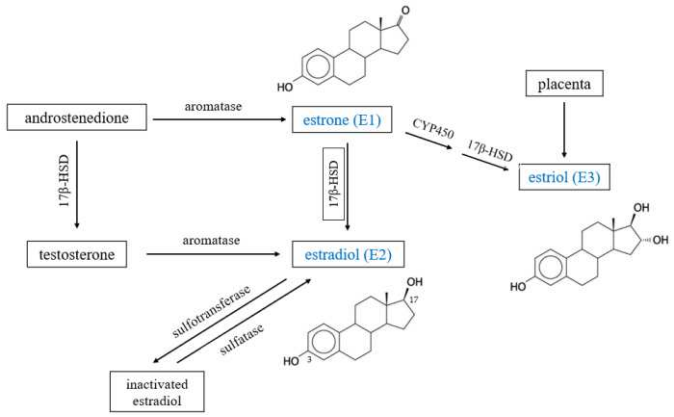


Fig. 4 A general scheme of estradiol's metabolism.

When compared to ERβ, ERα plays a more prominent role in the mammary gland and uterus as well as the regulation of metabolism. ERα LBD is composed of 12 α-helices (H1-H12) [99-103].

Estradiol (EST) binds as an agonist in the pocked formed by 22 residues. EST hydroxyl groups play a decisive role in the hormone positioning within the pocket. The hydroxyl group of the A ring [Fig. 5] creates a hydrogen bond with Glu353 from H3, Arg394 from H5 and water molecule, whereas hydroxyl group of the ring D creates a hydrogen bond with His 524 from H11 [99-103]. Creation of the hydrogen bond with H11 allows repositioning of the H12 what in turn generates a ligand-dependent activation function 2 (AF-2). It is necessary for the interaction with co-activators and later initiation of the intercellular signalling pathway [104,105].

Except for the already mentioned hydrogen bond interactions with Glu353, Arg394 and His524, EST molecule position is stabilized also by the π- π stacking with Phe404 [99-103].

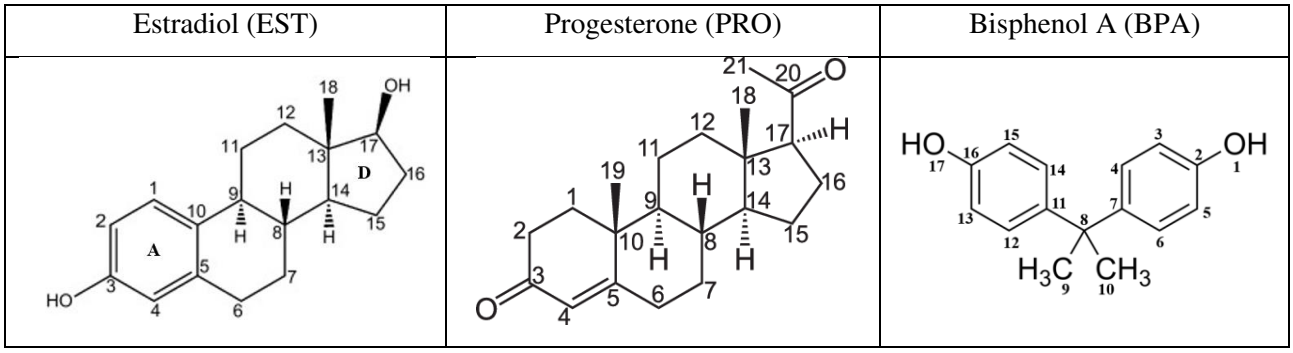


Fig. 5 Structures of selected EDCs.

Due to the structural similarity with EST, bisphenol A (BPA) [Fig. 5] binds to both types of ER. It displays 1000- to 2000-fold less affinity to ER then EST does [81]. BPA is ERα activator via the same mechanism as EST. Towards ERβ, it acts as an antagonist because it prevents LBD from obtaining the activated type of conformation [106]. BPA shows also a high binding affinity towards GPR30 [81]. This shows that the disruptive influence of BPA on the hormonal homeostasis happens via multiple mechanisms. It has been also proven that BPA interacts with other hormonal receptors like androgen, pregnane X, and peroxisome proliferator-activated receptors [81]. This example highlights to which extent a single EDC can disrupt the functioning of a human hormonal system.

On the contrary to BPA, progesteron (PRO) does not bind to ER. This is due to the absence of hydroxyl groups at carbons 3 and 17 of PRO molecule. PRO binds to progesterone receptor (PR). The molecule is anchored in the PR binding pocket via net of hydrogen bonds created around carboxyl oxygen attached to PRO's carbon 3 [107,108]. PRO binds to PR and causes activation of its transcriptional function in a mechanism similar to the one described for ER: AF-2 activity is mediated by a hormone-dependent interaction with steroid receptor coactivators (Src) [109].

Both ER and PR undergo a dimerization which happens after binding of the agonist [110,111]. With regards to this process, an important element of the ERα LBD structure is Tyr537. Its phosphorylation has been proven to influence the hormone binding, ER dimerization and transcriptional activity [112]. Src family tyrosine kinases were shown to specifically phosphorylate

ER's Tyr537 [113]. This estradiol-dependent ER phosphorylation at Tyr537 plays a crucial role in the nuclear export of ERα.

Cyclodextrins

Cyclodextrins (CDs) are cyclic oligosaccharides consisting of glucose (α-D-glucopyranoside) subunits joined by α-1,4 glycosidic bonds [Fig. 7]. The so-called native CDs are not substituted and are composed of 6 (α-CD), 7 (β-CD) or 8 (γ-CD) units. CDs are non-toxic and therefore can be used as drug delivery agents [114]. CDs are used in pharmaceutical formulations due to their ability to form inclusion complexes. Due to the presence of hydroxyl groups, the external fragments of CDs are polar. When a non-polar substance enters the molecular hole of CD, the formed host–guest complex is polar and more soluble than a separate guest molecule [115]. Therefore, CDs are commonly used to increase the solubility of API (Active Pharmaceutical Ingredient) or protect it from external factors like light, humidity and heat. Worldwide, more than 100 original drugs have been ever manufactured with CDs as excipients [116–118]. One of the APIs groups characterized by poor solubility in water are hormonal steroids like estradiol or progesterone. Encapsulation in CDs may enhance their solubility in water and as a result also their bioavailability. Based on the same principle of the encapsulation CDs found a second application which is their usage as toxin (e.g. EDCs) removing agents.

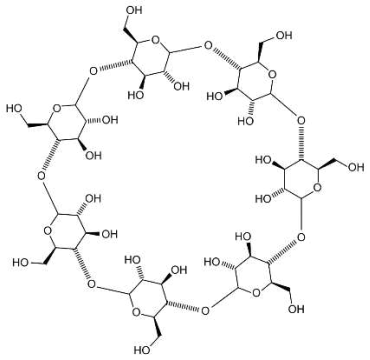


Fig. 7 Structure of β-CD.

According to the Web of Science, within the last decade each year more than 1000 articles concerning the CD-including drug delivery systems have been published and since 2018 this number is visibly rising. The topics encompass such inventions like liposomes+CD+ligand, nanotubes+CD+ligand or gold layer+CD+ligand. Already even a couple of review articles has been written on this subject [119-121]. According to the EMA (European Medicine Agency) [122], there are some CD-complexed drugs at the European market, formed with SBE-β-CD (sulfo-butyl-ether-CD) or 2-hydroxypropylo-β-CD (2-HP-β-CD). Currently, on the European market there is one CD-hormone medication. It is RM-β-CD nasal spray for hormone replacement therapy by 17-β-estradiol [123]. Nevertheless, still new attempts are made in this topic and CDs are generally considered as good non-toxic agents enhancing solubility of the low water soluble chemical compounds.

In terms of extraction, they are often used in organic solvents being attached to the chromatographic columns [124,125]. Both in the experimental and computational studies apart from the 'natural' CDs (α, β, γ) also the ones with attached different side chains are used, for instance the already mentioned SBE-β-CD [126,127] and 2-HP-β-CD [128-131] or 2,6-dimethylo-β-CD [132,133], methyl-β-CD [134,135]. Among all CDs, the most often used ones are 2-HP-β-CD and β-

CD. This is due to the fact that most of chemical compounds (potential drugs and toxins) are too big to enter the cavity of the α-CD. In turn, the γ-CD is in most of the cases too wide and therefore the binding affinity between the CD and the guest is weaker. The 2-HP-β-CD is typically chosen among the CD derivatives because, from a synthetic perspective, a structural alteration from the β-CD is relatively easy and still, in many cases, the 2-HP-β-CD's solubility enhancing abilities are sufficient enough. In experimental works methylation or 2-hydroxypropylation happens randomly. In the *in silico* research such attempt is not that common as it would require specifying places at which a side chain is added so it would not be 'random' anymore. It is more popular to use fully substituted CDs (per-methylated, per-2-HP-hydroxypropylated etc.) [136-138]. In this project β-CD was used.

AMOEBA forcefield

AMOEBA FF uses the concept of ***atomic multipoles***. Atomic multipole term defines that each atomic centre consists of partial charge, dipole vector and quadrupole tensor. For the dipole and quadrupole description ***local coordinate frames*** are constructed at each site. They are constructed according to the z-then-x convention [47,73], as described in Fig. 6. The multipole moments are derived directly from ab initio quantum mechanical electron densities for small molecules and molecular fragments. For this purpose the ***Distributed Multipole Analysis*** (DMA) of wavefunctions expressed in terms of Gaussian atomic orbitals is used. It is carried out in the Gaussian software (GDMA) [139].

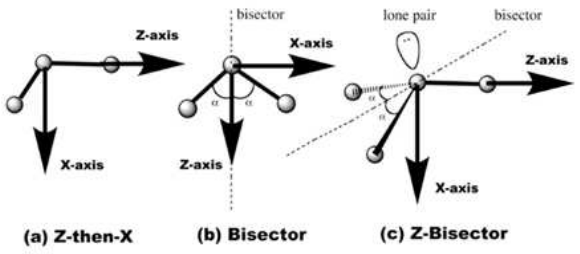


Fig. 6 Local coordinate frame definitions for atomic multipole sites.

Adapted from [32] under the CC BY 4.0. licence.

a) The Z-then-X frame is used for general sites, and with addition of a third orthogonal y-axis can treat chiral centers. The majority of AMOEBA multipole sites are defined using this local frame. (b) The Bisector frame is useful for molecules with 2-fold local symmetry or pseudo-symmetry, such as water and aliphatic methylene carbon atoms. (c) The Z-Bisector frame is used for sites such as the sulfur atom of dimethylsulfoxide, which have a distinct primary ("Z") axis and symmetry or pseudo-symmetry along a secondary direction.

Induced dipoles (μind i,α) are described by atomic polarizability (αi) and influence of the electric field on the atom i (Ei,α):

$$\mu ind\ i,\alpha = \alpha i\ Ei,\alpha. \quad (3)$$

Polarization term is defined as a sum of atomic multipoles' response terms for the electric field created by non-connected atoms and interaction terms between inducted atomic dipoles. Polarization is explicitly treated by mutual induction of dipoles at polarizable sites (located at atomic

centers) [47]. A point dipole moment is induced at each polarizable site with regards to the electric field experienced by that site, according to the eq. 4:

$$\mu_i^{ind} = \alpha_i\left(E_i^{dir} + E_i^{mut}\right) \quad (4)$$

where α_i is the atomic polarizability on site i; E_i^{dir} is the "direct" electric field generated by permanent multipoles of other sites; E_i^{mut} is the "mutual" field generated by induced dipoles of other sites [68].

In other words, induced dipoles produced at the atomic centers mutually polarize all other sites. Based on Thole's model, [140] polarization at a very short range is damped, what delivers energies in a better agreement with ab initio results and allows to avoid the so-called polarization catastrophe [71]. Atomic polarizabilities are assigned based on the element type of each atom [73]. When short-range polarization between bonded atoms is ignored, use of intramolecular polarization delivers only marginal improvement when compared with the nonpolarizable potentials. To overcome this problem, a group-based intramolecular polarization scheme has been introduced. Those ***polarization groups*** are usually functional groups with limited conformational degrees of freedom [32,71]. They are partitioned between rotatable bonds [73]. This concept prevents permanent multipoles from polarizing other atoms within their group but induced-induced polarization occurs between all atoms.

The ***polarization energy*** between induced dipoles and permanent multipole moments is computed fully between atoms separated by three (1-4) or more bonds, and completely neglected for any closer separation [71].

In the AMOEBA FF atomic interactions are defined as ***bonded and non-bonded interactions***, according by the following equations:

$$U = U_{bond} + U_{angle} + U_{b\theta} + U_{torsion} + U_{oop} + U_{vdw} + U_{ele\ perm} + U_{ele\ ind} \quad (5)$$

where the first five terms describe the short-range valence interactions (bond stretching, angle bending, bond-angle cross term, torsional rotation, out-of-plane bending, please see full equations: eq. 2-7) and the next three terms describe: nonbonded vdW and electrostatic contributions.

$$U_{bond} = K_b(b-b_0)^2[1 - 2.55(b-b_0) + 3.793125(b-b_0)^2] \quad (6)$$

$$U_{angle} = K_\theta(\theta-\theta_0)^2[1 - 0.014(\theta-\theta_0) + 5.6x10^{-5}(\theta-\theta_0)^2 - 7.0x10^{-7}(\theta-\theta_0)^3 + 2.2x10^{-8}(\theta-\theta_0)^4] \quad (7)$$

$$U_{b\theta} = K_{b\theta}[(b-b_0) + (b'-b'_o)](\theta-\theta_0) \quad (8)$$

$$U_{torsion} = \sum_n K_{n\varphi}\,[1 + \cos(n\varphi \pm \delta)] \quad (9)$$

$$U_{oop} = K_\chi \chi^2 \quad (10)$$

$$U_{vdw(ij)} = \varepsilon_{ij}\left(\frac{1.07}{\rho_{ij}+0.07}\right)^7 + \left(\frac{1.12}{\rho_{ij}^7+0.12} - 2\right) \quad (11)$$

Equations 6-11 describe: bond stretching, angle bending, bond-angle cross term, out-of-plane bending, torsional rotation energy and vdW terms in AMOEBA FF, where K_b is bond force constant, b-b0 is distance

from equilibrium after atom movement, K_θ is angle force constant, $\theta - \theta_0$ is angle from equilibrium between 3 bonded atoms, $K_{n\varphi}$ is dihedral force constant, n is multiplicity of the function, φ is dihedral angle, δ is phase shift, K_χ is out-of-plane bending constant, χ is an angle created between 4 atoms; R_{ij} is separation distance between atoms i and j ($\rho_{ij}=R_{ij}/R_{0ij}$ where R_{0ij} is minimum energy distance and is combined for heterogeneous atom pairs); εij is potential minimum combined for heterogeneous atom pairs

Additive forcefields

- **CHARMM forcefield**

In comparison to the polarizable FF, a potential energy function of an additive FF is composed as presented in equation 12, on the example of CHARMM FF (Chemistry at HARvard Macromolecular Mechanics). The main difference lies in the absence of the electrostatic contribution description which is a core element of a polarizable FF [141]. In this work, CHARMM FF has been used to evaluate the parametrization process of AMOEBA FF. [ttp://dx.doi.org/10.1016/j.bbagen.2014.08.004].

$$V = \sum_{bonds} k_b(b - b_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} k_\Phi[1 + \cos(n\Phi \pm \delta)] + \sum_{impropres} k_\omega(\omega - \omega_0)^2 + \sum_{Urey-Bradley} k_u(u - u_0)^2 + \sum_{nonbonded} \varepsilon \left[\left(\frac{R_{\min ij}}{r_{ij}}\right)^{12} - \left(\frac{R_{\min ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{\varepsilon r_{ij}} \quad (12)$$

where in bond stretches term: k_b is bond force constant, $b\text{-}b_0$ is distance from equilibrium after atom movement; in bond angles term: k_θ is angle force constant, $\theta - \theta_0$ is angle from equilibrium between 3 bonded atoms; in dihedrals (torsion angles) term: k_Φ is dihedral force constant, n is multiplicity of the function, Φ is dihedral angle, δ is phase shift; in impropers (out of plane bending) term: k_ω is force constant, $\omega\text{-}\omega_0$ is out of plane angle; Urey-Bradley term is cross-term accounting for angle bending using 1,3 nonbonded interactions: k_U is respective force constant, U is distance between 1,3 atoms in harmonic potential; last two terms account for nonbonded interactions between pairs of atoms i and j.

- **AMBER forcefield**

Another additive force field applied in this work is AMBER FF (Assisted Model Building with Energy Refinement) [142]. The potential energy function is calculated according to the eq. 13 which, similarly as in other additive force fields, consist of terms for bonds, angles, dihedrals, van der Waals interactions and electrostatics.

$$V = \sum_{bonds} k_b(b - b_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\Phi - \delta)] + \sum_{ij} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{r_{ij}} \quad (13)$$

These terms are derived with use of the Antechember software which uses general AMBER FF for organic molecules (GAFF) [143]. The terms are assigned based on the atoms connectivity [144].

In this work AMBER FF has been used due to its particularity: an adjustment called GLYCAM which is AMBER FF adapted for the carbohydrates [75]. In the newest version, GLYCAM06j, bond and valence angle deformation force constants, dihedral angle rotational barriers, electrostatic properties were obtained with QM calculations, as those parameters are hardly obtainable experimentally.

Partial atomic charges are derived by fitting to the QM molecular electrostatic potentials (ESP-fitting). However, in contrast to older GLYCAM versions, partial charges are not fitted to aliphatic hydrogen atoms. In the GLYCAM06j version, for the ono-bonded interactions the 1-4 scaling has been removed.

In this work GLYCAM force field was applied to perform MD/MMGBSA analysis of the cyclodextrin-estradiol complex. The underlying theory for the MD/MMGBSA calculations is described in the further part of this thesis.

Molecular Dynamics / MD-MMGBSA approach

- **Molecular Dynamics – receptors in solution and crystal structures**

For the Molecular Dynamics (MD) calculations the underlying physics is defined by the Newton equation of motion [145]. The analyzed system might be either objects placed into a solvent box which is replicated into infinity (for solutions, for instance: receptor simulation) or a whole crystal structure recreated thanks to an infinite replication of the crystal unit cell (for solid state). Calculations are performed in one of the ensembles: NVE (microcanonical), NVT (canonical), NPT (isothermal-isobaric), where N states for number of particles, V for volume, T for temperature, P for pressure. Each time, the given parameters (N, V, E, P) are restrained to the imposed values. First stage after solvent box / crystal unit cell preparation is system's heating up to the desired temperature and later two-stage equilibration, till firstly temperature and later pressure oscillates around the imposed values [145]. The following step is the production run. Positions and velocities from the MD trajectories which define movements of atoms, are used to compute the structural and thermodynamic properties. The above described method is referred to as a classical MD. Several variations have been already constructed, among them ab initio MD, which is said to be probably the most precise approach, as it starts from the QM-optimized structures, is however, restricted to small systems [**Publication 5: A Review on Combination of ab Initio Molecular Dynamics and NMR Parameters Calculations**]. Therefore, this method was not used in this work.

In all cases a proper representation of the entropy term is a crucial aspect. Its measurement is dependent on the space sampling. The extended space sampling methods are among others SMD (Steered Molecular Dynamics) and FEP (Free Energy Perturbation calculations). The idea of the former is based on application of the biased coordinates and the free energy of binding (ΔG) is calculated from the non-equilibrium work [146,147]. The principle of the latter is application of the biased paths and ΔG is calculated based on the alchemical transformation [148,149]. In the current work neither SMD nor FEP approach is used. However, this work is a preparation for future application of those methods to analyze both the EDC-ER and EDC-CD complexes. More detailed information on the topic can be found in **Publication 1: Application of Various Molecular**

Modelling Methods in the Study of Estrogens and Xenoestrogens and in the 'Conclusions and perspectives' part at the end of this work.

- **MD-MMGBSA -cyclodextrin complexes**

So far, in many cases in order to prepare a CD structure for further simulations, the geometry optimization has been performed using Molecular Mechanics (MM) methods. Often a special Glycam06 forcefield (adjusted AMBER forcefield) dedicated for carbohydrates has been used [75]. In the works published even a couple years ago it has been often referred to as a 'standard procedure'. The review about the computational methods used for CD-complexes simulations [**Publication 4: Application of Molecular Dynamics Simulations in the Analysis of Cyclodextrin Complexes**] cites at least 25 articles from the recent years where Glycam06 has been applied. However, now, the energy minimization of a CD structure can be handled by the DFT calculations which are much more accurate and therefore have been used in this project.

Alongside with MD the often used method is MMGBSA (MM Generalized Born Surface Area). This approach allows to obtain the free energy of binding (ΔG). Firstly, MD using an explicit solvent model is performed. Secondly, from the last snapshots of MD the solvent molecules are extracted. On these snapshots, MMGBSA calculation in the implicit solvent is conducted. In MMGBSA the entropy term (ΔH) is calculated as a sum of MM-based electrostatics energy term (bonded and non-bonded energy terms) and two solvation related energy terms (calculated in the implicit model) [eq. 14] [150]. In the eq. 14 ΔG_{pol} corresponds to the Generalized Born (GB) approximation of the Poisson-Bolzmann equation, which in turn describes the electrostatic environment of the solute in a solvent containing ions [150]. ΔG_{nonpol} relates to the Solvent Accessible Surface Area (SASA) which is an implicit approach describing the relationship between ΔG and surface area of a solute molecule.

$$\Delta H = \Delta E_{MM} + \Delta G_{pol} + \Delta G_{nonpol} \quad (14)$$

The change between explicit and implicit solvent model which happens before MMGBSA is performed, requires energies' reweighting and several approximations. What is more, for each simulated system several parameters must be arbitrary decided on. All these factors and the fact that the implicit model is less accurate then the explicit one, results in MMGBSA methods being very differently assessed: for some systems they reflect the experimental ΔG very accurately, for others not at all. A number of adjustments has been tried on the MMGBSA model, among others application of the polarizable FF QM/GBSA approach. MMGBSA is still a widely chosen method, especially to calculate ΔΔG in the protein-ligand systems, where the MMGBSA score is used to rank the ligands' binding affinity to the receptor. However, this approach has a better equivalent in form of the FEP calculations. This fact has been known for years but the FEP method is computationally demanding.

When it comes to the MD-MMGBSA calculations, there is not much information on the simulations concerning specifically CD complexes [151-154]. There is one relatively recent (publication year: 2018) example where the computation object is a CD complex with genistein, a natural EDC characterized by a structural similarity to EST. In this work a high level of theory, M06-2X/6-31+G(d,p), is applied to perform calculations on the snapshots extracted from MD [155]. However, as it will be presented later, for the purpose of this work the standard MMGBSA method was applied, without the QM approach after MD run.

Quantum Mechanical approaches

In this work for different purposes, two types of QM computational approaches have been used: semi-empirical methods and Density Functional Theory calculations (DFT).

Semi-empirical methods accuracy is generally considered to be lower than the accuracy of DFT. However, with regards to CD complexes no real comparison between different QM-based approaches has ever been made. Moreover, CD complexes are not small systems and still semi-empirical calculations are often a preferred approach. This has been shown through a thorough literature review **Publication 6: Current Status of Quantum Chemical Studies of Cyclodextrin Host-Guest Complexes**.

DFT approaches provide a high calculational accuracy. However, they are computationally costly, when compared to semi-empirical methods. For years DFT-based methods have been not-correct enough due to the neglection of the dispersion (London) effects [156]. For example, in the condensed matter studies, this was not a major problem in the case of systems characterized by strong electrostatic interactions such as ionic solids, while it was a serious limitation for molecular crystals, where dispersion forces such as van der Waals interactions greatly contribute to the overall binding energy. The most popular method to overcome this problem is the application of "dispersion corrections" (DFT-D), i.e. in the form $C6R^{-6}$ in the DFT formalism [157]. These semiempirical approaches provide the best compromise between the cost of first principles evaluation of the dispersion terms and the need to improve non-bonding interactions in the standard DFT description [158]. Implementation of the dispersion corrections (e.g. D3, TS, MBD) [159] made DFT approach one of the most desirable option for the analysis of small systems. However, it must be stated that application of the empirical dispersion corrections does not increase the accuracy of the results in 100% of cases. Therefore, their application should be tested for each new system.

DFT approach describes the total energy of the system (Et) by the Hohenberg-Kohn-Sham equation [145] [eq. 15]:

$$Et[\rho] = T[\rho] + U[\rho] + Exc[\rho] \quad (15)$$

where T stands for kinetic energy of non-interacting particles, U for classical electrostatic energy due to the Coulombic interactions, Exc for the exchange-correlation energy, ρ is charge density.

A crucial element, on which the accuracy of the DFT methods depends, is the exchange-correlation energy presented as the exchange-correlation functional [145] which can be approximated in several ways: as Local-Density Approximation (LDA), Generalized Gradient Approximation (GGA) e.g. PBE-TS, PBE-SOL or hybrid functionals, e.g. B3LYP, M06-2X. In other words, this energy term is represented as a functional of the electron density ρ.

In the DFT approach, the electronic structure is evaluated on the basis of a potential acting on the electrons in the system. The DFT potential is constructed as the sum of external potentials, determined solely by the structure of the system and an effective potential resulting from interelectronic interactions. All-electron DFT methods treat core and valence electrons in the same way. However, the DFT calculations can be very much simplified and accelerated if electrons are divided in two groups: valence electrons and inner core electrons. In most cases, the electrons of the inner shells (core electrons) are tightly bound and are not involved in the chemical binding. In most organic molecules, binding is solely due to the valence electrons [160]. This separation means that in a large number of cases the atom can be reduced to an ionic core that interacts with the valence electrons. In the pseudopotential approach, widely used for the solid-state DFT calculations, ion cores are considered to be frozen, meaning that the properties of solids are calculated on the assumption that the ion cores are not involved in chemical bonding and therefore they do not change as a result of structural modifications or presence of other atoms. A pseudopotential represents an effective interaction that approximates the potential felt by the valence electrons [161].

Another aspect which must be decided on and which has a huge influence on the calculation results, is choice of a basis set. A basis set is set of basis functions which represent the electronic wave function in form of the algebraic equations what makes them readable for a computer [145]. In the non-periodic DFT calculations the localized basis sets are used.

- **DFT calculations in solid state**

Solid state substances have either amorphous or crystalline character. In order to properly represent the crystalline ones during the calculations their periodicity must be taken into account. This happens when the periodic DFT approach is used. In such case, plane-wave basis sets are usually applied. They are commonly used in calculations involving three-dimensional periodic boundary conditions. The main advantage of a plane-wave basis sets is that it is guaranteed to converge in a smooth, monotonic manner to the target wavefunction [162]. Additional benefit resulting from the application of plane-wave basis set is the introduction of periodic conditions to the studied system. For accurate and computationally feasible approximation of a large system such as macroscopic

crystals, periodic boundary conditions are often applied using crystal unit cells as simulation boxes. During the computations only the properties of the original unit cell need to be calculated and then propagated in the chosen dimension. Additionally, the main advantage of imposing periodic boundary conditions relates to Bloch's theorem, which states that in a periodic system each electronic wavefunction can be written as a product of a cell-periodic part and a wavelike part. The cell periodic part can then be expanded using a basis set consisting of a discrete set of plane waves whose wave vectors are reciprocal lattice vectors of the crystal. Therefore, each electronic function can be written as a sum of plane waves [163]. Periodic DFT calculations are used among others in the procedure of the Crystal Structure Prediction, to explain the crystallization and solvation processes, analyze polymorphs, verify the experimentally obtained structures etc. A detailed description and numerous examples on the topic can be found in **Publication 3: Periodic DFT Calculations-Review of Applications in the Pharmaceutical Sciences** and **Publication 10: Pharmaceutical Hydrates Analysis—Overview of Methods and Recent Advances**.

A particular application of periodic DFT approach is calculation of NMR properties. NMR data is of high importance for the description of CD complexes. Only small number of these complexes has crystalline form and only for few of them it is possible to obtain a crystal of a size suitable for single-crystal X-ray measurements. Therefore, the ssNMR (solid state Nuclear Magnetic Resonance) technique is often the best choice to analyze the inner structure of the complex. Moreover, ssNMR technique delivers information unobtainable by any other experimental technique. In particular, ssNMR can provide the information on orientation of the guest molecule inside the cavity and the complex stability in the solid state. It also enables the quantitative analysis of the phases, especially the complexed and non-complexed guest molecules. In addition, this technique allows for the study of the local molecular dynamics of a guest molecules and the nature of intermolecular interactions between the host and the guest. The thorough description of the topic including numerous examples can be found in **Publication 7: A Review of Applications of Solid-State Nuclear Magnetic Resonance (ssNMR) for the Analysis of Cyclodextrin-Including Systems**.

Already for over a decade the computation of NMR shielding tensors is performed using the Gauge Including Projector Augmented Wave Density Functional Theory (GIPAW) method of Pickard et al. [164] and not by previously used Gauge Invariant Atomic Orbitals (GIAO) [165]. To compare the theoretical and experimental data, the calculated chemical shielding constants (σiso) are converted to chemical shifts (δiso) using the following equation:

$$\delta iso = (\sigma Gly + \delta Gly) — \sigma iso \quad (16)$$

where σGly and δGly stand for the shielding constant and the experimental chemical shift, respectively, of the glycine carbonyl carbon atom (176.50 ppm), if glycine is used as external standard.

The accuracy of combined ssNMR and DFT NMR (GIPAW calculation) is already confirmed so well that such an approach is used as a verifying tool for other techniques, like X-ray photoelectron spectroscopy (XPS) [166].

Combined ssNMR and DFT NMR approach can be especially helpful in the structural analysis in case of a huge disorder within the crystal structure. The GIPAW NMR calculations facilitate peak assignment in the ^{13}C CP MAS NMR spectra. Such approach has been used in this work and has been described in details in **Publication 8: 17-β-Estradiol—β-Cyclodextrin Complex as Solid: Synthesis, Structural and Physicochemical Characterization**.

- **DFT calculations in solution**

In the QM molecular modelling approaches, solvent is presented as a continuum using implicit models. There are two main types of implicit solvent models: Polarizable Continuum Model (PCM) and SMD (Solvation Model Density). In both cases, the structureless polarizable medium is characterized mainly by its dielectric constant ε. PCM is the most often used solvent model in the computational analysis of CD complexes and one of the most used continuum models in general. It defines the molecular free energy as a sum of electrostatic (es) and the dispersion-repulsion (dr) contributions to the free energy, and the cavitation energy (cav) [167] [eq. 17]:

$$Gsol = Ges + Gdr + Gcav \quad (17)$$

In order to calculate the Gcav, the surface of the van der Waals sphere is used. Van der Waals sphere is defined as a function of atom type, connectivity, overall charge of the molecule, and the number of attached hydrogen atoms. To obtain Gdr, the solvent accessible surface is used. Ges is obtained thanks to use of an approximate version of the solvent excluding surface constructed through scaling all radii by a constant factor (e.g. 1.2 for water) and then adding some more spheres not centered on atoms in order to arrive at a somewhat smoother surface [Fig. 8] [167].

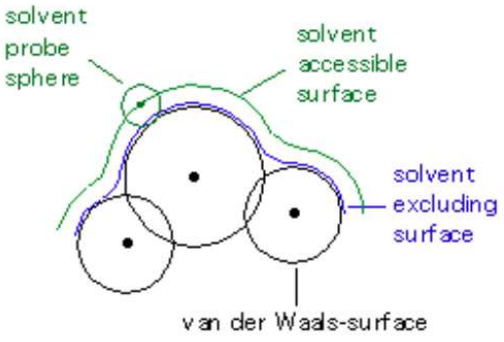


Fig. 8 Graphical representation of the PCM solvent model, description in the main text. Adapted from [167] under the CC BY 4.0. licence.

A different approach is presented by SMD. This model defines the free energy of solvation via two components: the one is electrostatic contribution arising from the self-consistent reaction field, the other comes from the short-range interactions between the solute and solvent molecules [168].

Results

For the first time the chosen EDCs (estradiol, progesterone and bisphenol A) and cyclodextrin have been subjected to the AMOEBA FF parametrization. The procedure was performed using Tinker software. The molecular information about the chosen molecules obtained after the parametrization process was compared with the data from the QM approaches and also using the classical CHARMM FF (NAMD software). The compared data stayed in a good agreement. 10-ns Molecular Dynamics simulation of EST with ERα was performed using Tinker-HP. The simulation was stable. Detailed information about methods and the results are presented in **Publication 2: Polarizable models for selected Endocrine Disrupting Chemicals and their hosts.**

For the first time, the crystal structure of the estradiol and β-cyclodextrin complex has been determined. Different approaches have been tested in order to obtain both crystalline and amorphous system. The complex has been analyzed using SCXRD, PXRD (powder X-ray diffraction), ^{13}C CP MAS ssNMR, FT-IR (Fourier transform infrared spectroscopy), TGA (thermogravimetric analysis), DSC (differential scanning calorimetry), Cryo-SEM experimental techniques as well as molecular modelling approaches: periodic DFT calculations and NMR parameters calculation (CASTEP software by BIOVIA). Detailed information about methods and the results are presented in **Publication 8: 17-β-Estradiol—β-Cyclodextrin Complex as Solid: Synthesis, Structural and Physicochemical Characterization**.

The EST-βCD complex has been also analyzed in the aqueous solution. Application of HRMS (high-resolution mass spectrometry) experimental technique allowed for the first time to thoroughly examine the structure of the complex and define its molar ratio as 1:2 (EST:βCD). Usage of the phase solubility phase studies delivered value of the complex stability constant what in turn enabled to obtain the experimental ΔG Gibbs free energy of the EST-βCD complex.
Moreover, the analyzed system was subjected to DFT and semi-empirical computational approaches (Gaussian16 software). The benchmark method was used to describe the influence of different computational QM-based parameters (B3LYP vs M06-2X functional / PM6 vs PM7 semi-empirical approaches; PCM / SMD water models / in vacuo; presence / absence of D3 dispersion correction) on the results concerning energy and thermodynamic properties. The parameters have been chosen based on the literature review **Publication 6: Current Status of Quantum Chemical Studies of Cyclodextrin Host-Guest Complexes**. At the end, Molecular Dynamics simulation and MMGBSA calculations were performed (AMBER software) to analyze the molar ratio and stability of the complex. Detailed information about methods and the results are presented in **Publication 9: 17-β-Estradiol—β-Cyclodextrin Complex as an aqueous solution: Structural and Physicochemical Characterization supported by MM and QM calculations**.

Conclusions and perspectives

The tasks defined at the beginning of this work have been successfully completed. For the first time, estradiol + β-cyclodextrin complex has been determined using both experimental and computational methods. The complex has been also analyzed in the aqueous solution using both experimental methods and a good variety of computational approaches which were compared to each other. For the first time the molar ratio of the complex has been definitely determined.

Estradiol, progesterone, bisphenol A and cyclodextrin have been successfully parametrized in the polarizable AMOEBA FF. Using AMOEBA FF, the estradiol + estrogen receptor system has been analyzed and the obtained simulation was stable.

This work constitutes a prelude to a complex analysis of the EDC-receptors and EDC-cyclodextrins systems what would be followed by formation of general guidelines on molecular modelling regarding such systems.

In the future, in the first place, progesterone + progesterone receptor and bisphenol A + estrogen receptor simulations using AMOEBA FF will be performed. Next steroidal hormones and selected small EDCs, like phthalates and polychlorinated biphenyls will follow. There is a need for further parametrization of such molecules and their simulation with the respective receptors. The former element will be easier now due to the already obtained steroidal fused rings parameters. These studies will not only deliver information on the applicability of polarizable force fields but will also contribute to the pre-experimental detection of possible EDCs and description of their mode of action.

With regards to the complexation between the low water solubility hormones and cyclodextrins, the list of both guests and hosts should be extended.

Potential objects of such studies would be steroidal pharmaceuticals such as progesterone, hydrocortisone, prednisolone, dexamethasone, testosterone etc. [Fig. 9].

First of all, till now structures of none of those hormones encapsulated in cyclodextrins have yet been identified. Secondly, similarly as in case of estradiol+β-cyclodextrin case, the stoichiometry of such complexes with β-cyclodextrin is not decisively determined, as the literature shows non-coherent data. So far, as a result of my additional research, two complexes: between progesterone and βCD as well as

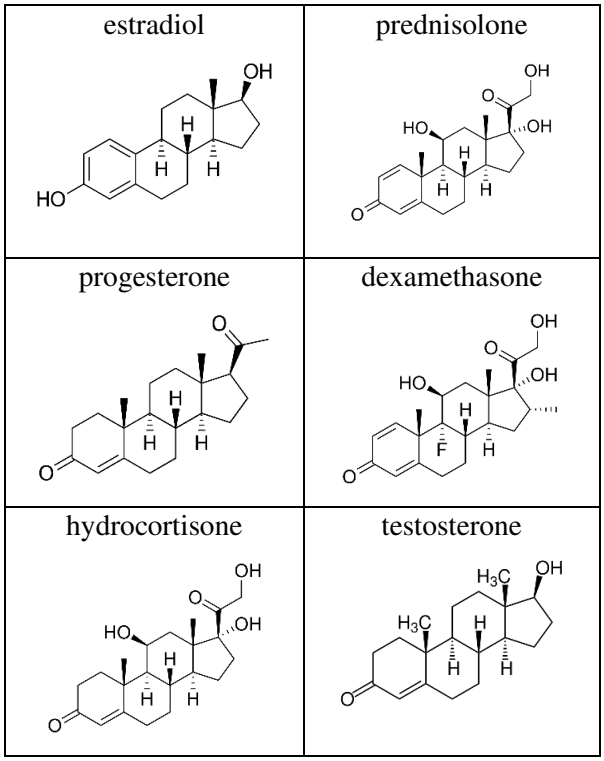


Fig. 9 Structures of selected EDCs, steroidal hormones of a pharmaceutical application.

between hydrocortisone and βCD have been already experimentally determined and analyzed using both experimental and computational methods. At the moment of writing this thesis, the results are not published yet.

The next step would be extension to other widely applied cyclodextrins like 2-HP-βCD. Additionally, the release of guests from the CD-complexes should be measured. Such a study has already been carried out for the estradiol+β-cyclodextrin complex as an addition to my PhD project. As it is described in the original publication attached to this thesis, two forms of this complex have been obtained: amorphous and crystalline. The first objective of the performed release study was to confirm that the solubility of the amorphous complex is higher, hence the release of the estradiol should be higher, as well. The second objective was to obtain the information how much the encapsulation in a cyclodextrin enhances the solubility of estradiol. Release study was performed in HCl solution of pH=2 according to the dissolution test for solid dosage forms as described in Ph. Eur. Monographs 2.9.3 [169] and 5.17.1. [170]. The paddle method was used.

The results confirmed that, as In the majority of cases, the solubility of the amorphous complex was higher than the crystalline one [Fig. 10]. However, the increase of the estradiol solubility after the encapsulation with cyclodextrin was very low [Fig. 10]. Those results have not been published yet, however the efforts to analyze different CD+steroidal pharmaceuticals may bring more welcomed results.

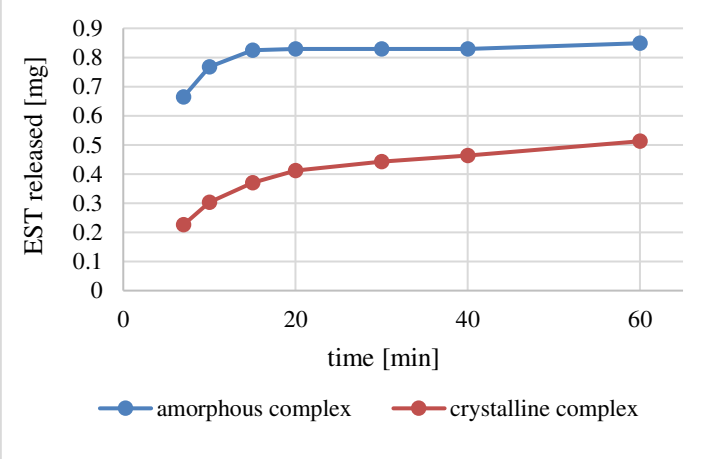


Fig. 10 EST release from amorphous and crystalline complex with β-cyclodextrin, unpublished results of PhD candidate.

Another group of EDCs which complexation with CDs might be useful are pesticides. The research in this topic has been already started. Thanks to the cooperation with the Agricultural University of Athens in Greece, a few complexes between CD and different derivatives of chlorophenoxyacetic acid have been already obtained during duration of my PhD project. The complexes have been analyzed with application of the molecular modelling approach. At the moment of writing this thesis, the results are not published yet.

All the above mentioned complexes should be analyzed using the in silico methods which were defined in this study as the most compatible with the experimental results. This will allow to create a good dataset of the results and confirm which methods are the most effective to predict structure, stoichiometry, stability and interactions within the CD-EDC complexes.

However, not only the already mentioned techniques should be applied. There are two MM-based extended space sampling methods which might be of an interest both for the CD-EDC and receptor-EDC complexes. Those methods are Steered Molecular Dynamics (SMD) and FEP (Free Energy Perturbation) calculations.

SMD applies an external steering force and in this way allows to move a ligand along a selected pathway (for example into and out of a host molecule). The moving is scheduled to stop at the given host-guest distances (called 'windows') in which in the equilibrated state, MD simulation is performed. The pulling velocity is applied to the selected ligand's atom [146,147]. SMD results have form of diagrams of the free binding energy vs host-guest distance, one for each window. WHAM (Weighted Histogram Analysis Method) [171] is used to connect these diagrams and arrive at one Potential of Mean Force profile (PMF) corresponding to the whole pulling process. Out of PMF, the overall ΔG is extracted [171].

SMD calculations deliver mechanistic information on the host-guest binding. In contrast to a binding site of a protein, each CD offers two entering modes: via its wider or narrower rim. This is well illustrated in the article from 2008 about β-CD and progesterone binding [172].
More than a decade ago, SMD has been checked for CD-complexes. One of the last articles using SMD for CD back then, in 2008 claimed that 'the energy analysis was in good agreement with the experimental results' (β-CD-progesterone complex) [172]. However, the CD input structure at the time could be optimized solely with the MM-based methods because the QM geometry optimization of such objects requires much more calculating power, the science lacked at the time. Though, the results obtained in 2008 may not be accurate. Now, it is possible to use the DFT-based methods for that purpose what means that the obtained results should be closer to the experimental data. In other words, SMD could be a good technique for the CD-complexes but it needs to be revisited with the new computer capabilities at hand.

Though, there are still just a few articles published on the subject. For the search 'cyclodextrin SMD' without any search constraints, the Web of Science database shows less than ten results, but interestingly, the used guests are similar to estradiol, for example pinostrobin in 2018 [173]. The newest article in the topic published in 2022 applied SMD for levodopa-CD complex [174]. All the cited studies omit description of a vital aspect which is geometry optimization. And it is already a well-known fact that the geometry of the initial structure has a huge impact on the SMD results. This is why the SMD method can be applicable for the CD-complexes only when the initial geometry is optimized with the newest QM-based approaches about which the benchmarking tests have been described in this thesis.

The next extended space sampling approach is FEP. It allows to calculate the difference of ΔG (ΔΔG) between two similar systems. This method is used for instance to calculate differences in

ligand binding to a wild and mutated protein or to obtain Potential of Mean Force profile for systems which differ in chemical structure, for example comparison of a couple of similar ligands binding to the same binding site [148,149].

In FEP, thermodynamic cycles of non-existing intermediate states are created [175] [Fig. 11]. At each state, after obtaining an equilibrium, the MD simulation is performed. Movement from one intermediate state to the other is regulated by the coupling parameter λ [176].

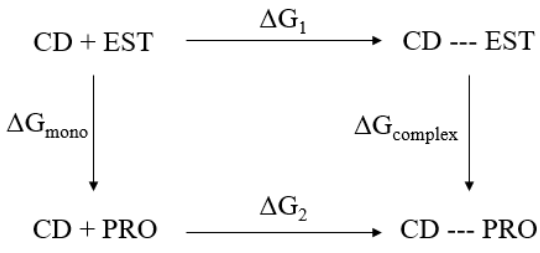


Fig. 11 $\Delta\Delta G = \Delta G_2 - \Delta G_1 = \Delta G_{complex} - \Delta G_{mono}$

Similarly to the SMD case, also FEP application for the CD systems is rare. Two articles refer to a double complexation of a ligand with a CD (imipramine [177], amphotericin B [178]). With regards to this thesis, the more interesting example is the FEP study for progesterone, testosterone and hydrocortisone [179] which delivers some concrete calculation parameters. However, this data has been published in 2009, so surely it needs to be revisited, taking into account even just the increase of the computational power which happened in the last decade. The last, and the most recent (2016), CD-including FEP study corresponds to S-β-CD complexed with either uranyl or uranyl ion [180]. This publication can be also a source of some basic calculational details but its objects are far different than EDCs.

The same methods, SMD and FEP, could be used also for the EDC-receptor systems. And both in CD-EDC and receptor-EDC cases, additive and polarizable force field (AMOEBA FF) should be used. The compilation of all above mentioned methods and objects, would deliver a complete view on the molecular modelling possibilities and challenges regarding the computational analysis of the interactions between EDCs and host molecules.

References

1. EC. Corrigendum to Commission Regulation (EU) 2018/605 of 19 April 2018 Amending Annex II to Regulation (EC) No 1107/2009 by Setting Out Scientific Criteria for the Determination of Endocrine Disrupting Properties. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R0605&rid=210
2. Darbre, P.D. The history of endocrine-disrupting chemicals. Current Opinion in Endocrine and Metabolic Research 2019, 7, 26-33.
3. Singleton D.W., David W.S. Xenoestrogen exposure and mechanisms of endocrine disruption. Front. Biosci. 2003, 8, 110–118.
4. Safe S., Khan S., Wu F., Li X. Chemical-induced estrogenicity in Xenoestrogens and phytoestrogens as SERMs and implications for risk assessment (Chapter 65). In Veterinary Toxicology Basic and Clinical Principles; Elsevier: Amsterdam, The Netherlands, 2007, pp. 811–822.
5. Frizzell C., Ndossi D., Verhaegen S., Dahl E., Eriksen G., Sørlie M., Ropstad E., Muller M., Elliott,C.T., Connolly L. Endocrine disrupting effects of zearalenone, alpha- and beta-zearalenol at the level of nuclear receptor binding and steroidogenesis. Toxicology Letters 2011, 206, 210–217.
6. Caldwell D.J., Mastrocco F., Nowak E., Johnston J., Yekel H., Pfeiffer D., Hoyt M., DuPlessie B.M., Anderson P.D. An assessment of potential exposure and risk from estrogens in drinking water. Environ Health Perspect. 2010, 118, 338-44.
7. Wee S.Y., Aris A.Z., Yusoff F.M., Praveena S.M. Occurrence of multiclass endocrine disrupting compounds in a drinking water supply system and associated risks. Sci Rep 2020, 10, 17755.
8. Aris A.Z., Shamsuddin A.S., Praveena S.M. Occurrence of 17α-ethynylestradiol (EE2) in the environment and efect on exposed biota: A review. Environ. Int. 2014, 69, 104–119.
9. Gonsioroski A., Mourikes V.E., Flaws J.A. Endocrine disruptors in water and their effects on the reproductive system. Int. J. Mol. Sci. 2020, 21, 1929.
10. Caldwell D.J., Mastrocco F., Anderson P.D., Länge R., Sumpter J.P. Predicted-no-effect concentrations for the steroid estrogens estrone, 17β-estradiol, estriol, and 17α-ethinylestradiol. Environ Toxicol Chem. 2012, 31, 1396-406.
11. Laurenson J.P., Bloom R.A., Page S., Sadrieh N. Ethinyl estradiol and other human pharmaceutical estrogens in the aquatic environment: a review of recent risk assessment data. AAPS J. 2014, 16, 299-310.
12. Scientific Committee on Health and Environmental Risks Opinion on ‘’Chemicals and the water framework directive: draft environmental quality standards’’ Ethinylestradiol, SCHER adopted this opinion at its 12th plenary on 30 March 2011 https://ec.europa.eu/health/scientific_committees/environmental_risks/docs/scher_o_146.pdf
13. Scientific Committee on Health and Environmental Risks Opinion on ‘’Chemicals and the water framework directive: draft environmental quality standards’’ 17β-estradiol, SCHER adopted this opinion at its 12th plenary on 30 March 2011 https://ec.europa.eu/health/scientific_committees/environmental_risks/docs/scher_o_131.pdf

14. The communication of the Commission of 7 November 2018 ‘Towards a comprehensive European Union framework on endocrine disruptors’ and Directive (EU) 2020/2184 of the European Parliament and the Council of 16 December 2020 on the quality of water intended for human consumption. https://eur-lex.europa.eu/eli/dir/2020/2184/oj

15. Guerrero-Gualan D., Valdez-Castillo E., Crisanto-Perrazo T., Toulkeridis T. Methods of removal of hormones in wastewater. Water 2023, 15, 353.

16. Tagliavini M., Weidler P.G., Njel C., Pohl J., Richter D., Böhringer B., Schäfer A.I. Polymer-based spherical activated carbon - ultrafiltration (UF-PBSAC) for the adsorption of steroid hormones from water: Material characteristics and process configuration. Water Res. 2020, 185, 116249.

17. Prokić D., Vukčević M., Kalijadis A., Maletić M., Babić B., Đurkić T. Removal of estrone, 17β-Estradiol, and 17α-ethinylestradiol from water by adsorption onto chemically modified activated carbon cloths. Fibers Polym 2020, 21, 2263–2274.

18. Agenda of the EU SCoPAFF Committee Standing Committee on Plants, Animals, Food and Feed Section Phytopharmaceuticals – Legislation 24 - 25 May 2023, point C.06, https://food.ec.europa.eu/system/files/2023-05/sc_phyto_20230524_ppl_agenda.pdf

19. Poulson B.G., Alsulami Q.A., Sharfalddin A., El Agammy E.F., Mouffouk F., Emwas A.-H., Jaremko L., Jaremko M. Cyclodextrins: Structural, Chemical, and Physical Properties, and Applications. Polysaccharides 2022, 3, 1.

20. Saokham P., Muankaew C., Jansook P., Loftsson T. Solubility of cyclodextrins and drug/cyclodextrin complexes. Molecules 2018, 23, 1161.

21. Thomas M.P., Potter B.V. The structural biology of oestrogen metabolism. J. Steroid Biochem. Mol. Biol. 2013, 137, 27–49.

22. Grandi G.;\, Napolitano A., Cagnacci A. Metabolic impact of combined hormonal contraceptives containing estradiol. Expert Opin. Drug Metab. Toxicol. 2016, 12, 779–787.

23. Santoro N., Epperson C.N., Mathews S.B. Menopausal Symptoms and Their Management. Endocrinol. Metab. Clin. N. Am. 2015, 44, 497–515.

24. Daulbayev C., Kaidar B., Sultanov F., Bakbolat B., Smagulova G., Mansurov Z. The recent progress in pitch derived carbon fibers applications. A Review. S. Afr. J. Chem. Eng. 2021, 38, 9–20.

25. Guo A., Gong X., He J., Guo Y., Ning L., Chen X., Xu J., Guo Y., Wang, H. Study on Co-crystals of Estradiol. Her. Med. 2021, 40, 1716–1723.

26. Wang J.R., Wang X., Yang Y., Chen X., Mei X. Solid-state characterization of 17β-estradiol co-crystals presenting improved dissolution and bioavailability. CrystEngComm 2016, 18, 3498–3505.

27. Ning L., Gong X., Li P., Chen X., Wang H., Xu J. Measurement and correlation of the solubility of estradiol and estradiol-urea co-crystal in fourteen pure solvents at temperatures from 273.15 K to 318.15 K. J. Mol. Liq. 2020, 304, 112599.

28. Du X., Li Y., Xia Y.L., Ai S.M., Liang J., Sang P., Ji X-L., Liu S-Q. Insights into protein-ligand interactions: mechanisms, models, and methods. Int J Mol Sci. 2016, 17, 144.

29. Turza A., Borodi G., Muresan-Pop M., Ulici A. Polymorphism and β-cyclodextrin complexation of methyldrostanolone. J. of Mol. Structure 2022, 1250, 131852.
30. Mahmoudi F., Shahraki M. Encapsulating and decontaminating of sarin by heptakis(2,3,6-tri-O-methyl)-β-cyclodextrin: MD simulations and QM calculations. Mol. Syst. Des. Eng. 2021, 6, 643-653.
31. Vanommeslaeghe K., Guvench O., MacKerell A.D. Jr. Molecular mechanics. Curr Pharm Des. 2014, 20, 3281-3292.
32. Ren P., Wu C., Ponder J.W. Polarizable atomic multipole-based molecular mechanics for organic molecules. J Chem Theory Comput. 2011, 7, 3143-3161.
33. Zhang C., Bell D., Harger M., Ren P. Polarizable multipole-based force field for aromatic molecules and nucleobases. J Chem Theory Comput. 2017, 13, 66-678.
34. Bradshaw R.T., Dziedzic J., Skylaris C.K., Essex J.W. The role of electrostatics in enzymes: Do biomolecular force fields reflect protein electric fields? J. Chem. Inf. Model. 2020, 60, 3131-3144.
35. Lemkul J.A. Pairwise-additive and polarizable atomistic force fields for molecular dynamics simulations of proteins. Prog Mol Biol Transl Sci. 2020;170:1-71.
36. Rick S.W., Stuart S.J. Potentials and algorithms for incorporating polarizability in computer simulations. Rev Comp Ch. 2002, 18:89–146.
37. Ponder J.W., Case D.A. Force fields for protein simulations. Adv. Protein Chem. 2003, 66, 27–85.
38. Warshel A., Kato M., Pisliakov A.V. Polarizable force fields: history, test cases, and prospects. J. Chem. Theory Comput. 2007, 3, 2034–2045.
39. Lopes P.E.M., Roux B., Mackerell A.D. Molecular modeling and dynamics studies with explicit inclusion of electronic polarizability: theory and applications. Theor Chem Acc. 2009, 124, 11–28.
40. Cieplak P., Dupradeau F.Y., Duan Y., Wang J. Polarization effects in molecular mechanical force fields. J. Phys-Condens Mat. 2009, 21, 333102 1–21.
41. Gresh N., Cisneros G.A., Darden T.A., Piquemal J.P. Anisotropic, polarizable molecular mechanics studies of inter- and intramoecular interactions and ligand-macromolecule complexes. A bottom-up strategy. J Chem Theory Comput. 2007, 3, 1960–1986.
42. Soderhjelm P., Ryde U. How accurate can a force field become? A polarizable multipole model combined with fragment-wise quantum-mechanical calculations. J. Phys. Chem. A. 2009, 113, 617–627.
43. Soderhjelm P., Ohrn A., Ryde U., Karlström G. Accuracy of typical approximations in classical models of intermolecular polarization. J. Chem. Phys. 2008, 128, 014102.
44. Holt A., Bostrom J., Karlström G., Lindh R. A NEMO potential that includes the dipole-quadrupole and quadrupole quadrupole polarizability. J. Comput. Chem. 2010, 31, 1583–1591.
45. Stern H.A., Rittner F., Berne B.J. Friesner R.A. Combined fluctuating charge and polarizable dipole models: application to a five-site water potential function. J. Chem. Phys. 2001, 115, 2237–2251.
46. Lamoureux G., MacKerell A.D., Roux B. A Simple polarizable model of water based on classical drude oscillators. J. Chem. Phys. 2003,119, 5185–5197.
47. Ren P.Y., Ponder J.W. Polarizable atomic multipole water model for molecular mechanics simulation. J. Phys. Chem. B. 2003, 107, 5933–5947.

48. Kaminski G.A., Stern H.A., Berne B.J., Friesner R.A. Development of an accurate and robust polarizable molecular mechanics force field from ab initio quantum chemistry. J. Phys. Chem. A. 2004, 108, 621– 627.

49. Anisimov V.M., Vorobyov I.V., Roux B., Mackerell Jr A.D. Polarizable empirical force field for the primary and secondary alcohol series based on the classical Drude model. J. Chem. Theory Comput. 2007, 3, 1927–1946.

50. Lopes P.E.M., Lamoureux G., Roux B., Mackerell Jr A.D. Polarizable empirical force field for aromatic compounds based on the classical Drude oscillator. J. Phys. Chem. B. 2007, 111, 2873–2885.

51. Harder E., Anisimov V.M., Whitfied T., Mackerell Jr A.D., Roux B. Understanding the dielectric properties of liquid amides from a polarizable force field. J. Phys. Chem B. 2008, 112, 3509–3521.

52. Shi Y., Wu C., Ponder J.W., Ren P. Multipole electrostatics in hydration free energy calculations. J. Comput. Chem. 2011, 32, 967–977.

53. Ren P.Y., Ponder J.W. Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. J. Comput. Chem. 2002, 23, 1497–1506.

54. Jiang J.L., Wu Y.B., Wang Z.X., Wu C. Assessing the performance of popular Quantum Mechanics and Molecular Mechanics methods and revealing the sequence-dependent energetic features using 100 tetrapeptide models. J. Chem. Theory Comput. 2010, 6, 1199–1209.

55. Jiao D., Golubkov P.A., Darden T.A., Ren P. Calculation of protein-ligand binding free energy by using a polarizable potential. P. Natl. Acad. Sci. USA. 2008, 105, 6290–6295.

56. Jiao D., Zhang J.J., Duke R.E., Li G., Schnieders M.J., Ren P. Trypsin-ligand binding free energies from explicit and implicit solvent simulations with polarizable potential. J. Comput. Chem. 2009, 30, 1701–1711.

57. Shi Y., Zhu C.Z., Martin S.F., Ren P. Probing the effect of conformational constraint on phosphorylated ligand binding to an SH2 domain using polarizable force field simulations. J. Phys. Chem. B. 2012, 116, 1716–1727.

58. Zhang, J.; Shi, Y., Ren P. Protein-ligand interactions. Wiley-VCH Verlag GmbH & Co. KGaA; 2012. Polarizable force fields for scoring protein–ligand interactions; p. 99-120.

59. Harder E., MacKerell A.D., Roux B. Many-body polarization effects and the membrane dipole potential. J. Am. Chem. Soc. 2009, 131, 2760–2761.

60. Bauer B.A., Lucas T.R., Meninger D.J., Patel S. Water permeation through DMPC lipid bilayers using polarizable charge equilibration force fields. Chem. Phys. Lett. 2011, 508, 289–294.

61. de Courcy B., Piquemal J.P., Garbay C., Gresh N. Polarizable water molecules in ligand-macromolecule recognition. Impact on the relative affinities of competing pyrrolopyrimidine inhibitors for FAK kinase. J. Am. Chem. Soc. 2010, 132, 3312–3320.

62. Grossfield A., Ren P.Y., Ponder J.W. Ion solvation thermodynamics from simulation with a polarizable force field. J. Am. Chem. Soc. 2003, 125, 15671–15682.

63. Jiao D., King C., Grossfield A., Darden T.A., Ren P. Simulation of Ca2+ and Mg2+ solvation using polarizable atomic multipole potential. J. Phys. Chem. B. 2006, 110, 18553–18559.

64. Yu H.B., Whitfield T.W., Harder E., Lamoureux G., Vorobyov I., Anisimov V.M., Mackerell A.D. Jr, Roux B. Simulating monovalent and divalent ions in aqueous solution using a Drude polarizable force field. J. Chem. Theory Comput. 2010, 6, 774–786.

65. Wu J.C., Piquemal J.P., Chaudret R., Ren P. Polarizable Molecular Dynamics simulation of Zn(II) in water using the AMOEBA force field. J. Chem. Theory Comput. 2010, 6, 2059–2070.

66. Zhang J., Yang W., Piquemal J.P., Ren P. Modeling structural coordination and ligand binding in zinc proteins with a polarizable potential. J. Chem. Theory Comput. 2012, 8, 1314–1324.

67. Patel S., Davis J.E., Bauer B.A. Exploring ion permeation energetics in gramicidin A using polarizable charge equilibration force fields. J. Am. Chem. Soc. 2009, 131, 13890–1.

68. Liu C., Qi R., Wang Q., Piquemal J.P., Ren P. Capturing many-body interactions with classical dipole induction models. J. Chem. Theory Comput. 2017, 13, 2751-2761.

69. Stern H.A., Rittner F., Berne B.J., Friesner R.A. Combined fluctuating charge and polarizable dipole models: Application to a fivesite water potential function. J. Chem. Phys. 2001, 115, 2237.

70. Lemkul J.A., Huang J., Roux B., MacKerell A.D. An empirical polarizable force field based on the classical Drude oscillator model: development history and recent applications. Chem. Rev. 2016, 116, 4983.

71. Shi Y., Xia Z., Zhang J., Best R., Wu C., Ponder J.W., Ren P. The polarizable atomic multipole-based AMOEBA force field for proteins. J. Chem. Theory Comput. 2013, 9, 4046-4063.

72. Zhang C., Lu C., Jing Z., Wu C., Piquemal J.P., Ponder J.W., Ren P. AMOEBA polarizable atomic multipole force field for nucleic acids. J. Chem. Theory Comput. 2018, 14, 2084-2108.

73. Wu J.C., Chattree G., Ren P. Automation of AMOEBA polarizable force field parameterization for small molecules. Theor Chem Acc. 2012, 131, 1138.

74. Rackers J.A., Wang Z., Lu C., Laury M.L., Lagardère L., Schnieders M.J., Piquemal J.P., Ren P., Ponder J.W. Tinker 8: software tools for molecular design. J. of Chem. Theory and Comp. 2018, 14, 5273-5289.

75. Kirschner K.N., Yongye A.B., Tschampel S.M., González-Outeiriño J., Daniels C.R., Foley B.L., Woods R.J. GLYCAM06: A generalizable biomolecular force field. Carbohydrates. J. Comput. Chem., 2008, 29, 622–655.

76. La Merrill M.A., Vandenberg L.N., Smith M.T., Goodson W., Browne P., Patisaul H.B., Guyton K.Z., Kortenkamp A., Cogliano V.J., Woodruff T.J., Rieswijk L., Sone H., Korach K.S., Gore A.C., Zeise L., Zoeller R.T. Consensus on the key characteristics of endocrine-disrupting chemicals as a basis for hazard identification. Nat Rev Endocrinol. 2020,16, 45-57.

77. Ahn C., Jeung E.B. Endocrine-Disrupting Chemicals and disease endpoints. Int J Mol Sci. 2023, 24, 5342.

78. Lee H.R., Jeung E.B., Cho M.H., Kim T.H., Leung P.C., Choi K.C. Molecular mechanism(s) of endocrine-disrupting chemicals and their potent oestrogenicity in diverse cells and tissues that express oestrogen receptors. J Cell Mol Med. 2013, 17, 1-11.

79. Sellami A., Montes M., Lagarde N. Predicting potential Endocrine Disrupting Chemicals binding to Estrogen Receptor α (ERα) using a pipeline combining structure-based and ligand-based in silico methods. Int. J. Mol. Sci. 2021, 22, 2846.

80. Yilmaz B., Terekeci H., Sandal S., Kelestimur F. Endocrine disrupting chemicals: exposure, effects on human health, mechanism of action, models for testing and strategies for prevention. Rev. Endocr. Metab. Disord. 2020, 21, 127–147.

81. Acconcia F., Pallottini V., Marino M. Molecular mechanisms of action of BPA. Dose response. 2015, 13, 1559325815610582.

82. Shanle E.K., Xu W. Endocrine Disrupting Chemicals targeting Estrogen Receptor signaling: identification and mechanisms of action. Chem. Res. Toxicol. 2011, 24, 6–19.

83. La Merrill M.A., Vandenberg L.N., Smith M.T., Goodson W., Browne P., Patisaul H.B., Guyton K.Z., Kortenkamp A., Cogliano V.J., Woodruff T.J., Rieswijk L., Sone H., Korach K.S., Gore A.C., Zeise L., Zoeller R.T. Consensus on the key characteristics of endocrine-disrupting chemicals as a basis for hazard identification. Nat. Rev. Endocrinol. 2020, 16, 45–57.

84. Gupta R.K., Archambeault D.R., Yao H.H.-C. Genetic mouse models for female reproductive toxicology studies. Comprehensive Toxicology, 2010, 561–575.

85. Dechering K., Boersma C., Mosselman S. Estrogen receptors alpha and beta: Two receptors of a kind? Curr. Med. Chem. 2000, 7, 561–76.

86. Ruff M., Gangloff M., Wurtz J., Moras D. Estrogen receptor transcription and transactivation Structure-function relationship in DNA- and ligand-binding domains of estrogen receptors. Breast Cancer Res. 2000, 2, 353–359.

87. Billon-Galés A., Krust A., Fontaine C., Abot A., Flouriot G., Toutain C., Berges H., Gadeau A.-P., Lenfant F., Gourdy P., Chambon P., Arnal J.F. Activation function 2 (AF2) of estrogen receptor-α is required for the atheroprotective action of estradiol but not to accelerate endothelial healing. Proc. Natl. Acad. Sci. USA 2011, 108, 13311–13316.

88. Gao L., Tu Y., Ågren H.A. Eriksson, L. Characterization of agonist binding to His524 in the Estrogen Receptor α ligand binding domain. J. Phys. Chem. B 2012, 116, 4823–4830.

89. Hu G., Wang J. Ligand selectivity of estrogen receptors by a molecular dynamics study. Eur. J. Med. Chem. 2014, 74, 726–735.

90. Gu X. Helix 12 in the human estrogen receptor (hER) is essential for the hER function by overcoming nucleosome repression in yeast. J. Cell. Biochem. 2002, 86, 224–238.

91. Shiau A.K., Barstad D., Loria P.M., Cheng L., Kushner P.J., Agard D.A., Greene G.L. The structural basis of Estrogen Receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. Cell 1998, 95, 927–937.

92. Bennink H.C., Verhoeven C., Zimmerman Y., Visser M., Foidart J.-M., Gemzell-Danielsson K. Clinical effects of the fetal estrogen estetrol in a multiple-rising-dose study in postmenopausal women. Matur. Eur. Menopause J. 2016, 91, 93–100.

93. Bennink F.C., Holinka C.F., Visser M., Bennink H.J.T.C. Maternal and fetal estetrol levels during pregnancy. Climacteric 2008, 11, 69–72.
94. Schreiner W.E. The ovary in Labhart, A. Clinical Endocrinology: Theory and Practice; Springer: Berlin/Heidelberg, Germany, 2012, p. 548.
95. Kuhl H. Pharmacology of estrogens and progestogens: Influence of different routes of administration. Climacteric 2005, 8 (Suppl. 1), 3–63.
96. Blackburn S. Maternal, Fetal, & Neonatal Physiology; Elsevier: Amsterdam, The Netherlands, 2014.
97. Fait T. Menopause hormone therapy: Latest developments and clinical practice. Drugs Context 2019, 8, 1–9.
98. Tofovic S.P., Jackson E.K. Estradiol metabolism: crossroads in pulmonary arterial hypertension. Int. J. Mol. Sci. 2019, 21, 116.
99. Brzozowski A.M., Pike A.C., Dauter Z., Hubbard R.E., Bonn T., Engström O., Ohman L., Greene G.L., Gustafsson J.A., Carlquist M. Molecular basis of agonism and antagonism in the oestrogen receptor. Nature. 1997, 389, 753-8.
100. Kumar R., Zakharov M.N., Khan S.H., Miki R., Jang H., Toraldo G., Singh R., Bhasin S., Jasuja R. The dynamic structure of the estrogen receptor. J. Amino Acids. 2011, 2011, 812540.
101. Yaşar P., Ayaz G., User S.D., Güpür G., Muyan M. Molecular mechanism of estrogen-estrogen receptor signaling. Reprod Med Biol. 2016;16, 4-20.
102. Tanenbaum D.M., Wang Y., Williams S.P., Sigler P.B. Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. Proc. Natl. Acad. Sci. USA. 1998, 95, 5998-6003.
103. Eiler S., Gangloff M., Duclaud S., Moras D., Ruff M. Overexpression, purification, and crystal structure of native ER alpha LBD. Protein Expr Purif. 2001, 22, 165-73.
104. Wärnmark A., Treuter E., Gustafsson J.A., Hubbard R.E., Brzozowski A.M., Pike A.C. Interaction of transcriptional intermediary factor 2 nuclear receptor box peptides with the coactivator binding site of estrogen receptor alpha. J. Biol. Chem. 2002, 277, 21862-8.
105. Sakkiah S., Selvaraj C., Guo W., Liu J., Ge W., Patterson T.A., Hong H. Elucidation of agonist and antagonist dynamic binding patterns in ER-α by integration of molecular docking, Molecular Dynamics simulations and Quantum Mechanical calculations. Int. J. Mol. Sci. 2021, 22, 9371.
106. Yuan M., Chen S., Zeng C., Fan Y., Ge W., Chen W. Estrogenic and non-estrogenic effects of bisphenol A and its action mechanism in the zebrafish model: An overview of the past two decades of work. Environ Int. 2023, 176, 107976.
107. Williams S.P., Sigler P.B. Atomic structure of progesterone complexed with its receptor. Nature. 1998, 393, 392-6.
108. Zheng L., Lin V.C., Mu Y. Exploring flexibility of Progesterone Receptor ligand binding domain using Molecular Dynamics. PLoS One. 2016, 11, e0165824.
109. Wardell S.E., Boonyaratanakornkit V., Adelman J.S., Aronheim A., Edwards D.P. Jun dimerization protein 2 functions as a progesterone receptor N-terminal domain coactivator. Mol. Cell Biol. 2002, 22, 5451-66.

110. Tamrazi A., Carlson K.E., Daniels J.R., Hurth K.M., Katzenellenbogen J.A. Estrogen receptor dimerization: ligand binding regulates dimer affinity and dimer dissociation rate. Mol. Endocrinol. 2002, 16, 2706-19.

111. Azeez J.M., Susmi T.R., Remadevi V., Ravindran V., Sasikumar Sujatha A., Ayswarya R.N.S., Sreeja S. New insights into the functions of progesterone receptor (PR) isoforms and progesterone signaling. Am. J. Cancer Res. 2021, 11, 5214-5232.

112. Arnold S.F., Vorojeikina D.P., Notides A.C. Phosphorylation of tyrosine 537 on the human estrogen receptor is required for binding to an estrogen response element. J Biol Chem. 1995, 270, 30205-12.

113. Castoria G., Giovannelli P., Lombardi M., De Rosa C., Giraldi T., de Falco A., Barone M.V., Abbondanza C., Migliaccio A., Auricchio F. Tyrosine phosphorylation of estradiol receptor by Src regulates its hormone-dependent nuclear export and cell cycle progression in breast cancer cells. Oncogene. 2012, 31, 4868-77.

114. Szejtli J. Introduction and general overview of cyclodextrin chemistry. Chem. Rev. 1998, 98, 1743–1754.

115. Jambhekar S.S., Breen P. Cyclodextrins in pharmaceutical formulations II: Solubilization, binding constant, and complexation efficiency. Drug Discov. Today 2016, 21, 363–368.

116. European Medicines Agency. Available online: https://www.ema.europa.eu/en (accessed on 28 January 2024).

117. U.S. Food & Drug Administration (FDA). Available online: https://www.fda.gov (accessed on 28 January 2024).

118. Pharmaceutical and Medical Devices Agency. Available online: https://www.pmda.go.jp/english/index.html (accessed on 28 January 2024)

119. Gadade D.D., Pekamwar S.S. Cyclodextrin based nanoparticles for drug delivery and theranostics. Adv. Pharm. Bull. 2020, 10, 166-183.

120. Wankar J., Kotla, N.G., Gera S., Rasala S., Pandit A., Rochev, Y.A. Recent Advances in Host–Guest Self-Assembled Cyclodextrin Carriers: Implications for Responsive Drug Delivery and Biomedical Engineering. Adv. Funct. Mat. 2020, 30, 1909049.

121. Gidwani B., Vyas A. A Comprehensive Review on Cyclodextrin-Based Carriers for Delivery of Chemotherapeutic Cytotoxic Anticancer Drugs. Biomed. Res. Int. 2015, 2015, 198268.

122. European Medicines Agency website https://www.ema.europa.eu/en accessed on 21.02.2022

123. European Medicines Agency website https://www.ema.europa.eu/en/cyclodextrins accessed on 21.02.2022

124. Xiao Y., Ng S-H., Tan T.H.Y., Wang Y. Recent development of cyclodextrin chiral stationary phases and their applications in chromatography. J. of Chrom. A 2012, 1269, 52-68.

125. Razak J.L., Doyen H.J., Lunte C.E. Cyclodextrin-modified micellar electrokinetic chromatography for the analysis of Esterom, a topical product consisting of hydrolyzed benzoylecgonine in propylene glycol. Electrophoresis. 2003, 24, 1764-1769.

126. Carvalho L.B.D., Burusco K.K., Jaime C., Venâncio T., Carvalho A.F.S.D., Murgas L.D.S., Pinto L.D.M.A. Complexes between methyltestosterone and β-cyclodextrin for application in aquaculture production. Carbohydr. Polym. 2018, 179, 386–393.
127. Rescifina A., Surdo E., Cardile V., Avola R., Eleonora Graziano A.C., Stancanelli R., Tommasini S., Pistarà V., Ventura C.A. Gemcitabine anticancer activity enhancement by water soluble celecoxib/sulfobutyl ether-β-cyclodextrin inclusion complex. Carbohydr. Polym. 2019, 206, 792–800.
128. Mokhtar M.S., Suliman F.O., Elbashir A.A. Experimental and molecular modeling investigations of inclusion complexes of imazapyr with 2-hydroxypropyl(β/γ) cyclodextrin. J. Mol. Liq. 2018, 262, 504–513.
129. Gieroba B., Kalisz G., Sroka-Bartnicka A., Płazińska A., Płaziński W., Starek, M., Dąbrowska M. Molecular structure of cefuroxime axetil complexes with α-, β-, γ-, and 2-hydroxypropyl-β-cyclodextrins: Molecular simulations and raman spectroscopic and imaging studies. Int. J. Mol. Sci. 2021, 22, 5238.
130. Kellici T.F., Ntountaniotis D., Leonis G., Chatziathanasiadou M., Chatzikonstantinou A.V., Becker-Baldus J., Glaubitz C., Tzakos A.G., Viras K., Chatzigeorgiou P. et al. Investigation of the interactions of silibinin with 2-hydroxypropyl-β-cyclodextrin through biophysical techniques and computational methods. Mol. Pharm. 2015, 12, 954–965.
131. Mayer B.P., Kennedy D.J., Lau E.Y., Valdez C.A. Solution-state structure and affinities of cyclodextrin: Fentanyl complexes by nuclear magnetic resonance spectroscopy and molecular dynamics simulation. J. Phys. Chem. B 2016, 120, 2423–2433.
132. Dang P., Ye R., Meng F., Han Y., Zhou Y., Gong X., Zhou B. Microencapsulation thermodynamics of methylated β-cyclodextrins with bile salt: Enthalpy, entropy, and solvent effect. J. Incl. Phenom. Macrocycl. Chem. 2017, 88, 181–189.
133. Han D., Han Z., Liu L., Wang Y., Xin S., Zhang H., Yu Z. Solubility enhancement of myricetin by inclusion complexation with heptakis-O-(2-hydroxypropyl)-β-cyclodextrin: A joint experimental and theoretical study. Int. J. Mol. Sci. 2020, 21, 766.
134. Melani F., Pasquini B., Caprini C., Gotti R., Orlandini S., Furlanetto S. Combination of capillary electrophoresis, molecular modeling and NMR to study the enantioselective complexation of sulpiride with double cyclodextrin systems. J. Pharm. Biomed. Anal. 2015, 114, 265–271.
135. Huang Z., Xu R., Ge X., Cheng J. Complexation of capsaicin with hydroxypropyl-β-cyclodextrin and its analytical application. Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 2019, 223, 117278.
136. Mahalapbutr P., Charoenwongpaiboon T., Phongern C., Kongtaworn N., Hannongbua S., Rungrotmongkol T. Molecular encapsulation of a key odor-active 2-acetyl-1-pyrroline in aromatic rice with β-cyclodextrin derivatives. J. Mol. Liq. 2021, 337, 116394.
137. Mahalapbutr P., Thitinanthavet K., Kedkham T., Nguyen H., Theu L.t.h., Dokmaisrijan S., Huynh L., Kungwan N., Rungrotmongkol T. A theoretical study on the molecular encapsulation of luteolin and pinocembrin with various derivatized beta-cyclodextrins. J. Mol. Struct. 2019, 1180, 480–490.
138. De Medeiros A.S.A., Zoppi A., Barbosa E.G., Oliveira J.I.N., Fernandes-Pedrosa M.F., Longhi M.R., da Silva-Júnior A.A. Supramolecular aggregates of oligosaccharides with co-solvents in ternary systems for the solubilizing approach of triamcinolone. Carbohydr. Polym. 2016, 151, 1040–1051.

139. Anthony J. Distributed Multipole Analysis of Gaussian wavefunctions GDMA version 2.3.0 https://www-stone.ch.cam.ac.uk/documentation/gdma/manual.pdf

140. Thole B.T. Molecular polarizabilities calculated with a modified dipole interaction. Chem. Phys. 1981, 59, 341-350.

141. Vanommeslaeghe K., MacKerell A.D. Jr. CHARMM additive and polarizable force fields for biophysics and computer-aided drug design. Biochim Biophys Acta. 2015, 1850, 861-871.

142. Case D.A., Cheatham T.E. 3rd, Darden T., Gohlke H., Luo R., Merz K.M. Jr, Onufriev A., Simmerling C., Wang B., Woods R.J. The Amber biomolecular simulation programs. J. Comput. Chem. 2005, 26, 1668-88.

143. Wang J., Wang W., Kollman P.A., Case D.A. Automatic atom type and bond type perception in molecular mechanical calculations. J. Mol. Graph. Model. 2006, 25, 247-60.

144. Wang J., Wolf R.M., Caldwell J.W., Kollman P.A., Case D.A. Development and testing of a general amber force field. J. Comput. Chem. 2004, 25, 1157-74.

145. Leach, A. Molecular Modelling: Principles and Applications. 2001, Publisher: Addison Wesley

146. Noh S.Y., Notman R. Comparison of umbrella sampling and steered molecular dynamics methods for computing free energy profiles of aromatic substrates through phospholipid bilayers. J. Chem. Phys. 2020, 153, 034115.

147. Park S., Khalili-Araghi F., Tajkhorshid E., Schulten K. Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. J. Chem. Phys. 2003, 119, 3559–3566.

148. Jorgensen W.L., Thomas L.L. Perspective on Free-Energy Perturbation calculations for chemical equilibria. J. Chem. Theory Comput. 2008, 4, 869-876.

149. King E., Aitchison E., Li H., Luo R. Recent developments in free energy calculations for drug discovery. Front. Mol. Biosc. 2021, 8, 712085.

150. Forouzesh N., Mishra N. An effective MM/GBSA protocol for absolute binding free energy calculations: a case study on SARS-CoV-2 spike protein and the human ACE2 receptor. Molecules. 2021, 26, 2383.

151. Fateminasab F., Bordbar A.K., Shityakov S., Saboury A.A. Molecular insights into inclusion complex formation between β-and γ-cyclodextrins and rosmarinic acid. J. Mol. Liq. 2020, 314, 113802.

152. Elias C., Katerina F., Athena A., Kostas, B. X-ray crystallography and molecular dynamics studies of the inclusion complexes of geraniol in β-cyclodextrin, heptakis (2,6-di-O-methyl)-β-cyclodextrin and heptakis (2,3,6-tri-O-methyl)-β-cyclodextrin. J. of Mol. Str. 2020, 1202, 127350.

153. Bethanis K., Christoforides E., Tsorteki F., Fourtaka K., Mentzafos D. Structural studies of the inclusion compounds of α-naphthaleneacetic acid in heptakis(2,6-di-O-methyl)-β-Cyclodextrin and heptakis(2,3,6-tri-O-methyl)-β-Cyclodextrin by X-ray crystallography and molecular dynamics. J. Incl. Phenom. Macrocycl. Chem. 2018, 92, 157–171.

154. Shityakov S., Salmas R.E., Durdagi S., Roewer N., Förster C., Broscheit J. Solubility profiles, hydration and desolvation of curcumin complexed with γ-cyclodextrin and hydroxypropyl-γ-cyclodextrin. J. Mol. Struct. 2017, 1134, 91–98.

155. Hanpaibool C., Chakcharoensap T., Arifin, Hijikata Y., Irle S., Wolschann P., Kungwan N., Pongsawasdi P., Ounjai P., Rungrotmongkol, T. Theoretical analysis of orientations and tautomerization of genistein in β-cyclodextrin. J. Mol. Liq. 2018, 265, 16–23.

156. Grimme S. Density functional theory with London dispersion corrections. WIRES Compt. Mol. Sc. 2011, 1, 211-228.

157. Grimme S., Antony J., Schwabe T., Mück-Lichtenfeld C. Density Functional Theory with dispersion corrections for supramolecular structures, aggregates, and complexes of (Bio)rganic molecules. ChemInform. 2007, 5, 741–758.

158. Grimme S. Density functional theory with London dispersion corrections. Wires Comput. Mol. Sci. 2011, 1, 211–228.

159. Grimme S., Hansen A., Brandenburg J.G., Bannwarth C. Dispersion-corrected mean-field electronic structure methods. Chem. Rev. 2016, 116, 5105-54.

160. Foglia N.O., Morzan U.N., Estrin D.A., Scherlis D.A., Lebrero M.C.G. Role of core electrons in Quantum Dynamics using TDDFT. J. Chem. Theory Comput. 2017, 13, 77–85.

161. Burnus T. Ph.D. Thesis. Universität Berlin; Berlin, Germany: 2004. Time-Dependent Electron Localization Function.

162. Lee K., Yu J., Morikawa Y. Comparison of localized basis and plane-wave basis for density-functional calculations of organic molecules on metals. Phys. Rev. B. 2007, 75, 045402.

163. Evarestov R.A. Quantum Chemistry of Solids. Springer; Berlin/Heidelberg, Germany: 2012. Basis sets and pseudopotentials in periodic LCAO calculations, pp. 281–326. (Springer Series in Solid-State Sciences).

164. Pickard C.J., Mauri F. All-electron magnetic response with pseudopotentials: NMR chemical shifts. Phys. Rev. B 2001, 63, 63–77.

165. Szeleszczuk Ł., Pisklak D.M., Zielińska-Pisklak M.A., Wawer I. Effects of structural differences on the NMR chemical shifts in cinnamic acid derivatives: Comparison of GIAO and GIPAW calculations. Chem. Phys. Lett. 2016, 653, 35–41.

166. Stevens J.S., Byard S.J., Muryn C.A., Schroeder S.L.M. Identification of protonation state by XPS, solid-state NMR, and DFT: Characterization of the nature of a new theophylline complex by experimental and computational methods. Phys. Chem. B. 2010, 114, 13961–13969.

167. LMU, Faculty for chemistry and pharmacy, group of prof. Hendrik Zipse https://zipse.cup.uni-muenchen.de/teaching/computational-chemistry-2/topics/the-polarizable-continuum-model-pcm/ accessed on 29.01.024

168. Marenich, A.V., Cramer, C.J., Truhlar, D.G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. J. Phys. Chem. B 2009, 113, 6378–6396.

169. European Pharmacopoeia 10.0 monographs 2.9.3 Dissolution test for solid dosage forms

170. European Pharmacopoeia 7.0 monographs 5.17 Recommendations on methods for dosage forms testing: recommendation on dissolution testing

171. Roux B. The calculation of the potential of mean force using computer simulations. Comp. Phys. Communications. 1995, 91, 275-282.

172. Caballero J., Zamora C., Aguayo D., Yañez C., González-Nilo F.D. Study of the interaction between progesterone and beta-cyclodextrin by electrochemical techniques and steered molecular dynamics. J Phys Chem B. 2008, 112, 10194-201.

173. Kicuntod J., Sangpheak K., Mueller M., Wolschann P., Viernstein H., Yanaka S., Kato K., Chavasiri W., Pongsawasdi P., Kungwan N., Rungrotmongkol T. Theoretical and experimental studies on inclusion complexes of pinostrobin and β-Cyclodextrins. Sci Pharm. 2018, 86, 5.

174. Rezaeisadat M., Salehi N., Bordbar A-K. Inclusion of levodopa into β-Cyclodextrin: A comprehensive computational study. ACS Omega 2021, 6, 23814–23825.

175. Oostenbrink C., van Gunsteren W.F. Free energies of ligand binding for structurally diverse compounds. PNAS 2005, 102, 6750-6754.

176. Pohorille A., Jarzynski C., Chipot C. Good practices in free-energy calculations. J. Phys. Chem. B. 2010, 114, 10235-53.

177. Sun T., Shao X., Cai W. Self-assembly behavior of β-cyclodextrin and imipramine. A Free energy perturbation study. Chem. Phys. 2010, 371, 84-90.

178. He J., Chipot C., Shao X., Cai W. Cooperative recruitment of amphotericin B mediated by a cyclodextrin dimer. J. Phys. Chem. C 2014, 118, 41, 24173–24180.

179. Cai W., Sun T., Liu P., Chipot C., Shao X. Inclusion mechanism of steroid drugs into beta-cyclodextrins. Insights from free energy calculations. J. Phys. Chem. B. 2009, 113, 7836-43.

180. Li L., Zhang Y., Li X., Shen S., Huang H., Bai Y., Liu H. Study on the interaction of uranyl with sulfated beta-cyclodextrin by affinity capillary electrophoresis and molecular dynamics simulation. Electrophoresis. 2016, 37, 2567-2573.

.

PhD project publications

I Publications which are the core of the PhD project (basis for the defence)

a) Original research
   - Mazurek, A.H., Thirion V., Szeleszczuk Ł., Piquemal J.-P., Clavaguera C., Simonson T. Polarizable models for selected Endocrine Disrupting Chemicals and their hosts. J. of Comp. Chem. SUBMITTED (in this document named as Publication 2)
   - Mazurek A.H., Szeleszczuk Ł., Bethanis K., Christoforides E., Dudek M.K., Zielińska-Pisklak M., Pisklak D.M. 17-β-Estradiol—β-Cyclodextrin Complex as Solid: Synthesis, Structural and Physicochemical Characterization. Molecules. 2023, 28, 3747. Doi: 10.3390/molecules28093747 (in this document named as Publication 8)
   - Mazurek A.H., Szeleszczuk Ł., Bethanis K., Christoforides E., Dudek M.K., Wielgus E., Pisklak D.M 17-β-Estradiol—β-Cyclodextrin Complex as an aqueous solution: Structural and Physicochemical Characterization supported by MM and QM calculations. J. of Mol. Structure 2024, 1313, 138710. Doi: 10.1016/j.molstruc.2024.138710 (in this document named as Publication 9)

b) Review articles
   - Mazurek A.H., Szeleszczuk Ł., Simonson T., Pisklak D.M. Application of Various Molecular Modelling Methods in the Study of Estrogens and Xenoestrogens. IJMS 2020, 21, 6411. DOI: 10.3390/ijms21176411 (in this document named as **Publication 1**)
   - Mazurek A.H., Szeleszczuk Ł. Current Status of Quantum Chemical Studies of Cyclodextrin Host–Guest Complexes. Molecules. 2022, 27, 3874. DOI: 10.3390/molecules27123874 (in this document named as **Publication 6**)

- For the above listed publications Polish Ministry of Science and Higher Education points is equal to 490 pkt
- For the above listed publications IF is equal to 18.6

II Additional publications

- Mazurek A.H., Szeleszczuk Ł., Pisklak D.M. Periodic DFT Calculations—Review of Applications in the Pharmaceutical Sciences. Pharmaceutics. 2020, 12, 415. DOI: 10.3390/pharmaceutics12050415 (in this document named as **Publication 3**)
- Mazurek A.H., Szeleszczuk Ł., Gubica T. Application of Molecular Dynamics Simulations in the Analysis of Cyclodextrin Complexes. IJMS. 2021, 22, 9422. DOI: 10.3390/ijms22179422 (in this document named as **Publication 4**)
- Mazurek A.H., Szeleszczuk Ł., Pisklak D.M. A Review on Combination of Ab Initio Molecular Dynamics and NMR Parameters Calculations. IJMS. 2021, 22, 4378. DOI: 10.3390/ijms22094378 (in this document named as **Publication 5**)
- Mazurek A.H., Szeleszczuk Ł. A Review of Applications of Solid-State Nuclear Magnetic Resonance (ssNMR) for the Analysis of Cyclodextrin-Including Systems. IJMS. 2023, 24, 3648. DOI: 10.3390/ijms24043648 (in this document named as **Publication 7**)
- Szeleszczuk Ł., Mazurek A.H., Milcarz K., Napiórkowska E., Pisklak D.M. Can We Predict the Isosymmetric Phase Transition? Application of DFT Calculations to Study the Pressure Induced Transformation of Chlorothiazide. IJMS. 2021, 22, 10100. DOI: 10.3390/ijms221810100
- Zielińska A., Mazurek A., Siudem P., Kowalska V., Paradowska K. Qualitative and quantitative analysis of energy drinks using 1H NMR and HPLC methods. J. Pharm. Biomed. Anal. 2022, 213, 114682. doi: 10.1016/j.jpba.2022.114682.
- Jurczak E., Mazurek A.H., Szeleszczuk Ł., Pisklak D.M., Zielińska-Pisklak M. Pharmaceutical Hydrates Analysis-Overview of Methods and Recent Advances. Pharmaceutics. 2020, 12, 959. doi: 10.3390/pharmaceutics12100959 (in this document named as **Publication 10**)

- Overall Polish Ministry of Science and Higher Education points is equal to 1428 pkt
- Overall IF is equal to 59.802

Appendix

The following appendix is composed of publications which are the basis for the defence. Those are:

1) Polarizable models for selected Endocrine Disrupting Chemicals and their hosts. (proof of submission)
2) 17-β-Estradiol—β-Cyclodextrin Complex as Solid: Synthesis, Structural and Physicochemical Characterization.
3) 17-β-Estradiol—β-Cyclodextrin Complex as an aqueous solution: Structural and Physicochemical Characterization supported by MM and QM calculations.
4) Application of Various Molecular Modelling Methods in the Study of Estrogens and Xenoestrogens.
5) Current Status of Quantum Chemical Studies of Cyclodextrin Host–Guest Complexes. Molecules.

Journal of Molecular Structure xxx (xxxx) 138710

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: www.elsevier.com/locate/molstruc

17-β-Estradiol—β-Cyclodextrin complex as an aqueous solution: Structural and physicochemical characterization supported by MM and QM calculations

Anna Helena Mazurek [a, b], Łukasz Szeleszczuk [a, *], Kostas Bethanis [c], Elias Christoforides [c], Marta Katarzyna Dudek [d], Ewelina Wielgus [d], Dariusz Maciej Pisklak [a]

[a] Department of Organic and Physical Chemistry, Faculty of Pharmacy, Medical University of Warsaw, Banacha 1 Str., 02-093 Warsaw, Poland
[b] Doctoral School, Medical University of Warsaw, Żwirki i Wigury 81 Str., 02-093 Warsaw, Poland
[c] Laboratory of Physics, Department of Biotechnology, Agricultural University of Athens, 11855 Athens, Greece
[d] Structural Studies Department, Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, Sienkiewicza 112 Str., 90-363 Łódź, Poland

ARTICLE INFO

Keywords:
Estradiol
Cyclodextrin
DFT calculations
Complex stability constant

ABSTRACT

17-β-estradiol (EST) is an Active Pharmaceutical Ingredient characterized by a low water solubility. Complexation with β-cyclodextrin (βCD) enhances its bioavailability, hence such complex is an interesting research object from pharmaceutical point of view. However, basic facts like description of complex's structure and definition of its molar ratio, were debatable already for decades. This work for the first time justifies the EST:βCD molar ratio as 1:2 using the HRMS (high-resolution mass spectrometry) and phase solubility studies. The latter are used to define complex stability constant, as well. The structure and stability is analyzed using a variety of computational approaches: Quantum Mechanics (QM) based methods (DFT, semiempirical approaches) and MD/MMGBSA approach. In case of the QM, for the first time in the computational analysis of cyclodextrin complexes, a thorough benchmarking test is presented. Different computational parameters (solvent model, presence/absence of dispersion correction etc.) are used. Obtained results are compared with the experimental data.

1. Introduction

Cyclodextrins (CDs) are oligosaccharides of a donut-like structure, which enables them to form inclusion complexes with non-polar substances. This characteristic is used by the pharmaceutical industry but at the same time it often poses a non-trivial questions about the structure and stability of the created complexes.

More precisely, CDs are cyclic structures composed of glucose subunits, joined by α−1,4 glycosidic bonds [1]. Because of the cyclic character of CDs, various chemical compounds can enter CD's void and this way inclusion complexes are created. Depending on the size of a chemical guest, different types of CDs are preferred. However, the most common one is a medium size beta-CD (βCD) which consists of 7 glucose units [1,2] (Fig. 1).

Due to the presence of hydroxyl groups, the external fragments of CDs are polar. When a non-polar substance enters the molecular hole of a CD, the formed host–guest complex is polar and more water soluble than a separate non-complexed guest molecule [2,3]. Therefore, knowing also that CDs are non-toxic for a human organism [4], CDs are commonly used in the pharmaceutical industry in order to increase the solubility of a complexed Active Pharmaceutical Ingredient (API) or protect it from external factors like light, humidity or heat [5-7]. An important API group characterized by a poor solubility in water are hormonal steroids. Encapsulation in CDs enhance their solubility in water and as a result also their bioavailability.

An example of the steroid hormones used as a medication is estradiol (EST) (Fig. 1). EST belongs to the estrogens group and is the most potent estrogen naturally produced by a human body [8]. Therefore, the first complexes between EST and a well-known βCD have been obtained already decades ago [9]. However, the structural analysis of such complexes happens not to be as straightforward as one would assume.

In 1997 in the article entitled ''Fluorometric Determination of Association Constants of Three Estrogens with Cyclodextrins'' the EST-βCD complex molar ratio has been defined as 1:1 [10]. This information has been used as a reference for instance in the following articles [11,12] by

* Corresponding author.
E-mail address: lukasz.szeleszczuk@wum.edu.pl (Ł. Szeleszczuk).

https://doi.org/10.1016/j.molstruc.2024.138710
Received 19 February 2024; Received in revised form 3 May 2024; Accepted 21 May 2024
0022-2860/

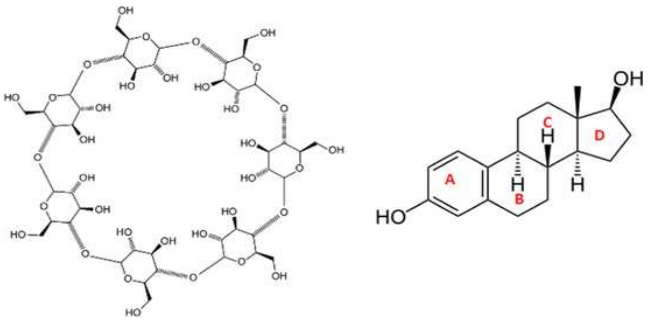


Fig. 1. Structures of β-CD and EST.

other scientists. Only two decades afterwards, this pre-defined EST-βCD complex molar ratio and its stability constant have been revisited [13,14]. However, again only either fluorescence or UV spectroscopy have been used for that purpose. Despite quite a few studies in this area, the 1:1 molar ratio of the complex in question has been accepted as a status quo and has never been questioned or verified by application of any other experimental methods such as high-resolution mass spectrometry (HRMS).

Recently, we have analyzed the EST-βCD complex in a solid state. For the first time, we have obtained and analyzed the structure of this complex's crystal structure. At the same time, it was one of the first ever obtained crystal structures of a steroid hormone complexed with any CD. The results have been published as [15] and the crystal structure has been deposited in the Cambridge Crystallographic Data Centre (CCDC). The results of the structural analysis in a solid state clearly indicate the 1:2 EST-βCD molar ratio, however they have also revealed significant structural disorder of this complex. Having obtained such results for a solid state, we were curious what is the molar ratio in the aqueous solution. Knowing the solid state structure of the complex, its 1:2 molar ratio after dissolution in water seems to be a scientifically sound hypothesis. Nevertheless, as we have gathered data on the previously performed experimental analyses repeatedly pointing out the 1:1 molar ratio, we decided against performing fluorescence or UV measurements and decided to concentrate on the molecular modelling techniques and HRMS spectrometry.

Complexes of various compounds with CDs have been analyzed using different molecular modelling approaches. At the beginning, computational methods used to predict the CD complexes structures and properties were the same as ones used for the analysis of much larger systems like receptors with ligands. Those were mainly Molecular Mechanics (MM) based methods and among them the most popular was molecular docking, sometimes followed by atomistic molecular dynamics (MD) simulations, at the same level of theory. However, in the recent years, with the increase of the computational power, such medium size systems like CD complexes started to be commonly analyzed using techniques based on the Quantum Mechanics (QM).

The conclusions of our recent review article [16] clearly show that the QM calculations era in the analysis of the CD complexes has already started more than a decade ago. The results show that the most commonly used are semi-empirical and Density Functional Theory (DFT) approaches. Among the former, PM6 and PM7 seem to be the most widely spread, also suggesting that they deliver the most appreciated results. When it comes to the DFT, a wider variety of computational options is used. The most frequent is application of B3LYP or M062X functionals both of the coming with or without dispersion correction. In terms of a solvent representation both possible approaches are practiced: either no solvent model is used or an implicit solvent model is applied.

As opposed to QM calculations, the Molecular Mechanics (MM) approach allows analysis of much larger systems but at the same time it delivers results of a significantly lower accuracy. However, for years CD complexes were too big for conducting QM calculations and Molecular Dynamics (MD) was the only option to obtain any information about the complexes' inner structure. Therefore, in the literature there are numerous examples of MD's application in the analysis of CD inclusion complexes. Even now this method is still in use. Therefore, we wanted to apply this well-described type of simulations in our study and compare them with the QM approaches.

Having gathered all the data on the previously conducted experimental analyses of the EST-βCD complex and having in mind our results concerning the solid state, we decided to try to determine the structure and the molar ratio of this complex. The aim of this study was to perform a benchmark studies on this topic using semi-empirical and DFT computational approaches. Additionally, we have managed to experimentally reveal the true complex's molecular ratio using a technique which has been never used in this particular case.

2. Materials and methods

2.1. Sample preparation

The EST-βCD complex was obtained by a method which is commonly used to obtain the CD inclusion complexes, a slow-cooling crystallization technique. This approach has already been used by us in a previous study describing the SCXRD analysis [15]. 60 mg of βCD was mixed in a flask with 1 mL distilled water and put into 70 °C water for 20 s to obtain a clear solution. Then the contents of the flask were poured into a beaker. In accordance with the molar mass of βCD and EST, the respective amount of EST was added to the beaker to maintain the 1:1 molar ratio. The beaker was put on a magnetic stirrer and left at room temperature for 15–20 min until a clear solution was obtained. Afterwards, the contents of the beaker were poured into a glass tube. The beaker was poured along with 0.5–1.0 mL water, which was also added to the glass tube. The glass tube was held in 70 °C water for 20 s to obtain a clear solution. Later, the tube was closed and put into a 70 °C water bath. A slow, gradual cooling process was performed over 10 days, reaching a temperature of 24 °C on the 10th day. At the end, a rotary evaporator was used.

2.2. HRMS

HRMS measurements were performed using Synapt G2-Si mass spectrometer (Waters) equipped with an ESI source and quadrupole-time-of-flight mass analyser. The mass spectrometer was operated in the positive ion detection mode. The optimized source parameters were: capillary voltage 3.0 kV, cone voltage 50 V, source temperature 110 °C, desolvation gas (nitrogen) flow rate 650 L/h with the temperature 450 °C, nebulizer gas pressure 6.5 bar. All samples were dissolved in water–methanol solution (1:1) and infused through a standard electrospray ion source into the instrument. The scan range was *m/z* 500–4000 and the acquisition method run time was 2 min. Mass calibration was performed using a cesium iodide solution. To ensure accurate mass measurements, data were collected in centroid mode and mass was corrected during acquisition using leucine enkephalin solution as an external reference (Lock-SprayTM) which generated reference ion at *m/z* 556.2771 Da ($[M + H]+$) in positive ESI mode. The results of the measurements were processed using the MassLynx 4.1 software (Waters) incorporated with the instrument.

2.3. Phase solubility studies

The UV–visible (UV–Vis) spectrophotometer (BioBase BK-S380, China) was utilized to assess the properties of EST and its inclusion complex. The EST showed a visible absorption peak at 280 nm. To generate a calibration curve, five standard solutions of EST in methanol

were measured, each replicated three times. The concentrations of the standards used were 0.1, 0.25, 0.50, 0.75, and 1.00 mM.

Phase solubility studies were carried out following the procedure outlined by Higuchi & Connors (1965) [17]. An excess amount of EST (50 mg) was added to 10 mL of deionized water containing various concentrations ranging from 0.20 to 20.00 mM for β-CD. The mixtures were then subjected to agitation using an orbital shaker (PHOENIX Instrument Laboratory Shaker RS-OS 5; Berlin, Germany) at 25 °C for 48 h to reach equilibrium. Afterwards, the collected samples were filtered through a 0.45 μm filter and assayed using a UV–Vis spectrophotometer at 280 nm.

2.4. QM calculations

All of the QM calculations were performed using the Gaussian 16 software [18]. All electron DFT computations were done employing the 6–311G(d,p) basis set and B3LYP or M062X functional, while the semi-empirical calculations have been done using PM6 and PM7 approaches. Those four methods were used either with or without Grimme's dispersion force corrections (D3). All of the calculations were performed either in vacuo or using one of the implicit solvation models: Polarizable Continuum Model (PCM) [19] or SMD (Solvation Model Density) [20], each time choosing water as the solvent with dielectric constant 78.540. For the details of computational models, please see Table 1 and Table 2. At the review stage, additional calculations have been performed at the ωB97X-D/6–31G(d,p)-PCM-Water level [21].

Vibrational frequencies were calculated to estimate thermodynamic parameters, including Zero Point Vibrational Energy (ZPVE) and Gibbs free energy (ΔG) at 298.15 K and 101.325 kPa.

According to our recent review, different types of QM approaches are commonly used when modeling CDs inclusion complexes. In this work we decided to use those methods which are the most commonly encountered in the recent literature. The goal was to compare the geometrically optimized 1:2 and 1:1 systems from a quantitative and qualitative perspective. For this reason, we have analyzed the structural aspects as well as values of the energy ΔE and Gibbs free energy ΔG defined as

$$\Delta E = E_{com} - (E_{EST} + nE_{CD}) \quad (1)$$

where E_{com} is energy of the EST-βCD complex, E_{EST} is the energy of EST, E_{CD} is the energy of βCD and n is the number of βCD molecules forming the complex, in this case $n = 1$ or $n = 2$

$$\Delta G = G_{com} - (G_{EST} + nG) \quad (2)$$

where G_{com} is free enthalpy of the EST-βCD complex, G_{EST} is the free enthalpy of EST, G_{CD} is the free enthalpy of βCD and n is the number of βCD molecules forming the complex, in this case $n = 1$ or $n = 2$

2.5. Molecular dynamics (MD) simulations

In addition to the crystallographically determined structure of the EST/β-CD inclusion complex, which is characterized by a host: guest ratio of 2:1 and a head-to-head inclusion mode, two additional models of the same complex assuming host: guest ratio of 1:1 but two different inclusion modes (down and up) were investigated through MD simulations. The initial coordinates for these latter two complexes were generated using molecular docking with Autodock Vina [22], where the simulation boxes were defined around the coordinates of the CD centers with size 40 Å in each direction and a grid spacing of 0.375 Å. The Lamarckian genetic algorithm in AutoDockTools was utilized for this purpose, enabling effective management of numerous degrees of freedom. The 3D structure of estradiol was retrieved from the PubChem database (Compound CID: 5757) [23], while the coordinates of β-CD correspond to its crystal structure in complex with EST [15]. The docking runs were set to 10 and the produced models with the most favorable binding energies for each one of the two inclusion modes were chosen for subsequent analysis.

The AMBER 12 software package [24] was used for the simulation of the EST/β-CD inclusion complexes in a aqueous environment. Three simulations were performed. In the first case, the stating 3D model was provided by the crystallographically determined atomic coordinates of a β-CD dimer[15], which includes one EST guest molecule (site A) inside the formed dimeric cavity. Thus, the host: guest stoichiometry of the entire system in the simulation was 2: 1. In the other two cases, monomers of EST/β-CD inclusion complexes with different inclusion modes from docking analysis were used as the starting models.

The geometry of EST was optimised following the AM1BCC methodology with the program Antechamber [25]. xLeaP, the GUI version of AMBER's LeaP program, was utilized for system preparation. The GLYCAM-06j [26] force field, which is suitable for β-CD atoms' treatment and the generalized AMBER (GAFF) (for the guest molecule) were applied for the simulation. Additionally, the TIP3P water model [27] was used to solvate the CD dimer in a periodic, octahedral box forming a 12 Å thick water shell around the structure.

Minimization and MD calculations that resulted in a single trajectory of the hydrated inclusion complex system were performed with Sander. The particle mesh Ewald summation approach [28] was followed in order to handle the long-range electrostatic interactions with a 10 Å cut-off limit for the direct space sum. Hydrogen bonds were handled using the SHAKE algorithm [29]. The simulation protocol was as following: Energy minimization for hydrogens and waters using 1000 steps of steepest descent (SD) followed by 500 steps of conjugated gradient (CG) methods, while the rest non-hydrogen atoms were fixed with positional restraints of 50 kcal mol 1 Å 2. Heating equilibration up to 300 K of the water in the canonical (NVT) ensemble for 50 ps using positional restraints and the Berendsen thermostat algorithm with coupling constants of 0.5 ps to control temperature and pressure. Energy minimization of all system atoms with weak positional restraints (10 kcal mol 1 Å 2), gradual temperature increase from 5 to 300 K with 10 kcal mol-1 Å-2 restraints on the atoms of the system followed by

Table 1
DFT computational approaches.

1	2	3	4	5	6	7	8	9	10	11	12	13
B3LYP In vacuo	M062X In vacuo	B3LYP-D3 In vacuo	M062X-D3 In vacuo	B3LYP PCM	M062 PCM	B3LYP-D3 PCM	M062X-D3 PCM	B3LYP SMD	M062X SMD	B3LYP-D3 SMD	M062X-D3 SMD	ωB97X-D PCM

Table 2
Semi-empirical computational approaches.

A	B	C	D	E	F	G	H	I	J	K	L
PM7 In vacuo	PM7-D3 In vacuo	PM6 In vacuo	PM6-D3 In vacuo	PM7 PCM	PM7-D3 PCM	PM6 PCM	PM6-D3 PCM	PM7 SMD	PM7-D3 SMD	PM6 SMD	PM6 -D3 SMD

gradual release of the restraints in successive steps at 300 K in NVT ensemble and finally density equilibration in the isobaric-isothermal (NPT) ensemble for 250 ps. Subsequently, production runs of the system under the NPT ensemble using a Berendsen-type algorithm with coupling constants of 1.0 ps were carried out under physiological conditions until reaching 12 ns.

The MD outputs were processed through the cpptraj module [30] of AMBER 12, to calculate the structural analyses (RMSD, distances, H-bonding). Moreover, the total guest binding energy ΔGbind, including the entropic term (ΔS) (calculated with the nmode module of AMBER 12) and the analysis of its components ΔEMM (changes in the gas-phase molecular mechanics (MM) energy) and ΔGsolv (solvation free energy) were computed with the aid of MM/PBSA.py script [31] implemented in AMBER 12. VMD [32] was also used for visualization and structural analyses of the MD trajectories.

3. Results and discussion

3.1. HRMS

As it was already mentioned, in our previous article concerning the EST-βCD complex in the solid state, we reported that the guest:host molar ratio in a crystal form is 1:2 [15]. However, we were interested whether 1:2 is the only form present in the aqueous solution of the complex, as well. Hence, we have defined 3 possible scenarios. The first option was presence of just 1:1 molar ratio complex, as it was stated many times in the literature. The second option was presence of both 1:1 and 1:2 complexes, where the 1:2 ratio would be the effect of the increasing concentration happening due to the solvent evaporation. According to that hypothesis, the amount of 1:2 complex should increase with the decrease of water content, resulting in the solely 1:2 complex in the solid state. And finally, the third possibility was presence of the identical stoichiometry as observed in a solid state which is 1:2 molar ratio.

To verify which scenario is correct, we have used the high-resolution mass spectrometry (HRMS). So far, this method has not been used often to define the molar ratio of the cyclodextrin complexes, however there are already some examples of its successful application for this purpose, published in the recent years [33,34]. The advantage of this method over the UV or fluorescence spectroscopy is that HRMS delivers a direct answer about the complex's molar ratio.

The results of the HRMS measurement of EST-βCD complex are presented in Fig. 2 and Table 3. EST forms stable associate with two βCD which can be detected in the mass spectra as ions corresponding to the protonated complex $[EST\text{-}2\beta CD + H]^+$ at *m/z* 2541.92 and the sodium adduct of complex $[EST\text{-}2\beta CD + Na]^+$ at *m/z* 2563.91. These observations suggest an interaction with 1:2 stoichiometry. No peaks corresponding to the association of one EST and one βCD were detected revealing the absence of 1:1 stable noncovalent complexes.

3.2. Phase solubility studies

The phase-solubility profile of EST in aqueous solution of successively increased β-CD concentrations at 25 °C (Fig. 3) indicates a B_S-type system which is usually observed with natural CDs, especially β-CD [34]. EST solubility increases linearly with increasing β-CD concentration in the range of 0.2 - 1.4 mM due to the formation of 1:1 EST:β-CD molar ratio inclusion complexes at the first stage. At the end of this linear portion, the maximum solubility S_{max} of EST is achieved. The solubility of the 1:1 EST:β-CD complex $S_{1:1}$ can be calculated as: $S_{1:1} = S_{max} - S_0$, where S_0 is the solubility of EST in pure water determined by the intercept of the phase solubility diagram.

Additional CD does not further increase the EST solubility and a first plateau is observed at β-CD concentration of 2 to 10 mM. This is due to the limited 1:1 complex solubility and/or the formation of 1:2 EST:β-CD molar ratio complexes. As the Gibbs phase rule indicates [35] only one

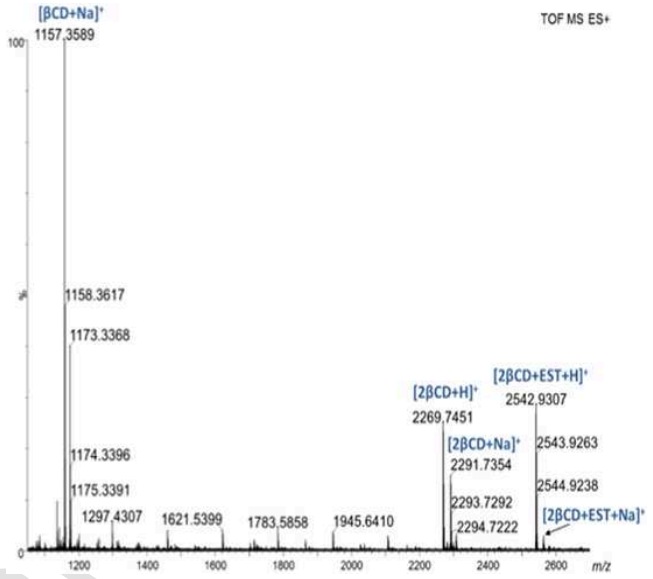


Fig. 2. HRMS spectrum of EST-βCD complex.

Table 3
The major peaks (*m/z* values and molecular formulas) from the HRMS measurement of the EST-βCD complex.

	Elemental composition	*m/z* $[M + H]^+$ or $[M+Na]^+$	
		Calculated	Found
$[\beta CD + Na]^+$	$C_{42}H_{71}O_{35}$	1157.3595	1157.3589
$[2\beta CD + H]^+$	$C_{84}H_{141}O_{71}$	2269.7474	2269.7451
$[2\beta CD + Na]^+$	$C_{84}H_{140}O_{70}Na_1$	2291.7293	2291.7354
$[2\beta CD + EST + H]^+$	$C_{102}H_{165}O_{72}$	2541.9250	2541.9248
$[2\beta CD + EST + Na]^+$	$C_{102}H_{164}O_{72}Na_1$	2563.9069	–

discrete complex may precipitate at the plateau segment of the diagram. Thus, at even higher β-CD concentrations (12 – 20 mM), where a second plateau is observed, the solubility approximates that of the pure 1:2 complex $S_{1:2}$. The scheme of the probable complex formation is given below as eq. 4.

For the two-step association process of EST complexation with β-CD, the apparent stability constants, K_1, K_2 of the following equilibria:

$$EST + CD \overset{K_1}{\leftrightarrow} EST/CD \tag{3a}$$

$$EST/CD + CD \overset{K_2}{\leftrightarrow} EST/CD2 \tag{4a}$$

and the $K_{overall} = K_1 \cdot K_2$ where estimated according to Liu et al. [35]

$$K_1 = \frac{slope}{S_0\,(1 - slope)} = 21,599\ M^{-1} \tag{3b}$$

$$K_2 = \frac{S_{1:2} \cdot\ slope}{{S_{1:1}}^2\,(1 - slope)} = 757\ M^{-1} \tag{4b}$$

where the slope was obtained from the linear part of the diagram and S_0, $S_{1:1}$, $S_{1:2}$ as described above. Thus, $K_{overall} = 1.6 \pm 0.4 \cdot 10^7\ M^{-2}$

This value is in the same order of magnitude to those estimated for progesterone, testosterone and cortisone by Liu et al. [36]. In that work the formation of inclusion complexes of steroids with β-CD at stoichiometric ratio of 1:2 was shown and its dependence on the steroid structure was discussed.

The *K* can be used to calculate Gibbs free energy (ΔG) according to the Eq. (5):

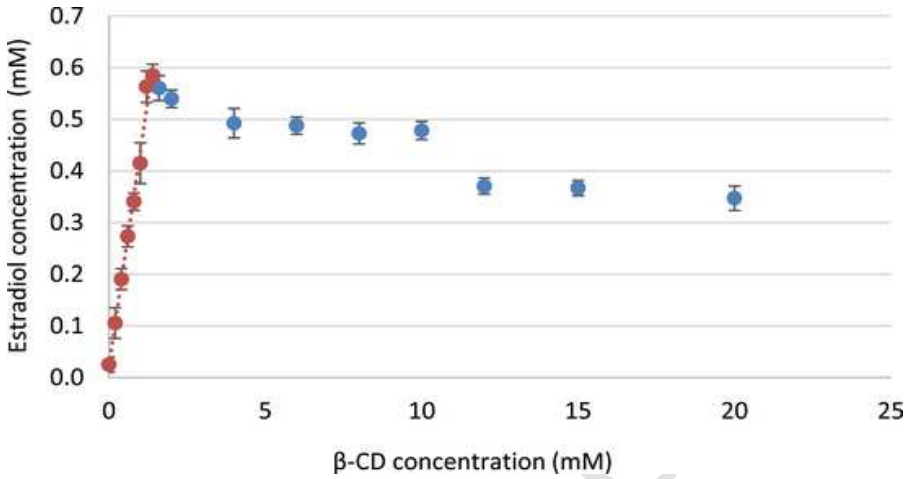


Fig. 3. Phase solubility diagram of EST/β-CD system in water at 25 °C ($n = 3$). The linear portion of the diagram (red line) was used for the calculation of K_1.

$$\Delta G = -RT\ln K \quad (5)$$

where R is gas constant (8.314 J mol^{-1} K^{-1}) and T is temperature (298 K).

This calculation reveals that the obtained here ΔG is equal to −9.92 kcal/mol.

3.3. QM calculations

As described above, the determination of this complex's stoichiometry is experimentally difficult and not straightforward task. Therefore, having in mind a huge applicability of the molecular modelling in the analysis of structure and properties of the CD complexes reviewed by us in the last year [16] we decided to apply these techniques also for this purpose. Therefore, the next step of our work was assessment of the QM approaches to check if they can properly foresee the host-guest molecular ratio and the complex association constant.

Knowledge about the crystal structure of a studied system significantly facilitates the calculations as it can be used to set the initial geometry of the complex. Thankfully, in our previous work we have determined the crystal structure of the EST-βCD complex [15]. In the current study it has been used as a starting point for all of the calculations.

Since we wanted to validate whether QM calculations can be used to predict the molar ratio and structure of the most stable complex, we have prepared 3 types of systems. The first one was 1:2 molar ratio complex and the experimental crystallographic structure of the hydrate of EST-βCD complex, after removing the water molecules, was used as an input for the computations. As EST is a molecule with quite limited conformational space, but also it is characterized by a structural anisotropy, to represent 1:1 molar ratio we needed two models called ''head up'' and ''head down'', as presented in Fig. 4. ''Head up'' is the case when EST's 5-carbon ring (steroidal D ring) goes through the wider CD's rim and the ''head down'' option is the case when it is the EST's 6-carbon ring (steroidal A ring) that protrudes through CD's wider rim. Those structures were based directly on the 1:2 experimental crystallographic data. We obtained them by removing each time different CD molecule from the 1:2 system.

3.4. DFT calculations

We have decided to apply Density Functional Theory (DFT) and semi-empirical methods. The chosen computational approaches were already presented in Table 1 and Table 2. In contrast to the majority of previously published cases of the CD complexes analysis by the means

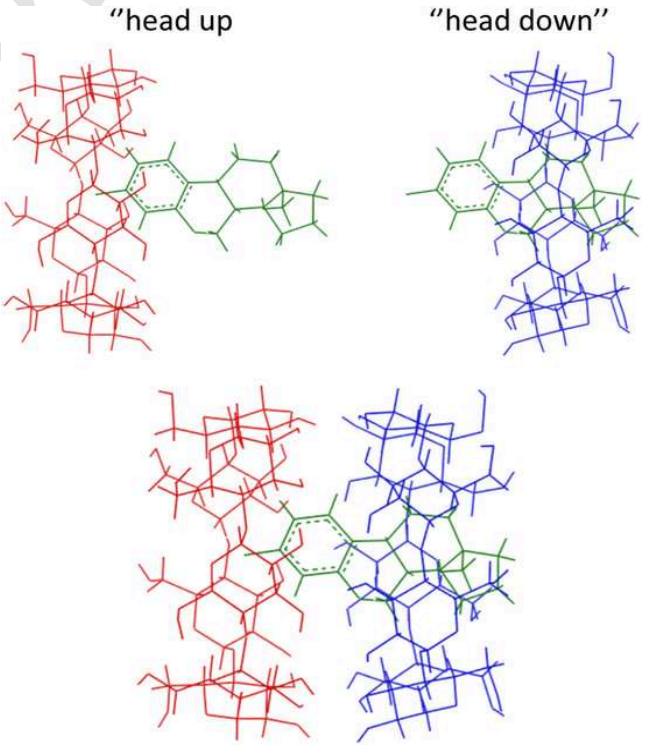


Fig. 4. Initial, non-optimized structures of 1:1 and 1:2 complexes, based on the experimental crystallographic data.

of DFT [16] where usually either 6–31 G or 6–31G(d) basis sets have been used, in this study for all DFT calculations we have used a relatively large 6–311G(d,p) basis set. Even if such approach has significantly elongated the computational time, it was a conscious choice. This way we have eliminated a risk of the influence of the basis set of the insufficient size on the obtained results.

Except of the choice of the functional in DFT (here: B3LYP and M062X) and type of the semi-empirical method (here: PM6 and PM7), one of the crucial decisions was about the type of the solvent representation. We have tested all 3 most common options: lack of solvent representation, PCM and SMD implicit solvent models.

PCM is the most often used solvent model in the computational analysis of CD complexes [37]. A different approach is presented by SMD. This model defines the free energy of solvation via two components: the one is electrostatic contribution arising from the self-consistent reaction field, the other comes from the short-range interactions between the solute and solvent molecules [20].

Second choice of the calculations parameters was the implementation or lack of the dispersion correction. Noncovalent forces like London and van der Waals interactions are crucial for the formation and stability of the CD inclusion complexes. This aspect is not included in the calculations using exchange-correlation functionals, as the long-range electron correlation effect, known as the London part of the dispersion energy term, is not included in the Kohn–Sham DFT equation. For years this was a real issue influencing the accuracy of DFT calculations. Nowadays, different dispersion corrections are available. According to our literature research [16], in almost all studies concerning CD complexes only Grimme dispersion correction (here: D3) was used. Hence, in this study we also apply only this type of dispersion correction and compare the results to the calculations where no dispersion correction was implemented.

The results of the DFT calculations are presented in Table 4. It is clearly visible that each of three variables used in this study and described above: type of functional, solvation scheme, dispersion correction had an influence on the results.

From the energetic point of view, in 6 out of 13 cases clearly the preferred structure is the '1:1, either "head up'' (or "head down". As we know from the HRMS studies, this is not consistent with the reality. The approaches which properly predict the 1:2 complex stoichiometry are models numbered 3, 5–8, 11 and 13. On this example we can see that applying dispersion correction improve the accuracy of results when the system is treated as the in vacuo one (models 1 and 3). Simultaneously, PCM solvent model works well also when no dispersion correction has been used (Models 5 and 6) [38].

Among them there is one scheme, the scheme number 5 (B3LYP-PCM, without dispersion correction), that most significantly favours the stability of the 1:2 molar ratio complex over the 1:1 ones, both in terms of energy and free enthalpy of complexation.

3.5. Semi-empirical calculations

From the energetic point of view, the results of the semi-empirical calculations are diversified, Table 5. In 10 out 12 cases, the 1:2 molecular ratio complex has been defined as the most stable one, what stays in accordance with the experimental results. However, in those two cases where the 1:1 complex stoichiometry has been favoured, difference between the 1:1 most preferred complex and 1:2 complex is small and has the value of 0.59 kcal/mol and 3.15 kcal/mol for C and F schemes, respectively. In other words, in general, when taking into account the energetic aspect, the applied semi-empirical approaches properly predicted which molar ratio describes a complex of the highest stability.

11 out 12 models show that all three options: 1:1 head up, 1:1 head down and 1:2, are energetically stable. Only in model K the value of ΔE is positive for 1:1 head down complex. In some computational schemes differences between these three possible options are almost neglectable, for instance in model C. Whereas in other cases, the differences are much bigger, like in model L, where this difference reached up to around 37 kcal/mol.

On the contrary, from thermodynamic point of view, the majority of models pointed out the 1:1 complex as the most stable. The exceptions are models A, B, I and K. However, similarly to the ΔE results, in almost all cases, all three structural options have been defined as probable. Only in model K some of the ΔG values are positive. Model K is also the only computational scheme which shows a distant difference between both 1:1 stoichiometries and 1:2 stoichiometry and distinctively favours the 1:2 complex molar ratio. Hence, we can assume that the K computational model (PM6-SMD, without dispersion correction) predicted the experimental results in a most accurate way.

Except for this one K model, in rest of the cases the differences in values between 1:1 head up and 1:1 head down options within one computational method are similar. This suggests that there is a similar probability of creation of 1:1 head up and 1:1 head down complexes.

All those findings may be taken as a guide to create the hypothesis on the 1:2 molar ratio complex formation path. Most probably, the complexation happens in two steps. The first one, is creation of the 1:1 complex. The second step is association of the second βCD to the already existing 1:1 EST-βCD system. Which complex out of those two, 1:1 head up or 1:1 head down, is formed at the beginning can be deducted from the ΔG values. The complex characterized by a lower value should be created as the first one. According to both the most accurate DFT approach, 5 (B3LYP PCM), as well as to the MD MMGBSA results, the 1:1 head up complex is the more stable one. Therefore, it is highly probable that the complex formation occurs according to the scheme (Eq. (6)) presented below.

$$EST + \beta CD \rightleftharpoons \mathbf{1:1\ up} \overset{\beta CD}{\rightarrow} \mathbf{1:2} \tag{6}$$

With the first step being reversible and second irreversible reaction. Such predictions about the mechanism of the complex's formation are possible only thanks to application of the molecular modelling approach. The structural analysis of the results can be find in the Supporting Information.

3.6. Computational thermodynamic results vs experimental data

As described in the previous section, obtention of the experimental complex stability constant allowed to define ΔG, which in case of EST-βCD complex is equal to −9.92 kcal/mol. Among the tested QM approaches, the ones which favour the most the 1:2 molar ratio are those in which the PCM correction has been applied However, the calculated ΔG overestimate the experimental ones Interestingly, the values close to the experimental one (−9.92 kcal/mol) have been obtained using M062X in vacuo (−9.18 kcal/mol) and PM7-D3 PCM approaches (−9.13 kcal/mol). Unfortunately, at the same time, both of them suggest that the 1:1 is more stable ratio than 1:2. Here, it should also be noted that β-CD is surrounded by hydration waters, which could be included in the equation of the complex formation. It was shown previously that an agreement between theoretical and experimental entropy data for inclusion complex formation was only attained when explicit water molecules were included [39].

3.7. MD simulations

MD simulations in aqueous media were carried out for the 3 types of systems, i.e. the crystallographically determined 1:2 guest:host complex and the two 1:1 monomeric complexes of opposite EST accommodation in the β-CD cavity (noted as "head up" and "head down" in Fig. 7). By monitoring the frames during the time interval of the simulations the following observations were made:

In the case of the 1:2 complex, where the starting model was retrieved from the crystal structure (Fig. 5a), the β-CD dimer encapsulating an EST molecule is preserved in the time frame of the simulation. In the absence of crystal contacts and in the presence of the surrounding water molecules, the guest EST rotates around and moves along the 7-fold molecular β-CD axis. However, it was observed a clear tendency of the guest's steroidal A-ring to be accommodated near the narrow rim of β-CD and its D-ring closer to the dimeric interface region. The measured distance between the center of mass (COM) of the A-ring and the O4n atoms mean plane of the host β-CD1, in whose cavity the A-ring is located, fluctuates around 1 Å, whereas that between the D-ring COM and the O4n plane of the other host (β-CD2), fluctuates in the range of 1 to 3 Å (Fig. 5b and 5c).

Table 4

DFT calculations results. Yellow colour indicates the lowest value of ΔE within the given method (within the column). Blue colour indicates the lowest value of ΔG within the given method (within the column). ΔG and ΔE "1:2 from 1:1 up" – the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "up" is used as an initial structure for the second step; ΔG and ΔE "1:2 from 1:1 down" - the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "down" is used as an initial structure for the second step.

		B3LYP In vacuo	M062X In vacuo	B3LYP -D3 In vacuo	M062X -D3 In vacuo	B3LYP PCM	M062X PCM	B3LYP -D3 PCM	M062X -D3 PCM	B3LYP SMD	M062X SMD	B3LYP -D3 SMD	M062X -D3 SMD	ωB97X-D PCM
model	kcal/mol	1	2	3	4	5	6	7	8	9	10	11	12	13
1:1 up	ΔE	-6.47	-37.35	-54.67	-49.14	-6.93	-31.49	-30.41	-45.89	-5.06	-30.35	-48.12	-44.56	-36.80
	ΔG	8.00	-12.28	-31.73	-34.28	6.12	-8.18	-9.64	-19.64	7.44	-10.08	-24.81	-21.35	-17.91
1:1 down	ΔE	0.44	-27.64	-42.20	-47.06	1.87	-30.82	-30.88	--43.37	1.99	-30.88	-49.20	-44.17	-41.70
	ΔG	15.40	-12.10	-21.32	-23.36	16.93	-8.22	-10.9	-19.09	14.24	-11.01	-26.97	-19.61	-19.90
1:2	ΔE	2.12	-26.46	-64.81	-47.99	-61.42	-81.56	-41.75	-118.54	5.59	-20.69	-65.69	-36.78	-120.68
	ΔG	13.33	-9.18	-38.67	-26.28	-47.69	-28.08	-49.66	-34.58	15.18	-0.75	-44.76	-19.26	-65.87
	ΔE	8.59	10.89	-10.14	1.15	-54.49	-50.07	-11.34	-72.65	10.65	9.66	-17.57	7.78	-83.88
1:2 from 1:1 up	ΔG	5.33	3.1	-6.94	8	-53.81	-19.9	-40.02	-14.94	7.74	9.33	-19.95	2.09	-47.96
	ΔE	1.68	1.18	-22.61	-0.93	-63.29	-50.74	-10.87	-161.91	3.6	10.19	-16.49	7.39	-78.98
1:2 from 1:1 down	ΔG	-2.07	2.92	-17.35	-2.92	-64.62	-19.86	-38.76	-15.49	0.94	10.26	-17.79	0.35	-45.97

Table 5

Semi-empirical calculations results. Yellow colour indicates the lowest value of ΔE within the given method (within the column). Blue colour indicates the lowest value of ΔG within the given method (within the column). ΔG and ΔE "1:2 from 1:1 up" – the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "up" is used as an initial structure for the second step; ΔG and ΔE "1:2 from 1:1 down" - the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "down" is used as an initial structure for the second step.

		PM7 In vacuo	PM7 -D3 In vacuo	PM6 In vacuo	PM6 -D3 In vacuo	PM7 PCM	PM7 -D3 PCM	PM6 PCM	PM6 -D3 PCM	PM7 SMD	PM7 -D3 SMD	PM6 SMD	PM6 -D3 SMD
model	kcal/mol	A	B	C	D	E	F	G	H	I	J	K	L
1:1 up	ΔE	-59.50	-40.53	-11.68	-37.84	-54.53	-35.59	-7.45	-34.91	-51.21	-32.38	-5.08	-24.27
	ΔG	-37.13	-27.15	-38.64	-54.75	-30.07	-30.93	-10.23	-43.58	-34.98	-17.14	7.82	-24.52
1:1 down	ΔE	-54.76	-38.20	-11.05	-35.41	-43.27	-38.76	-8.96	-35.52	-47.85	-34.58	86.58	-27.98
	ΔG	-37.14	-43.81	-25.83	-31.87	-25.31	-42.85	-5.93	-53.45	-41.29	-18.23	9.60	-54.32
1:2	ΔE	-84.25	-61.17	-11.09	-61.45	-60.84	-35.61	-18.59	-54.94	-61.54	-40.77	-19.91	-64.91
	ΔG	-71.81	-48.48	5.34	-41.83	-43.04	-9.13	0.62	-34.81	-48.49	-23.39	-3.33	-39.68
1:2 from 1:1 up	ΔE	-24.75	-20.64	0.59	-23.61	-6.31	-0.02	-11.14	-20.03	-10.33	-8.39	-14.83	-40.64
	ΔG	-34.68	-21.33	43.98	12.92	-12.97	21.8	10.85	8.77	-13.51	-6.25	-11.15	-15.16
1:2 from 1:1 down	ΔE	-29.49	-22.97	-0.04	-26.04	-17.57	3.15	-9.63	-19.42	-13.69	-6.19	-106.49	-36.93
	ΔG	-34.67	-4.67	31.17	-9.96	-17.73	33.72	6.55	18.64	-7.2	-5.16	-12.93	14.64

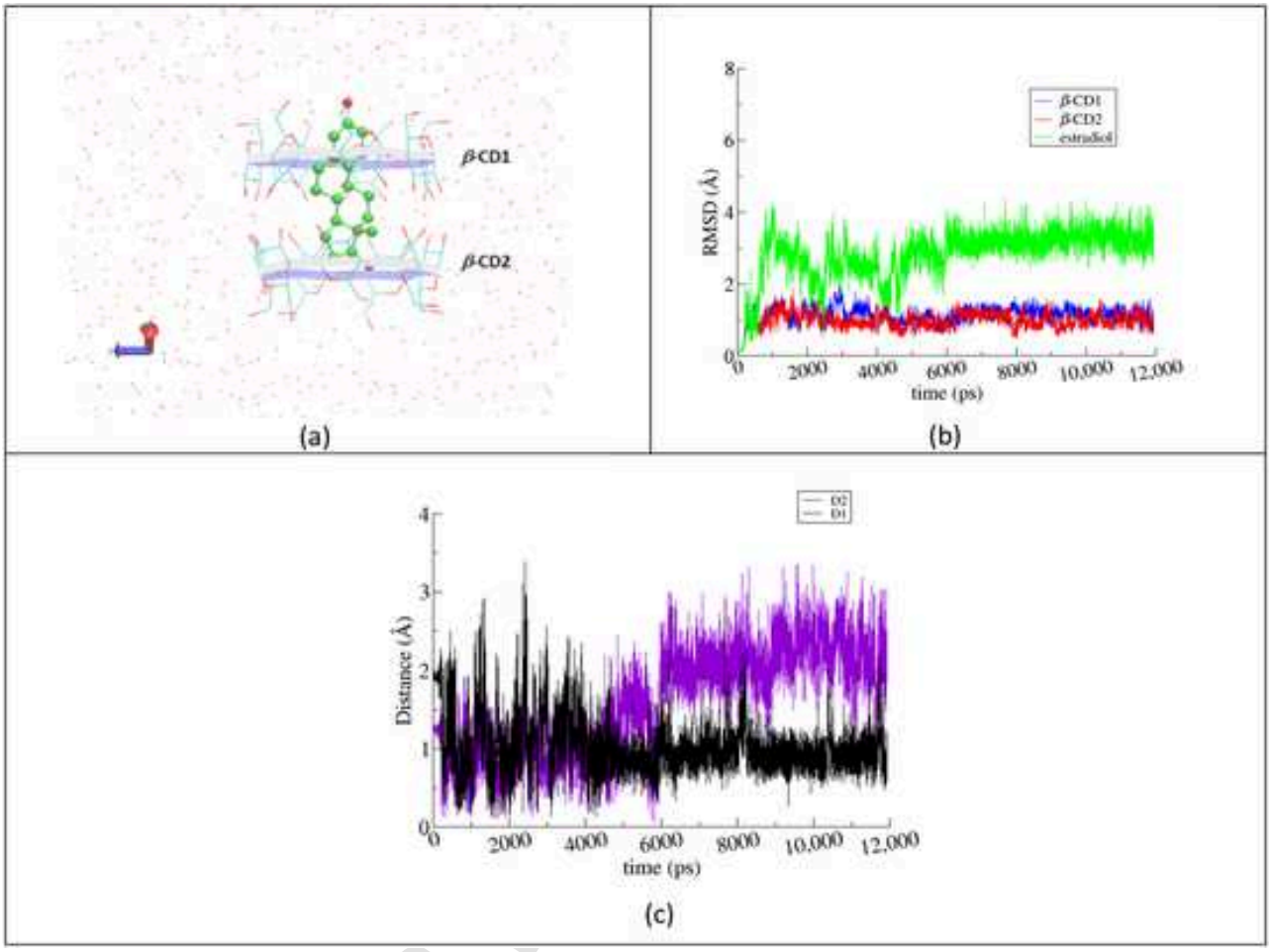


Fig. 5. (a) The starting model of the 1:2 inclusion complex based on the crystallographically determined atomic coordinates (CSD Refcode: OFANUI). (b) RMSD evolution of the host and guest molecules of the complex between the simulated states and the first frame of the simulations. (c) Distance D1 between the center of mass (COM) of the steroidal A-ring of EST and the O4*n* atom mean plane of the host β-CD1; distance D2 between COM of the steroidal D-ring of EST and the O4*n* atom mean plane of the other host of the dimeric cavity, β-CD2.

For the 1:1 molar ratio complex, both "head up" and "head down" models were examined. In the case of the "head down" binding mode, the A-ring of the guest cannot be stabilized in the host's wide rim with the rest part of EST protruding from the narrow rim of the host. Thus, EST is swiftly displaced from its initial location, exposing its A-ring to the solvent and accommodating the D-ring in the cavity (Fig. 6a). This behavior is also reflected in the high EST mobility displayed in the respective RMSD plot (Fig 6b).

On the other hand, in the case of the "head up" binding mode, the A-ring of EST which is initially exposed to the solvent by protruding from the narrow β-CD rim, it is accommodated quickly in the narrow β-CD rim where it remains relatively stable in the time frame of the simulation (Fig. 7a). The respective RMSD plot for the molecules of the system clearly shows a lower mobility of the guest compared to that of the "head down" system.

From all the above, it is concluded that the formation of a 1:1 inclusion complex of the "head up" binding mode is favored over the "head down" binding mode. Moreover, the dynamic behavior of the examined 1:2 complex, showed a very stable complex that tends to retain the accommodation of the guest (with its A-ring near the narrow rim and the D-ring near the interface of the hosts' dimer) according to that of the "head up" binding mode. These findings support the proposed complex formation described in the scheme of eq. 4.

Molecular Mechanics/Generalized Bohrn surface area (MM/GBSA) calculations [40], performed for the 3 examined types of complexes, further verify the above conclusions. The estimated host–guest binding affinities, as listed in Table 6, were extracted from 10,000 snapshots over the last 10-ns of the MD simulations. As expected, the lower ΔG-bind value is estimated for the 1:2 complex mainly due to the extended van der Waals interactions between the guest and the dimeric host, that significantly decrease the averaged change of van der Waals energies (ΔEvdW) upon EST inclusion in the host dimeric cavity. By comparing the ΔGbind values estimated for the systems of common 1:1 guest:host stoichiometry ("head down" and "head up"), the considerably lower ΔGbind value of the "head up" system indicates a more stable inclusion complex.

3.8. Computational stability results vs experimental data

As described in the previous section, acquisition of the experimental complex stability constant allowed to define ΔG, which in case of EST-βCD complex is equal to −9.92 kcal/mol. This value is within the limits of uncertainty of ΔGbind for 1:2 complex, −13.66 ± 4.20 kcal/mol (Table 6). However, the values of ΔGbind obtained for 1:1 complex, either head-up or head-down orientations, −2.27 ± 2.39 and −6.25 ± 2.40 respectively, although not strictly within the uncertainty limits, are also close to the experimentally determined one. Therefore,

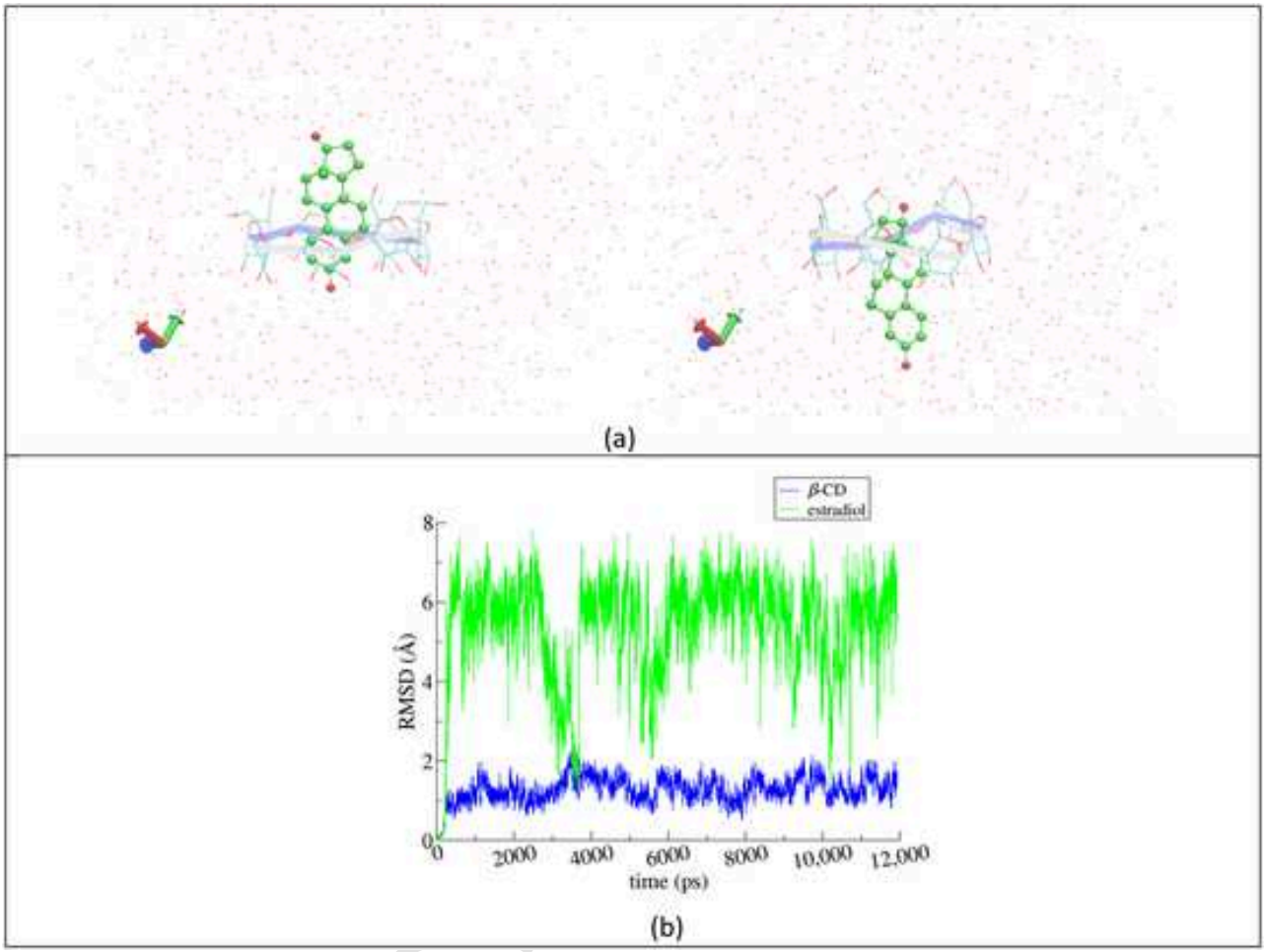


Fig. 6. (a) Two representative snapshots, at 0 and the 10th ns of the "head down" system simulation.
(b) RMSD plot for the host and guest molecule of the complex. The guest cannot be tightly stabilized in this orientation, thus exhibiting high mobility.

while the MD/MMGBSA method allows to properly indicate the order of the magnitude of ΔGbind, it should be supported by the experimental analysis such as HRMS to confirm the complex ratio.

4. Conclusions

Even though the existence of the EST-βCD complex in water solution has been known for decades, and despite quite some studies performed to define its molar ratio, this complex's structure remained not properly determined until now. The hypothesis that the knowledge in this context might be not complete, has occurred after determination of the EST-βCD crystal structure where the molar ratio was found to be 1:2 (EST:βCD).

In this work, thanks to application of the HRMS approach, it has been indisputably proven that the EST-βCD complex molar ratio is 1:2 and not as previously assumed 1:1. Moreover, the phase solubility studies confirmed these results. This type of experiments has been performed before, however, never the 1:2 molar ratio has been taken into account as a possible description of this system. In other words, the indisputable HRMS measurement results prompted the revision of the phase solubility studies. This allowed to properly define complex stability constant K, what in turn delivered information about the Gibbs free energy value of the complex formation.

The structure and thermodynamics of the complex were further analyzed using various QM (DFT, semi-empirical) and MM (MD/MMGBSA) approaches. Tests on the application of different computation parameters such as presence/absence of dispersion correction, choice of implicit solvent model or DFT functional, have been performed.

Possession of the credible experimental data allowed to assess the computational approaches. While some of the "static" QM methods properly indicated the correct host: guest ratio at the same time they failed to accurately predict the Gibbs free energy of complexation. On the other hand, QM methods that properly described the value of ΔG of 1:2 complex formation, such as M062X in vacuo, favored the 1:1 stoichiometry, which was experimentally excluded. The MD/MMGBSA method, although performed at the lower level of theory, accurately predicted the stability constant of the complexed but was not conclusive to indicate the formation of either 1:1 or 1:2 complex.

This leads to the conclusion that among tested computational approaches, there are some which are able to properly predict the composition of such complex and some that can assess the stability of the studied system. However, there is not single method that would allow to reproduce both the stoichiometry and the thermodynamic stability of the complex at the same time.This study finally describes the structure and thermodynamics of the EST-βCD complex in aqueous solution and delivers experiment-based information about the formation of this complex.

CRediT authorship contribution statement

Anna Helena Mazurek: Writing – original draft, Software, Investigation, Funding acquisition, Data curation, Conceptualization. **Łukasz**

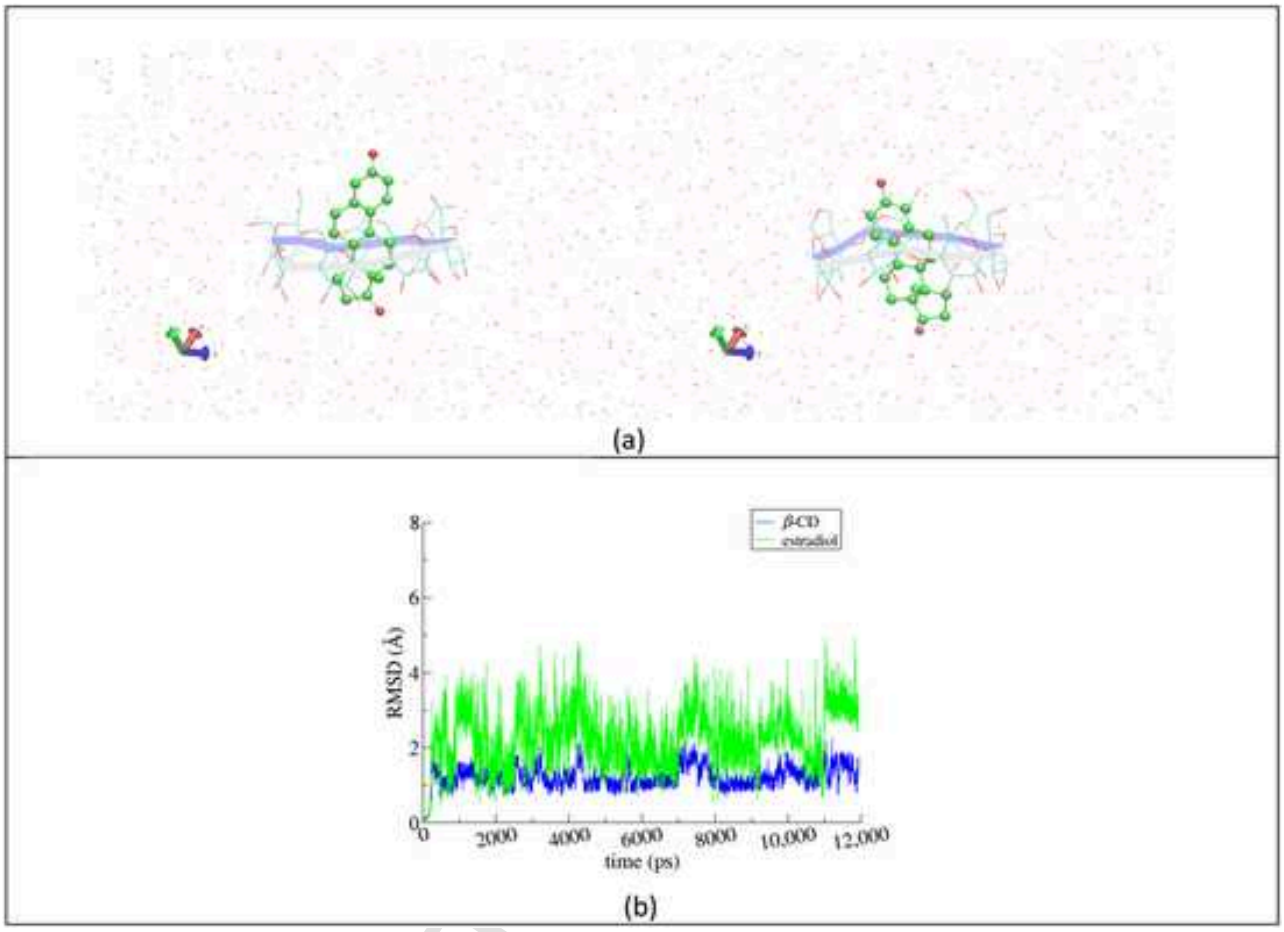


Fig. 7. (a) Two representative snapshots, at 0 and the 10th ns of the "head up" system simulation.
(b) RMSD plot the for host and guest molecule of the complex. The mobility of the guest is clearly lower than that of EST in the "head down" system simulation.

Table 6
Binding free energies and their standard deviations (kcal/mole) resulting from MM/GBSA analysis of the inclusion compounds of EST in β-CD with guest: host ratios of 1:2, 1:1 ("head down" mode) and 1:1 ("head up" mode), respectively.

	EST/β-CD (1:2)	EST/β-CD (1:1 "head down")	EST/β-CD (1:1 "head up")
ΔE_{vdW}	−47.31 ± 2.41	−29.10 ± 2.31	−30.46 ± 1.68
ΔE_{ele}	−4.15 ± 2.78	−2.12 ± 2.13	−2.09 ± 1.72
ΔE_{MM} [a]	−51.46 ± 3.58	−31.23 ± 3.21	−32.55 ± 2.38
ΔG_{GB}	23.22 ± 2.98	15.63 ± 2.80	13.45 ± 1.99
$\Delta G_{nonpolar}$	−4.45 ± 0.22	−2.94 ± 0.16	−3.03 ± 0.12
$\Delta G_{solvation}$ [b]	18.77 ± 2.94	12.69 ± 2.72	10.42 ± 1.94
ΔH [c]	−32.69 ± 4.63	−18.54 ± 2.06	−22.13 ± 1.97
T·ΔS [d]	−19.03 ± 3.00	−16.26 ± 1.21	−15.88 ± 1.36
ΔG_{bind} [e]	−13.66 ± 4.20	−2.27 ± 2.39	−6.25 ± 2.40

ΔE_{vdW} = van der Waals contribution from molecular mechanics; ΔE_{ele} = electrostatic energy as calculated by the molecular mechanics force field; ΔG_{GB} = the electrostatic solvation energy (polar contribution) calculated using the GB model; $\Delta G_{nonpolar}$ = nonpolar contribution to the solvation free energy, calculated by the solvent-accessible surface area (SASA) method;.

[a] $\Delta E_{MM} = \Delta E_{vdW} + \Delta E_{ele}$;.
[b] $\Delta G_{solvation} = \Delta G_{GB} + \Delta G_{nonpolar}$;.
[c] $\Delta H = \Delta G_{solvation} + \Delta E_{MM}$;.
[d] $T{\cdot}\Delta S$ entropic term calculated by normal mode analysis;.
[e] $\Delta G_{binding} = \Delta H - T{\cdot}\Delta S$.

Szeleszczuk: Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Conceptualization. **Kostas Bethanis:** Writing – original draft, Methodology, Investigation, Conceptualization. **Elias Christoforides:** Software, Resources, Project administration, Methodology. **Marta Katarzyna Dudek:** Software, Methodology, Conceptualization. **Ewelina Wielgus:** Writing – review & editing, Validation, Project administration. **Dariusz Maciej Pisklak:** Writing – review & editing, Writing – original draft, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the Medical University of Warsaw, Poland [grant number WF7/1/F/MB/N/23].

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.molstruc.2024.138710.

References

[1] J. Szejtli, Introduction and general overview of cyclodextrin chemistry, Chem. Rev. 98 (1998) 1743–1754.
[2] B.G. Poulson, Q.A. Alsulami, A. Sharfalddin, E.F. El Agammy, F. Mouffouk, A.-H. Emwas, L. Jaremko, M. Jaremko, Cyclodextrins: structural, chemical, and physical properties, and applications, Polysaccharides 3 (2022) 1.
[3] G. Crini, Review: a history of cyclodextrins, Chem. Rev. 114 (2014) 10940–10975.
[4] T. Irie, K. Uekama, Pharmaceutical applications of cyclodextrins. III. Toxicological issues and safety evaluation, J. Pharm. Sci. 86 (1997) 147–162.
[5] V. Aiassa, C. Garnero, M.R. Longhi, A. Zoppi, Cyclodextrin multicomponent complexes: pharmaceutical applications, Pharmaceutics 13 (2021) 1099.
[6] S.S. Jambhekar, P. Breen, Cyclodextrins in pharmaceutical formulations I: structure and physicochemical properties, formation of complexes, and types of complex, Drug Discov. Today. 21 (2016) 356–362.
[7] S.S. Jambhekar, P. Breen, Cyclodextrins in pharmaceutical formulations II: solubilization, binding constant, and complexation efficiency, Drug Discov. Today. 21 (2016) 363–368.
[8] M.P. Thomas, B.V. Potter, The structural biology of oestrogen metabolism, J. Steroid Biochem. Mol. Biol. 137 (2013) 27–49.
[9] Salole E.G. Estradiol, Analytical profiles of drug substances, 1986, 15, 283–318.
[10] N. Sadlej-Sosnowska, Fluorometric determination of association constants of three estrogens with cyclodextrins, J Fluoresc 7 (1997) 195–200.
[11] C. Yañez, J. Basualdo, P.J. Jara-Ulloa, A. Squella, Inclusion complexes of estrone and estradiol with β-cyclodextrin: voltammetric and HPLC studies, J. Phys. Org. Chem. 20 (2007) 499–505.
[12] R.L. Pérez, G.M. Escandar, Spectrofluorimetric study of estrogen-cyclodextrin inclusion complexes in aqueous systems, Analyst 138 (2013) 1239–1248.
[13] D.H. Schwarz, A. Engelke, G. Wenz, Solubilizing steroidal drugs by β-cyclodextrin derivatives, Int J Pharm 531 (2017) 559–567.
[14] Z.Y. Lin, X.X. Wang, S.B. Kou, J.H. Shi, Exploring the inclusion interaction of estradiol with β-CD and HP-β-CD with the help of molecular dynamics simulation as well as multi-spectroscopic approaches, Spectrochim Acta A Mol Biomol Spectrosc 269 (2022) 120764.
[15] A.H. Mazurek, Ł. Szeleszczuk, K. Bethanis, E. Christoforides, M.K. Dudek, M. Zielińska-Pisklak, D.M. Pisklak, 17-β-Estradiol—B-Cyclodextrin complex as solid: synthesis, structural and physicochemical characterization, Molecules 28 (2023) 3747.
[16] A.H. Mazurek, Ł. Szeleszczuk, Current status of quantum chemical studies of cyclodextrin host–guest complexes, Molecules 27 (2022) 3874.
[17] T. Higuchi, K.A. Connors, Phase solubility techniques, Adv. Anal. Chem. Instrument. 4 (1965) 117–212 - Open Access Library Available online: http://www.oalib.com/references/7163685. (accessed on 26 January 2024).
[18] Software https://gaussian.com/products/ accessed on 24th January 2024
[19] M. Cossi, N. Rega, G. Scalmani, V. Barone, Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model, J. Comput. Chem. 24 (2003) 669–681.
[20] A.V. Marenich, C.J. Cramer, D.G Truhlar, Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions, J. Phys. Chem. B 113 (2009) 6378–6396.
[21] J.-D. Chai, M. Head-Gordon, Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections, Phys. Chem. Chem. Phys. 10 (2008) 6615.
[22] J. Eberhardt, D. Santos-Martins, A.F. Tillack, S. Forli, AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings, J. Chem. Inf. Model. 61 (2021) 3891–3898.
[23] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B.A. Shoemaker, P.A. Thiessen, B. Yu, et al., PubChem 2023 update, Nucleic Acids Res. 51 (2023) D1373–D1380.
[24] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the amber biomolecular simulation package, Wiley Interdiscip. Rev.: Comput. Mol. Sci. 3 (2012) 198–210.
[25] J. Wang, W. Wang, P. Kollman, D. Case, ANTECHAMBER: an accessory software package for molecular mechanical calculations, J. Chem. Inf. Comput. Sci. - JCISD (2000) 222.
[26] K.N. Kirschner, A.B. Yongye, S.M. Tschampel, J. Gonzalez-Outeirino, C.R. Daniels, B.L. Foley, R.J. Woods, GLYCAM06: a generalizable biomolecular force field. Carbohydrates, J. Comput. Chem. 29 (2008) 622–655.
[27] P. Mark, L. Nilsson, Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K, J. Phys. Chem. A 105 (2001) 9954–9960.
[28] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[29] V. Kräutler, W.F. van Gunsteren, P.H. Hünenberger, A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations, J. Comput. Chem. 22 (2001) 501–508.
[30] D.R. Roe, T.E. Cheatham 3rd, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theory Comput. 9 (2013) 3084–3095.
[31] B.R. Miller 3rd, T.D.J. McGee, J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBSA.Py: an efficient program for end-state free energy calculations, J. Chem. Theory Comput. 8 (2012) 3314–3321.
[32] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (33–38) (1996) 27–28.
[33] G. Zengin, Yildiztugay E. Nilofar, A. Bouyahya, H. Cavusoglu, R. Gevrenova, D. Zheleva-Dimitrova, A comparative study on UHPLC-HRMS profiles and biological activities of inula sarana different extracts and its beta-cyclodextrin complex: effective insights for novel applications, Antioxidants (Basel) 12 (2023) 1842.
[34] T.K. Špehar, M. Pocrnić, D. Klarić, B. Bertoša, A. Čikoš, M. Jug, J. Padovan, S. Dragojević, N. Galić, Investigation of praziquantel/cyclodextrin inclusion complexation by NMR and LC-HRMS/MS: mechanism, solubility, chemical stability, and degradation products, Mol. Pharm 18 (2021) 4210–4223.
[35] M.E. Brewster, T. Loftsson, Cyclodextrins as pharmaceutical solubilizers, Adv. Drug Deliv. Rev. 59 (2007) 645–666.
[36] F.Y. Liu, D.O. Kildsig, A.K. Mitra, Beta-cyclodextrin/steroid complexation: effect of steroid structure on association equilibria, Pharm. Res 7 (1990) 869–873.
[37] M. Cossi, N. Rega, G. Scalmani, V. Barone, Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model, J. Comput. Chem. 24 (2003) 669–681.
[38] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132 (2010) 154104.
[39] J.F. Lopes, Nascimento C.S. Jr, C.P.A. Anconi, H.F.D. Santos, W.B Almeida, Inclusion complex thermodynamics: the β-cyclodextrin and sertraline complex example, J. Mol. Graph. Model 62 (2015) 11–17.
[40] S. Genheden, U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, Expert Opin. Drug Discov. 10 (2015) 449–461.

molecules

MDPI

Article

17-β-Estradiol—β-Cyclodextrin Complex as Solid: Synthesis, Structural and Physicochemical Characterization

Anna Helena Mazurek [1,2], Łukasz Szeleszczuk [1,*], Kostas Bethanis [3], Elias Christoforides [3], Marta Katarzyna Dudek [4], Monika Zielińska-Pisklak [5] and Dariusz Maciej Pisklak [1]

1 Department of Organic and Physical Chemistry, Faculty of Pharmacy, Medical University of Warsaw, Banacha 1 Str., 02-093 Warsaw, Poland; anna.mazurek@wum.edu.pl (A.H.M.)
2 Doctoral School, Medical University of Warsaw, Żwirki i Wigury 81 Str., 02-093 Warsaw, Poland
3 Laboratory of Physics, Department of Biotechnology, Agricultural University of Athens, 11855 Athens, Greece
4 Structural Studies Department, Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, Sienkiewicza 112 Str., 90-363 Łódź, Poland
5 Department of Pharmaceutical and Biomaterials Chemistry, Faculty of Pharmacy, Medical University of Warsaw, Banacha 1 Str., 02-093 Warsaw, Poland
* Correspondence: lukasz.szeleszczuk@wum.edu.pl; Tel.: +48-501-255-121

Abstract: 17-β-estradiol (EST) is the most potent form of naturally occurring estrogens; therefore, it has found a wide pharmaceutical application. The major problem associated with the use of EST is its very low water solubility, resulting in poor oral bioavailability. To overcome this drawback, a complexation with cyclodextrins (CD) has been suggested as a solution. In this work, the host–guest inclusion complex between the ß-CD and EST has been prepared using four different methods. The obtained samples have been deeply characterized using ^{13}C CP MAS solid state NMR, PXRD, FT-IR, TGA, DSC, and SEM. Using SCXRD, the crystal structure of the complex has been determined, being to the best of our knowledge the first solved crystal structure of an estrogen/CD complex. The periodic DFT calculations of NMR properties using GIPAW were found to be particularly helpful in the analysis of disorder in the solid state and interpretation of experimental NMR results. This work highlights the importance of a combined ssNMR/SCXRD approach to studying the structure of the inclusion complexes formed by cyclodextrins.

Keywords: cyclodextrin; estradiol; DFT; SCXRD; solid state NMR

Citation: Mazurek, A.H.; Szeleszczuk, Ł.; Bethanis, K.; Christoforides, E.; Dudek, M.K.; Zielińska-Pisklak, M.; Pisklak, D.M. 17-β-Estradiol—β-Cyclodextrin Complex as Solid: Synthesis, Structural and Physicochemical Characterization. *Molecules* **2023**, *28*, 3747. https://doi.org/10.3390/molecules28093747

Academic Editors: Rosa Iacovino, Marina Isidori and Margherita Lavorgna

Received: 31 March 2023
Revised: 20 April 2023
Accepted: 24 April 2023
Published: 26 April 2023

1. Introduction

17-β-estradiol, EST (Figure 1), is the most potent form of naturally occurring estrogens [1]; therefore, it has found wide application in hormonal contraception, hormone replacement therapy (HRT), and treatment of menopausal and postmenopausal symptoms [2]. Oral administration of EST in a solid dosage form is the most favorable form of HRT [3]. While in the European Pharmacopoeia only the hemihydrate form of EST is described, recently its anhydrous form was successfully obtained [4]. Moreover, numerous cocrystals of EST have been designed [5,6] to solve one of the major problems associated with the application of EST: its poor oral bioavailability caused by very low water solubility—0.2–5 μg mL^{-1} [7].

Cyclodextrins (CDs) are cyclic oligosaccharides consisting of a macrocyclic ring formed by glucose subunits joined by α-1,4 glycosidic bonds. CDs are primarily used in pharmaceutical formulations due to their unique properties, resulting in their ability to form inclusion complexes [8]. The desirable properties of CDs in the pharmaceutical field can be explained at the molecular level. CD molecules resemble a "doughnut" ring, in which small, non-polar substances such as EST can be entrapped. The external fragments of CD molecules are polar due to the presence of hydroxyl groups. When a non-polar substance (e.g., an EST) enters the molecular hole of cyclodextrin, the formed host–guest complex

is polar (at outside) and, therefore, is more soluble than the separated guest molecule. Therefore, CDs are commonly used in pharmaceutical formulations as they are able to increase the solubility of APIs, protect them against external factors, such as light, humidity, and heat, or can even mask unpleasant smells or flavors of drugs. Currently, more than 100 original drugs are manufactured with CDs as excipients [9–11].

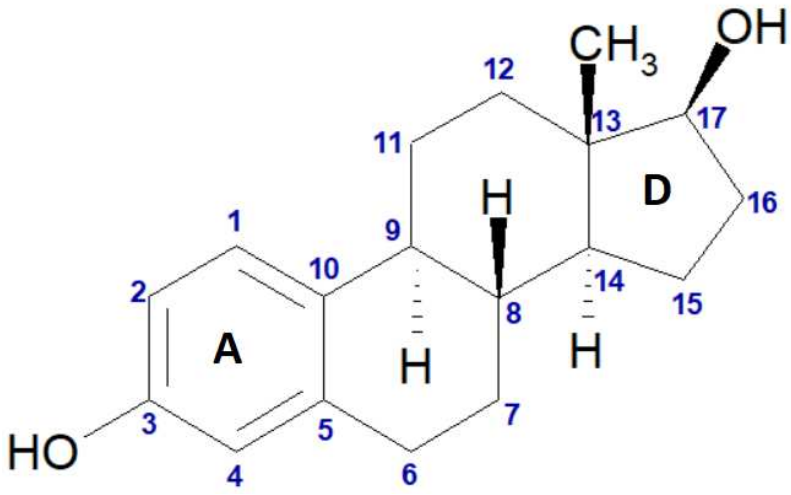


Figure 1. Chemical structure of 17-β-estradiol (EST) with atom numbering. "A" and "D" represent the symbols of the particular rings within the structure.

Multiple preparation methods for CD inclusion complexes are being exploited, such as the solvent evaporation method, grinding method, ultrasonic method, and freeze-drying method. It has been shown in many examples that the method of complex preparation may have a major impact on the obtained form of the final product [12–14].

Only a small amount of CD complexes have been reported with their crystal structures. This is caused by the fact that many of these complexes are either amorphous or polycrystalline, and even for the crystalline complexes, it is usually very hard to obtain a crystal of a size suitable for single-crystal X-ray measurements [15]. However, in order to fully understand the aforementioned changes resulting from complexation, knowledge of the molecular structure of CD complexes is crucial.

To achieve this goal—that is, to understand the structure and dynamics of the CD-based complexes—multiple computational and analytical methods are usually applied, some of which have been recently reviewed by us [16–18]. Among the analytical methods most commonly used to study these kinds of materials in a solid state are Fourier-transform infrared spectroscopy (FT-IR) and powder X-ray diffraction (PXRD), together with thermo-analytical techniques such as differential scanning calorimetry (DSC) and thermogravimetric analysis (TGA). Moreover, the application of solid-state nuclear magnetic resonance (ssNMR) can provide essential information, unobtainable by other methods [16]. From the theoretical approaches, the most important and accurate methods are obtained through the use of quantum chemical calculations, usually at the density functional theory (DFT) level [19]. A recent review [17] revealed that the application of quantum chemical calculations in studies of CD complexes can be essential, providing results unobtainable by any other method, both experimental and computational. In particular, density functional theory (DFT) methods are among the most accurate and most frequently used to model such systems, with the PBE functional being the method of choice when modeling solid state structures. However, to the best of our knowledge, such computations have not been performed on CD inclusion complexes yet.

The properties and structures of the complexes of EST with various CDs have been studied extensively for the last 30 years [20–27]. For example, in a recent work [28], single-crystal X-ray diffraction (SCXRD) results of the EST/ß-CD complex were presented, determining the unit cell and the crystallographic space group of the crystal structure. In

this work, only the atomic coordinates of the host molecule were determined, whereas the encapsulated hormone was not possible to be modeled due to its high disorder. Thus, although a 2:1 host:guest stoichiometry of the complex was estimated (based on the residual density calculated by that incomplete model), the structural information provided was limited to the host molecular arrangement in the crystalline state. On the other hand, in the present work, the crystal structure of the EST/ß-CD complex was fully determined and the atomic positions of both host and guest molecules reveal the orientation of the guest in the ß-CD host dimeric cavity and give valuable information about the intermolecular (host–guest and guest–guest) interactions and the arrangement of the full complex units in the crystalline state.

The aim of our work was to obtain the 17-β-estradiol–β-cyclodextrin complex (EST/ß-CD) by means of four different preparation methods. The obtained samples have been extensively characterized by means of various analytical (ssNMR, FT-IR, SCXRD, PXRD, SEM, DSC, TGA) and computational (periodic DFT) methods to extend knowledge of this complex formation and to study how the method of preparation influences the final results.

2. Results and Discussion

2.1. SCXRD Results

As was mentioned above, a single crystal analysis of the EST/ß-CD complex has been presented in a previous work [28] where only the host and the water oxygen atoms were located via the collected diffraction data using a Mo X-ray source, whereas the encapsulated hormone was not possible to be modeled as the residual electron density appearing within the host cavity was very low ($\Delta\rho \leq 1$ e·Å^3) due to the high disorder of the included estradiol. Similar to that work, the EST/ß-CD crystal structure presented here was found to belong to the monoclinic system with space group C2, with roughly the same lattice parameters (Table 1). However, in addition to the host and water oxygen atoms, the coordinates of the encapsulated estradiol atoms were successfully determined. This is likely due to the higher diffracted intensities collected by using a Cu *Ka* X-ray source, resulting in significant higher and discrete difference electron density peaks that allowed for the modeling of the disordered guest.

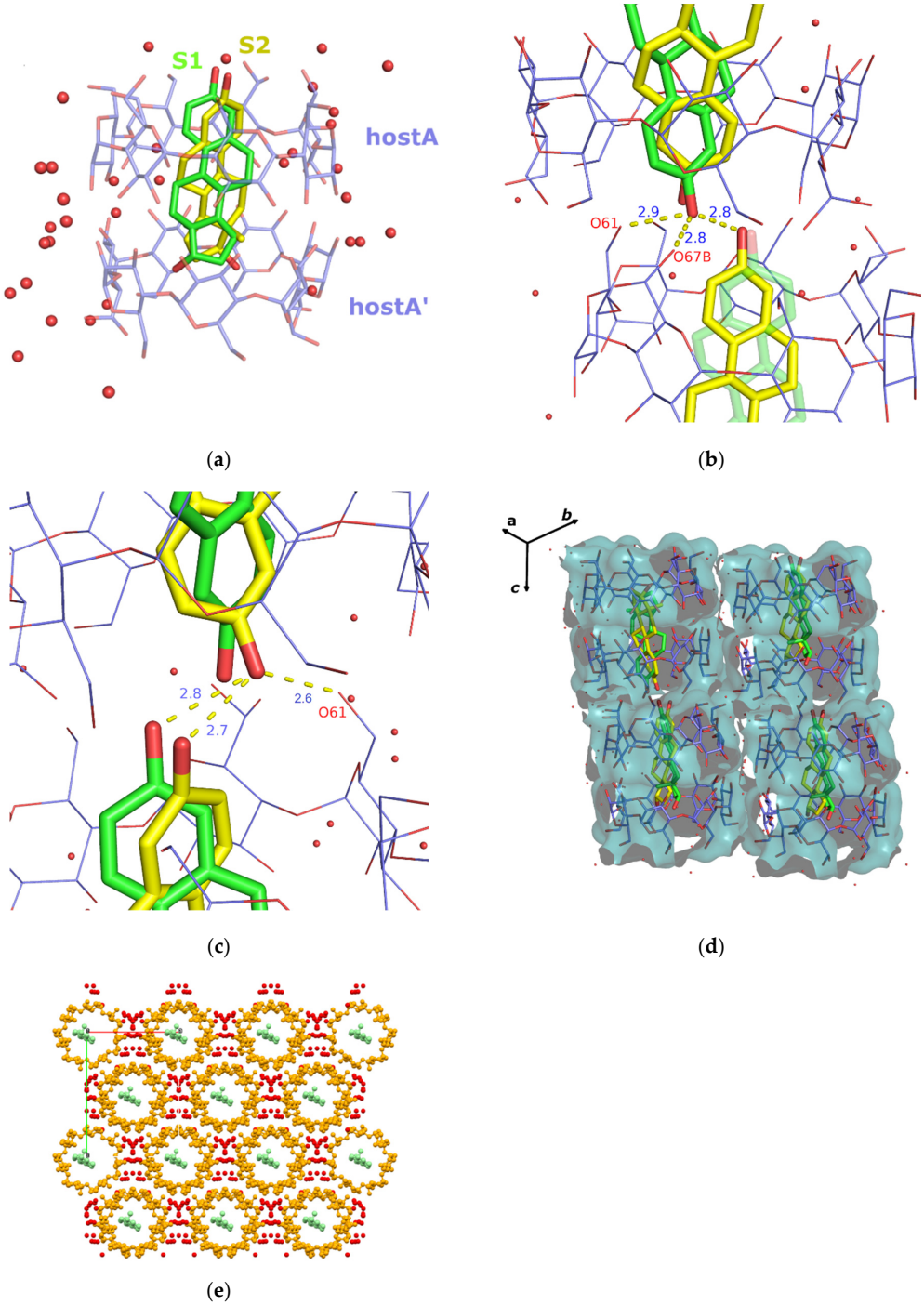
The determined asymmetric unit of the EST/ß-CD crystal structure contains one host, one guest (with s.o.f. of 0.5), and 10.5 water oxygens distributed over 17 sites (no water hydrogens were included). As the complex crystallizes in the C2 space group, the host:guest stoichiometry is 2:1, with two symmetry-related hosts (denoted as hostA and hostA') forming a classic "head-to-head dimer" stabilized by the well-known intermolecular hydrogen bonds between their secondary hydroxyls. The geometric features of the host molecule are reported in Supplementary Table S1, indicating that β-CD, upon complexation with estradiol, adopts the usual torus-like macrocycle shape and the round conformation due to the formation of the commonly observed intramolecular interglucose O3(n)-H···O2(n+1) hydrogen bonds. The encapsulated estradiol molecule is found disordered over two sites (site S1 and site S2, with s.o.f of 0.25 each). Both occupied sites have the same orientation: the guest is accommodated "axially" inside the dimeric β-CD cavity, with the hydroxyl of its A ring protruding from the primary rim of the host β-CD and its D ring being buried into the dimeric hydrophobic cavity (Figure 2a). More specifically, the mean plane of the estradiol aromatic ring system forms an angle of 95.795 (2)° and 82.404 (7)° in the case of the S1 and S2 occupied sites, respectively, with the mean plane of the glucosidic O4n atoms of the hosts. The oxygen atom of the guest's A ring hydroxyl is located at a distance of 0.638 (5) Å for S1 and 0.428 (3) Å for S2 above the mean plane of the O6n atoms of hostA, whereas the oxygen of the guest's D ring hydroxyl is found near the O4n atom plane of hostA' (the distance between the oxygen and the mean plane of the O4n atoms of hostA' being 0.1642 (12) Å and 0.112 (8) Å for S1 and S2, respectively). The protruding hydroxyl of the guest's A ring is at hydrogen bond distance from the primary hydroxyls of the host and the protruding hydroxyl of the guest of the consecutive complex unit. In particular, the A

ring hydroxyl of the guest occupying the S1 site can be hydrogen bonded with the fully occupied O(61)H, the 40% occupied O(67B)H of the host (1 − x, y, 1 − z), and the guest's A ring hydroxyl occupying the S2(1 − x, y, 1 − z) site (Table S2 and Figure 2b). However, as the distance between the two consecutive S1 and S1(1 − x, y, 1 − z) sites is just 2.4 Å, the guest is sterically forbidden to occupy the S1 site in two successive complex units. On the other hand, the A ring hydroxyl of the S2 site can be hydrogen bonded with the O(61)H of the host and the guest occupying the S2 site in the successive (1 − x, y, 1 − z) complex unit (Figure 2c). Thus, the arrangement of the encapsulated estradiol in the consequent dimers could be that of S1-S2(1 − x, y, 1 − z) or S2-S2(1 − x, y, 1 − z) but not S1-S1(1 − x, y, 1 − z) (left column of Figure 2d), whereas the S1-S1(x, y, −1 + z), S1-S2(x, y, −1 + z), and S2-S2(x, y, −1 + z) (right column of Figure 2d) arrangement is also possible.

Table 1. Crystallographic parameters of the EST/ß-CD inclusion complex [29].

Crystal Data	**EST/ß-CD**
CCDC No.	2250781
Complex formula in the asymmetric unit	$(C_{42}H_{70}O_{35})\cdot 0.5(C_{18}H_{24}O_2)\cdot 10.5H_2O$
Formula weight	1439.16
Crystal system, space group	Monoclinic, *C*2
Temperature (K)	100
a, *b*, *c* (Å)	19.1245 (15), 24.4180 (18), 15.6004 (11)
α, β, γ (°)	109.500 (5)
V (Å^3)	6867.2 (9)
Z	4
Radiation type	Cu *Ka*
μ (mm^{-1})	1.09
Crystal size (mm^3)	0.4 × 0.27 × 0.13
Data collection	
T_{min}, T_{max}	0.593, 0.754
No. of measured, independent, and observed [I > 2σ(I)] reflections	105,516, 11,945, 10,450
R_{int}	0.074
$(\sin \theta/\lambda)_{max}$ (Å^{-1})	0.595
Refinement	
R_1 [$F^2 > 2\sigma(F^2)$], $wR_2(F^2)$, GooF	0.095, 0.269, 1.04
No. of reflections	11,945
No. of parameters	950
No. of restraints	202
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.78, −0.44

The "head-to-head" β-CD dimers are arranged according to the channel (CH) packing mode (Figure 2) along the c-axis, the distance and the shift between the centroids of two successive dimers being 15.600(4) and 3.0647(7) Å, respectively. The adjacent channels are stacked via bridge water molecules and host–host intermolecular hydrogen bonds (Figure 2e).



(**a**) (**b**)

(**c**) (**d**)

(**e**)

Figure 2. (**a**) Estradiol disordered over two sites (S1 green and S2 yellow) is fully encapsulated inside the cavity of a classic "head-to-head" β-CD host dimer with a 2:1 host:guest stoichiometry. (**b**) Hydrogen bonds between the protruding oxygen atom of the A ring hydroxyl of the guest occupying the S1 site (green) and primary host–guest (occupying the S2 site) of the consecutive complex unit. (**c**) Hydrogen bonds between the guest occupying the S2 site (yellow) and the host–guest (both occupied sites) of the consecutive complex unit. (**d**) β-CD dimers forming channels along the c-axis. Left-hand side: S1-S2(1 − x, y, 1 − z) and S2-S2(1 − x, y, 1 − z) possible arrangements of the encapsulated estradiol molecules in the channel. Right-hand side: S1-S1(x, y, −1 + z), S1-S2(x, y, −1 + z), and S2-S2(x, y, −1 + z) possible arrangements of the encapsulated estradiol molecules in the channel. (**e**) Crystal packing of the EST/β-CD complex view perpendicular to the ab plane. In all figures, hydrogen atoms are omitted for clarity.

2.2. Periodic DFT Calculation Results

As described in Section 2.1 (SCXRD results), due to the disorder in the atomic positions of EST in the crystal structure, two significantly different orientations of the neighboring guest molecules exist in the solid state. Therefore, using the experimental crystal structure and choosing the proper guest molecules, we have created two model periodic structures for the DFT calculations. In the first one, the A-rings of neighboring EST molecules are located close to each other (Figure 3). For the purposes of this study, we have named this structure DAAD. This structure can be also found in the Supplementary Material ESI as DAAD.cif.

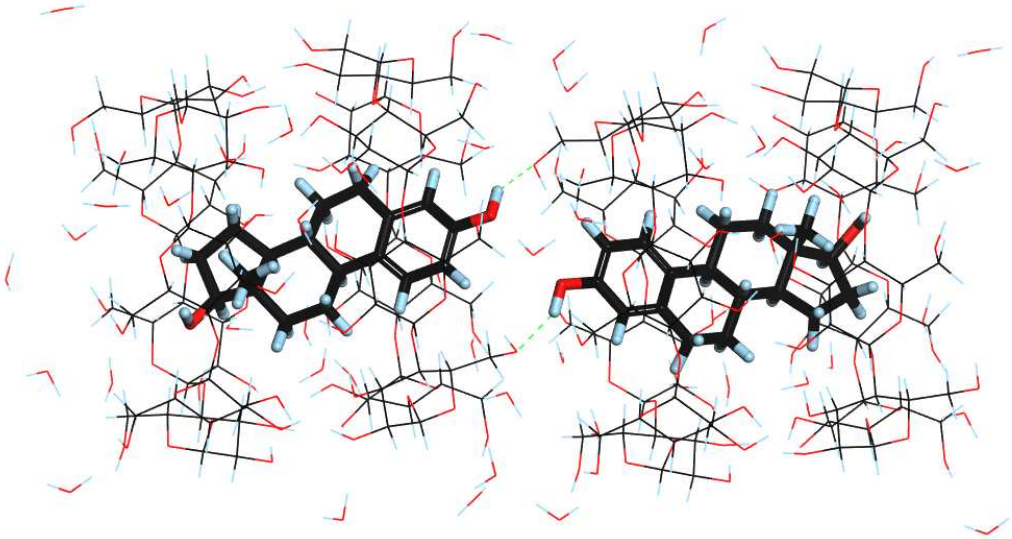


Figure 3. DFT-optimized structure of DAAD. Hydrogen bonds formed between EST and BCD are indicated as green dashed lines.

The other possible orientation was named ADAD, as in this one, the D ring of one molecule of EST is always located next to the A ring of the second EST molecule (Figure 4). This structure can also be found in the Supplementary Material as ADAD.cif. The unit cell dimensions of both DAAD and ADAD were exactly the same and can be found in Table 2.

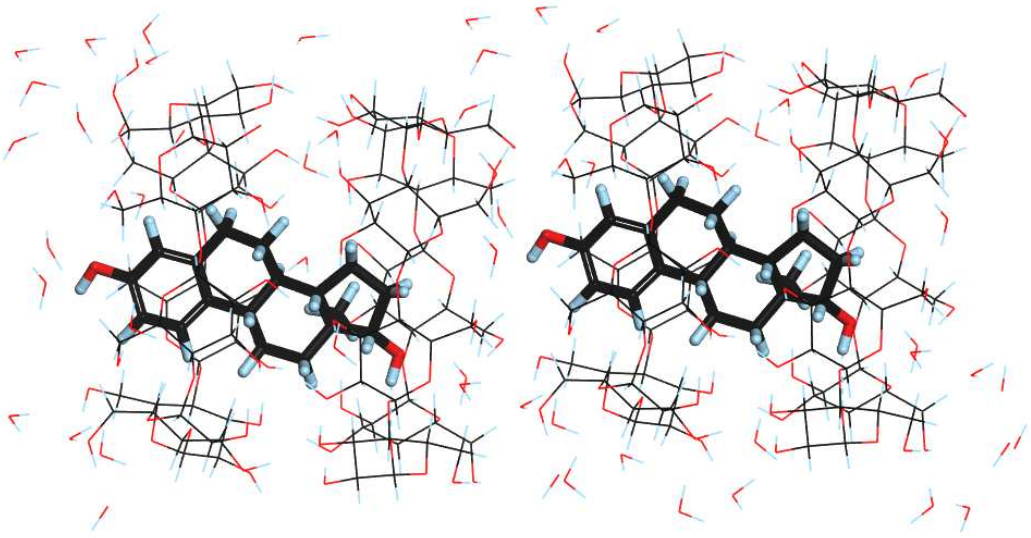


Figure 4. DFT-optimized structure of ADAD.

Table 2. Unit cell parameters and relative energy for the modeled structures, ADAD and DAAD, compared with the initial experimental structure used for geometry optimization (Exp.).

	Exp.	ADAD	DAAD
a [Å]	15.515011	15.742737	15.484059
b [Å]	15.515011	15.314397	15.484028
c [Å]	31.188400	31.718429	31.699391
α [°]	101.86701	100.93157	101.31461
β [°]	101.86701	99.83545	101.31466
γ [°]	103.84663	104.42976	101.44811
Relative energy [kcal/mol]		0	−5.152

It should be noted here that the DFT calculations performed in this study were done directly on the crystal structures, taking into account the periodicity of the studied system and, explicitly, water molecules. In addition, the initial structures for the geometry optimization calculations were taken directly from the SCXRD measurements, without adding any additional atoms. Therefore, the preparation of the structures for the calculations included solely removing chosen guest molecules and, in some cases, the water molecules if they were disordered over two neighboring positions. This resulted in the same stoichiometry of ADAD and DAAD, which enabled direct comparison of the energy of the studied systems. In addition, despite the symmetry found in both ADAD and DAAD, the structures were optimized without any constraints resulting from their corresponding crystal groups, with both of the structures being treated as P1 systems. This has been done purposely to enable the molecules to relax independently.

The results of the calculations (Table 2) show only slight changes in the unit cell dimensions resulting from the geometry optimization. Both the increase (i.e., 'a' for ADAD, 'c' for both of the structures) as well as the decrease (i.e., 'a' for DAAD and 'b' for both of the structures) of some unit cell lengths were observed. This is in agreement with the experimental SCXRD results, indicating that the structure was disordered and various guest orientations in the solid complexes exist. Despite similar unit cell dimensions, the optimized structures differed in their energies, indicating that DAAD is the more stable one by approximately 5 kcal/mol. Although no intermolecular interactions between guest molecules were observed in either of the structures, hydrogen bonds between the EST C3 hydroxyl group and primary hydroxyl groups of BCD were observed in ADAD (Figure 4). It should also be noted that after unit cell optimization, the symmetry in ADAD was no longer observed as the optimized 'a' and 'b' lengths differed significantly.

The next step in the DFT calculations was computation of the NMR chemical shielding constants for the optimized structures of ADAD and DAAD (which can be found in the Supplementary Material as ADADopt.cif and DAADopt.cif) using the GIPAW method. The isotropic chemical shielding values were then converted into chemical shifts to facilitate peak assignment of the NMR spectra and to compare the differences between the corresponding experimental and theoretical values obtained for the two optimized models.

2.3. ^{13}C CP MAS Solid State NMR Analysis

As described both in the introduction as well as in Section 4, in this study, the EST/ß-CD complexes were prepared using four different methods (LYS, STAND, MECH, STAND-SHORT). Detailed information on how exactly the samples were prepared can be found in Section 4.2.

To explore whether there are any structural differences between the complexes obtained in different ways, we have chosen the ^{13}C CP MAS solid state NMR analysis. The application of this method to the study of CD-based complexes has been recently reviewed

by us [16]. The spectra of the complexes (LYS, STAND, MECH, STANDSHORT) and reactants (EST, ß-CD) are presented in Figures 5–7. The spectra of the EST and ß-CD were recorded to facilitate the observation of changes in the chemical shifts and shapes of signals occurring upon complexation.

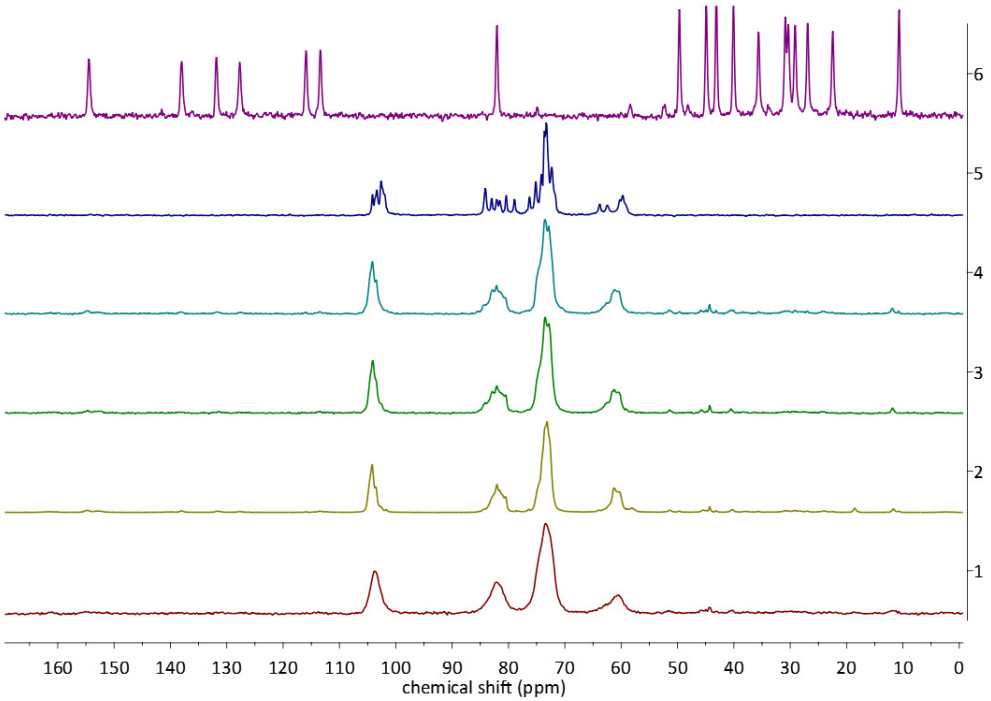


Figure 5. ^{13}C CP MAS NMR spectra of the EST (violet), ß-CD (dark blue), STAND (blue), STANDSHORT (green), MECH (olive green), and LYS (red).

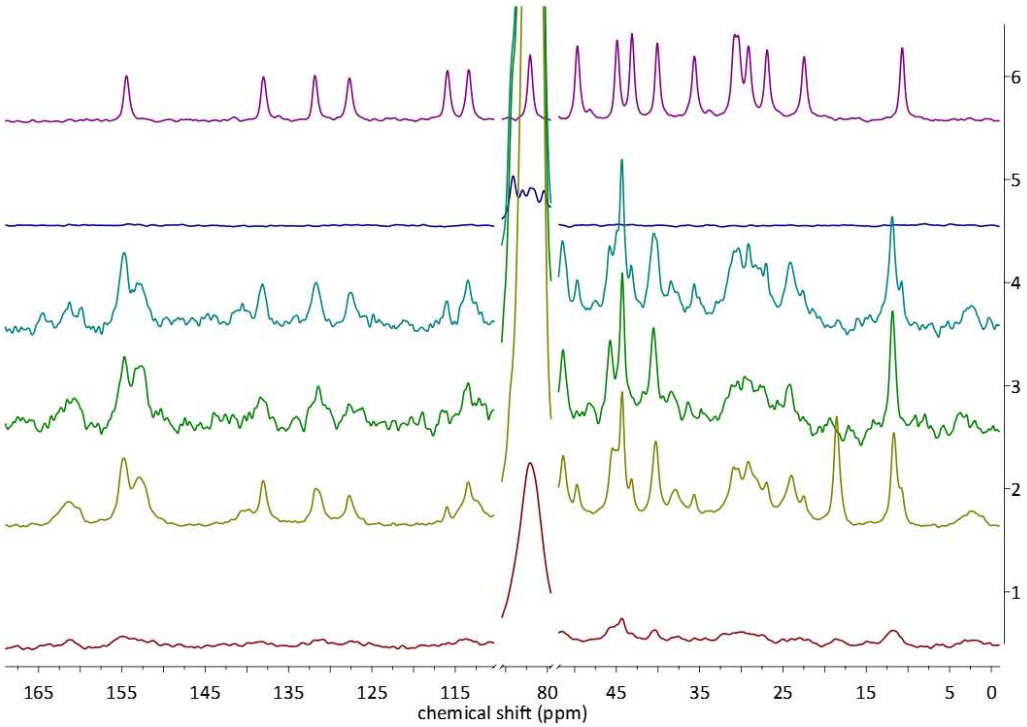


Figure 6. ^{13}C CP MAS NMR spectra of the EST (violet), ß-CD (dark blue), STAND (blue), STANDSHORT (green), MECH (olive green), and LYS (red). Only chosen regions of the spectra are presented for better visualization of the signals originating from EST.

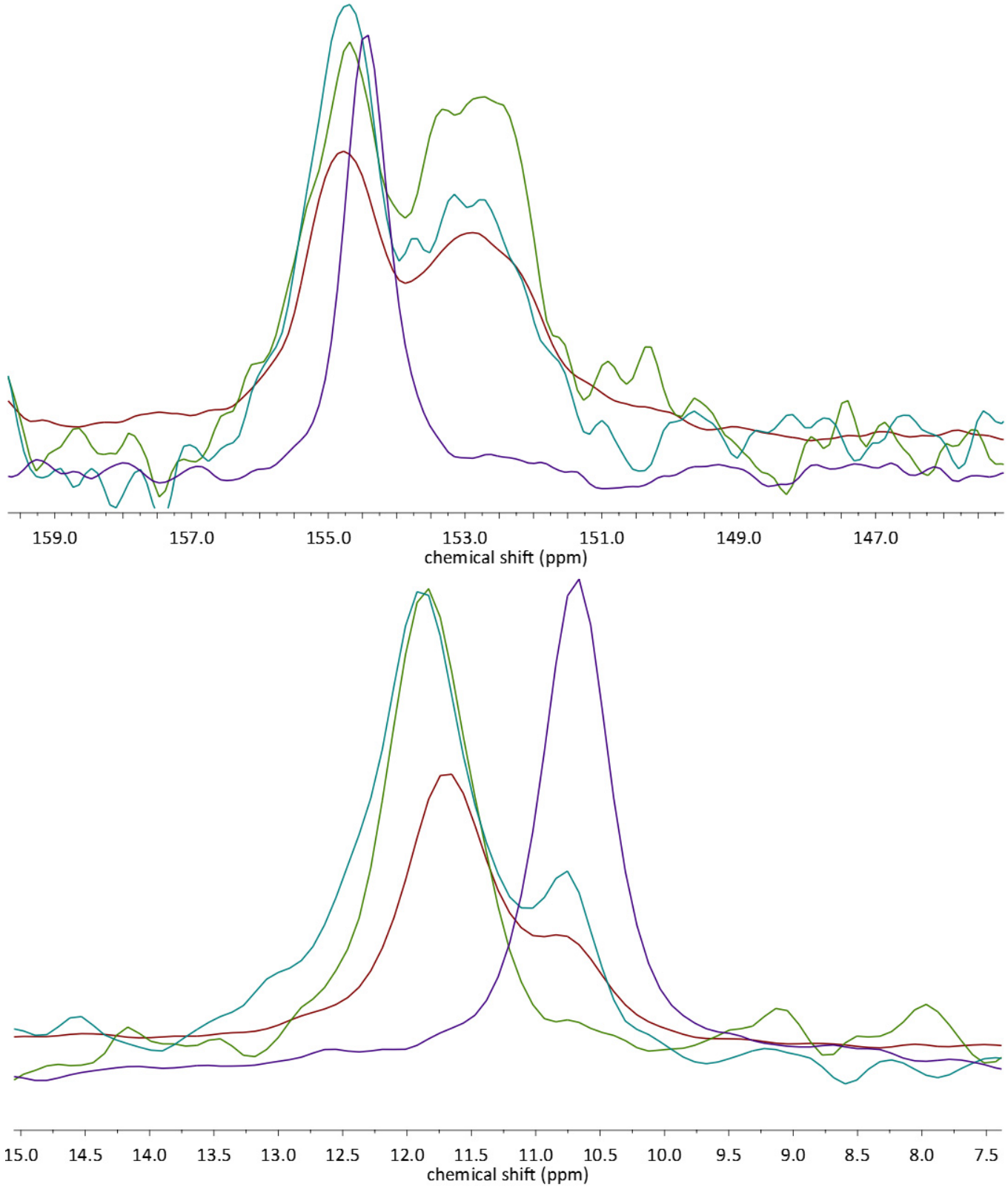


Figure 7. ^{13}C CP MAS NMR spectra of the EST (violet), STAND (blue), STANDSHORT (green), and MECH (red). Only chosen regions of the spectra are presented for better visualization of the signals originating from EST.

At first glance, in the stacked spectra, scaled in a way that the highest peaks have the same intensity (Figure 5), changes in the shape and number of signals originating from the

ß-CD are well visible. However, due to the significantly lower molecular mass of guest than host, the intensities of the signals of EST were very low. After increasing the intensity of the peaks (Figure 6), signals from EST have been revealed in all of the spectra of the complexes, with the exception of one obtained using the LYS method. In the LYS spectrum, signals originating from EST carbon atoms are either not visible or, in the best cases, very broad and low, i.e., in the 40–45 ppm region. This indicates that, as anticipated, the LYS method resulted in the amorphization of the sample, which was further confirmed by PXRD analysis (see Section 2.3).

In the discussion below, we will focus on the signals originating from the guest molecule, EST. This is justified for several reasons. First, the signals of EST, both in the non-complexed and in the EST/ß-CD forms, are usually well separated and sharp. Second, the chemical shifts of the EST carbon atoms occur in a wide range, 10–155 ppm, while all the signals from ß-CD can be found in a much wider range, 60–105 ppm. Moreover, the signals from the ß-CD are characterized by much larger FWHM and are highly overlapping; therefore, their analysis would not be possible without the ambiguous deconvolution.

We have started the NMR analysis from the 145–160 ppm range (Figure 7). In this region of the spectra, two overlapping peaks can be observed, with their maxima, respectively, at 152.9 and 154.8 ppm. Initially, after comparison with the spectrum of EST, in which a peak occurs at 152.8, we have assumed that these peaks originate from the complexed (154.42 ppm) and non-complexed (152.9 ppm) EST molecules. However, the intensities of the signals and areas under them were similar, which could mean that only around half of the EST was successfully complexed. Eventually, after careful analysis of the other regions of the spectra, 7.5–15 ppm (Figure 7), we have changed our initial assumptions. In this aliphatic region, the change in the chemical shift of the EST methyl group can be observed. The complexation resulted in the downfield shift of the single peak by c.a. 1 ppm, from 10.67 to 11.71 ppm. Still, even in some (MECH, STAND) of the spectra of complexes, the signal of the non-complexed EST methyl group could be observed. These observations allowed us to draw two conclusions. First, the two peaks in the spectra of complexes, located at 152.9 and 154.8 ppm, originate from the same carbon atom (C3) of the crystallographically nonequivalent EST molecules. This conclusion was also supported by the results of GIPAW calculations (Table 3), as the calculated chemical shifts for the C3 in ADAD and DAAD differ significantly. The other conclusion was that since a different amount of noncomplexed EST could be detected in the analyzed spectra, the yield of the complexation depends on the choice of the preparation method. As the ratio of the intensities of the peak at 10.67 to 11.71 was found to decrease in the order MECH → STAND → STANDSHORT, the yield was also decreasing in the same manner.

Upon complexation, the chemical shifts of some of the EST signals have only slightly changed (i.e., those from C1, C2, C4–C11). These carbon atoms form the A and B ring of EST and do not form any significant intermolecular interactions with other atoms; also, the conformation of these rings is highly rigid. The most apparent changes in the chemical shift values were observed for the signals occurring in the 21–53 ppm region (Figure 8). In the assignment of these signals, the results of GIPAW NMR calculations were found to be particularly useful. Additionally, the changes between the spectra of MECH, STAND, and STANDSHORT observed in the 10.67–11.71 ppm region (Figure 7) were found to be similar to those in the 21–53 ppm region. For example, the C14 signal has a 49.65 ppm shift in the spectrum of EST and 51.39 ppm in the spectrum of EST/ß-CD. In the spectra of MECH and STAND, the low-intensity signal from the noncomplexed EST can be observed, while in the spectrum of STANDSHORT, it is no longer visible. Similar observations were made for C15 signals, occurring at 22.45 and 23.98 ppm in the complexed and non-complexed forms, respectively (Figure 8).

Table 3. Experimental (exp.) and theoretically calculated (GIPAW) ^{13}C chemical shifts (δ) of EST and its complex with β-CD. Due to the presence of two EST molecules in the asymmetric unit of both ADAD and DAAD, two sets of values, (1) and (2), have been obtained for the first (1) and second (2) molecule present in the unit cell.

Atom Number	δ EST Exp.	δ EST GIPAW	δ EST Exp.—δ EST GIPAW	δ EST + ß-CD Exp.	δ ADAD GIPAW (1)	δ ADAD GIPAW (2)	(EST + ß-CD Exp.)—ADAD GIPAW (1)	(EST + βCD Exp.)—ADAD GIPAW (2)	δ DAAD GIPAW (1)	δ DAAD GIPAW (2)	(EST + ß-CD Exp.)—DAAD GIPAW (1)	(EST + ß-CD Exp.)—DAAD GIPAW (2)	δ EST Exp.—(EST/ß-CD Exp.)
1	127.65	129.59	−1.94	127.78	127.77	127.76	0.01	0.02	124.65	124.71	3.13	3.07	−0.13
2	113.33	111.18	2.15	113.43	109.86	109.82	3.57	3.61	109.61	109.65	3.82	3.78	−0.1
3	154.42	156.06	−1.64	154.8/152.9	157.21	157.23	−2.41	−2.43	155.72	155.71	−2.92	−2.91	1.62
4	115.89	118.35	−2.46	115.97	113.2	113.2	2.77	2.77	113.24	113.27	2.73	2.7	−0.08
5	137.94	139.1	−1.16	137.98	140.12	139.98	−2.14	−2.00	139.55	139.65	−1.57	−1.67	−0.04
6	30.83	30.75	0.08	30.92	28.79	28.81	2.13	2.11	28.58	28.56	2.34	2.36	−0.09
7	30.36	29.98	0.38	30.36	25.09	25.09	5.27	5.27	26.49	26.49	3.87	3.87	0.00
8	40.05	38.54	1.51	40.31	38	37.94	2.31	2.37	37.13	37.12	3.18	3.19	−0.26
9	44.89	43.52	1.37	45.5	45.24	45.2	0.26	0.30	43.73	43.74	1.77	1.76	−0.61
10	131.77	133.53	−1.76	131.84	132.39	132.31	−0.55	−0.47	130.46	130.52	1.38	1.32	−0.07
11	26.9	25.78	1.12	26.95	26.27	26.29	0.68	0.66	23.63	23.61	3.32	3.34	−0.05
12	35.59	36.07	−0.48	37.76	35.08	35.15	2.68	2.61	35.84	35.84	1.92	1.92	−2.17
13	43.09	42.62	0.47	44.3	43.5	43.38	0.8	0.92	42.08	42.07	2.22	2.23	−1.21
14	49.65	48.09	1.56	51.39	51.23	51.21	0.16	0.18	50.06	50.04	1.33	1.35	−1.74
15	22.45	21.06	1.39	23.98	20.61	20.61	3.37	3.37	21.45	21.43	2.53	2.55	−1.53
16	29.12	28.09	1.03	29.18	26.62	26.61	2.56	2.57	29.53	29.52	−0.35	−0.34	−0.06
17	82.04	82.63	−0.59	82.06	84.46	84.38	−2.4	−2.32	85.32	85.28	−3.26	−3.22	−0.02
18	10.67	7.54	3.13	11.71	8.46	8.5	3.25	3.21	8.54	8.52	3.17	3.19	−1.04

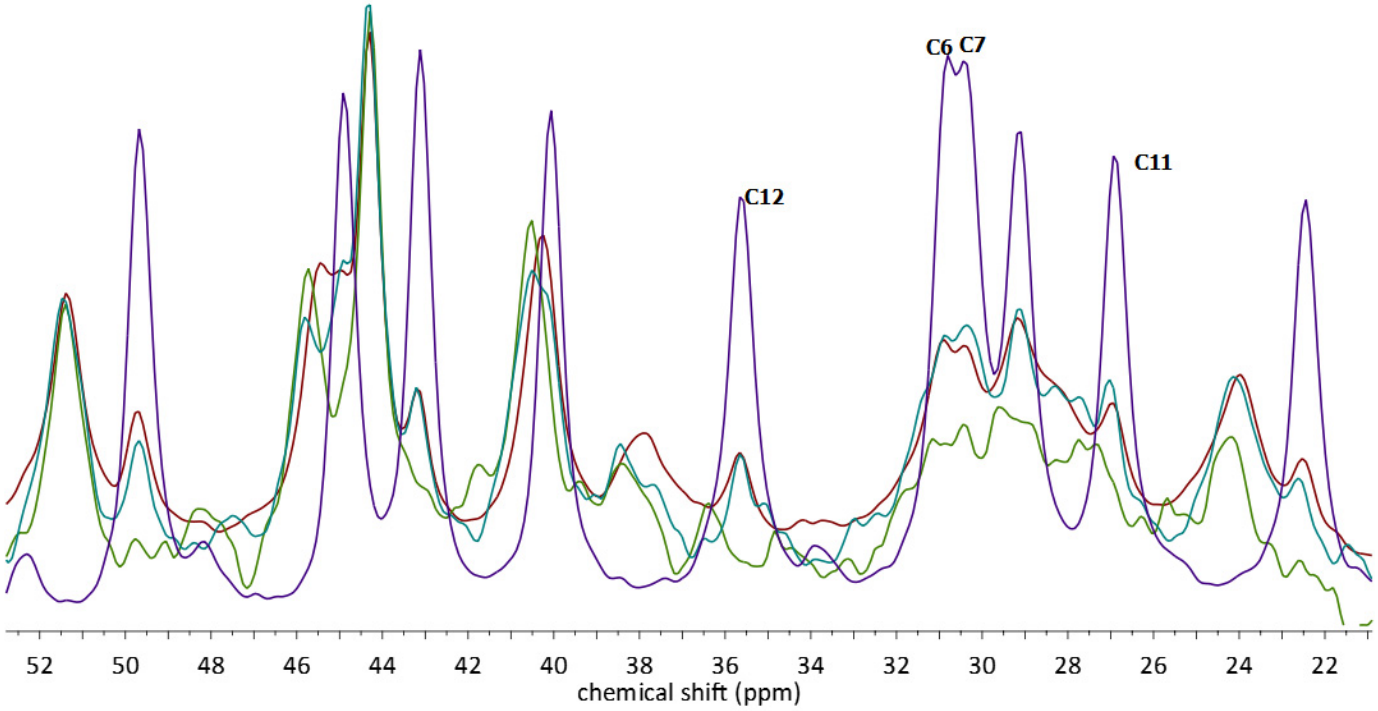


Figure 8. ^{13}C CP MAS NMR spectra of the EST (violet), STAND (blue), STANDSHORT (green), and MECH (red). Only chosen regions of the spectra are presented for better visualization of the signals originating from EST.

As mentioned above, the results of the GIPAW NMR calculations (Table 3) were found to be in very good agreement with corresponding experimental ones and facilitated proper signal assignment. The obtained differences between the experimental and theoretical values of δ for the complexes were found to be at a level similar to those of EST, not exceeding 4 ppm and, in most cases, below 3 ppm, with an exception for the C7 signal of ADAD. It should be noted, however, that in the spectra of the complexes, the peaks originating from C6, C7, C11, and C12 are overlapping and of a low intensity. This indicates the high level of dynamic disorder in this part of the guest molecule. All four atoms are chemically similar, as they are all secondary and form six-membered rings. Another explanation for this observation can be a dynamic of the C ring of EST. It has been reported previously that the C ring of EST can adopt either a chair or boat conformation, depending on its crystal form or, in the case of a solution, on the solvent [30]. It is therefore possible that the chair–boat conformational dynamics of the EST C ring can occur in the EST/ß-CD complex, which would explain the shape of the signals from C11 and C12.

The FT-IR spectra (Figure 9) of both ß-CD and EST have been found to be very similar to those reported previously [24]. The observed differences might have been caused by either the method of spectrum registration or the different degree of crystallinity. As in the case of the ^{13}C CP MAS NMR results, the FT-IR spectra of the complexes prepared by different methods have been found to be similar. However, there are also some noticeable differences among them. The signal at 3384 cm^{-1} is much narrower in LYS than in other cases. In addition, on the slope of the signal with a maximum at 2925 cm^{-1}, the small signals of EST, overlapped by the wide signal of ß-CD, are the most visible in the spectrum of STAND. Moreover, in the fingerprint area (500–850 cm^{-1}), the spectrum of LYS is flatter than the spectra of the complexes prepared by other methods. In all of the studied spectra of the complexes, broad signals in the range of 3000–3500 cm^{-1} can be found. These signals originate from multiple hydrogen bonds present in the studied system. These bonds differ in their length, energy, spectroscopic frequency, and intensity, as shown in previous works [31–34].

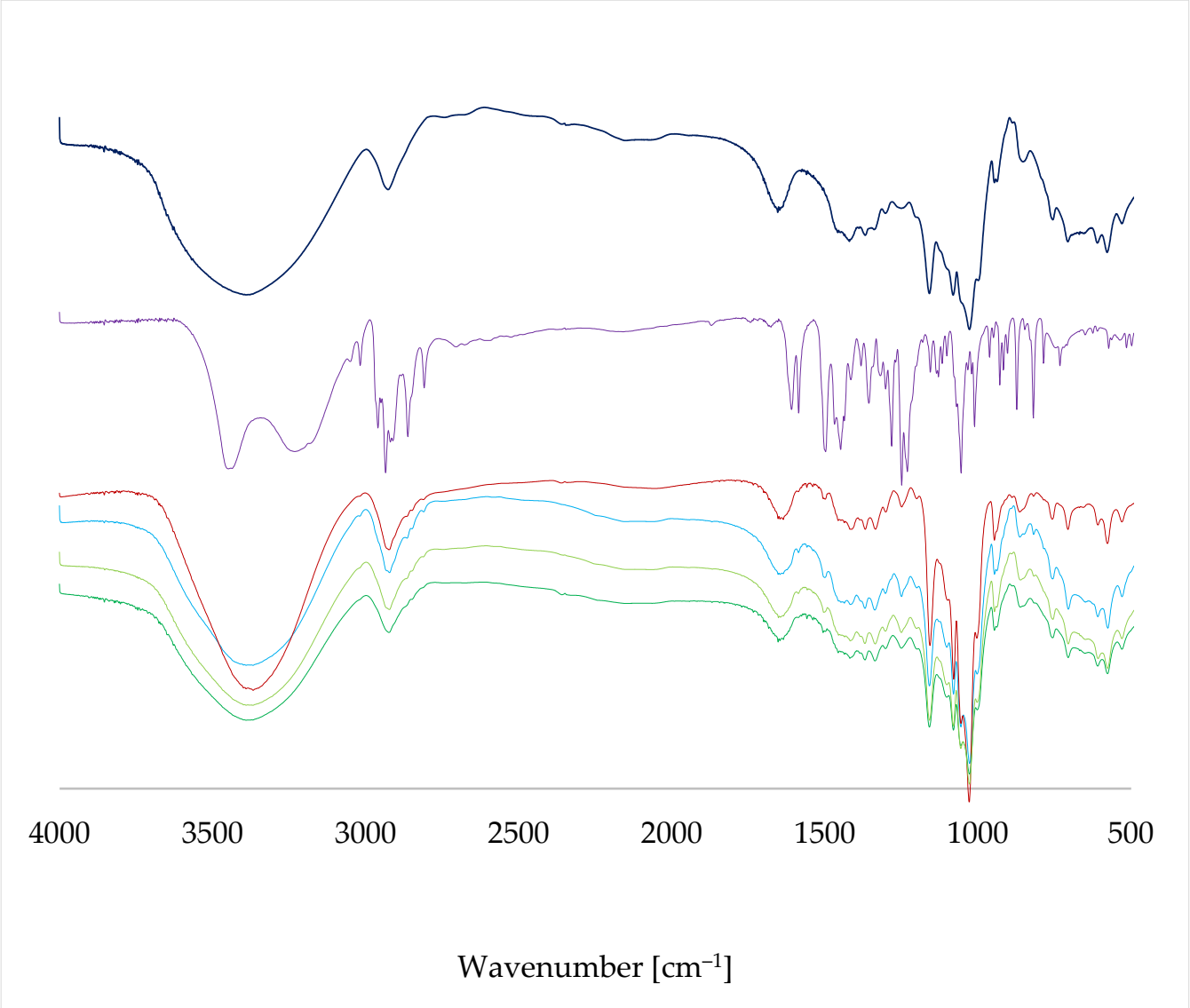


Figure 9. FT-IR spectra of ß-CD (dark blue), EST (violet), LYS (red), STAND (blue), MECH (olive green), and STANDSHORT (green).

Similar to the spectroscopic method results, the PXRD patterns (Figure 10) of the complexes are quite similar, with the exception of LYS. The lack of reflexes in the PXRD pattern of LYS and the characteristic halos indicate that the sample obtained by this method is amorphous. The PXRD patterns of STAND, MECH, and STANDSHORT are similar to the theoretical ones, simulated using the experimental crystal structure of EST/ß-CD. Characteristic reflexes can be found at the 2Θ values of 6.47, 7.25, 9.8, and 11.95 deg. In addition, two groups of signals in the ranges of 14.7–15.7 and 17.5–18.8 can be found both in the theoretical and experimental patterns. In the STAND, MECH, and STANDSHORT group, the pattern of STANDSHORT is slightly different than the other two. For example, in the PXRD pattern of STANDSHORT, there are no signals at 10.75 and 12.535, which are both present in the other two patterns. These reflexes are also present in the pattern of ß-CD, which indicates that in the samples of both STAND and MECH, some 'free' crystalline ß-CD can be found.

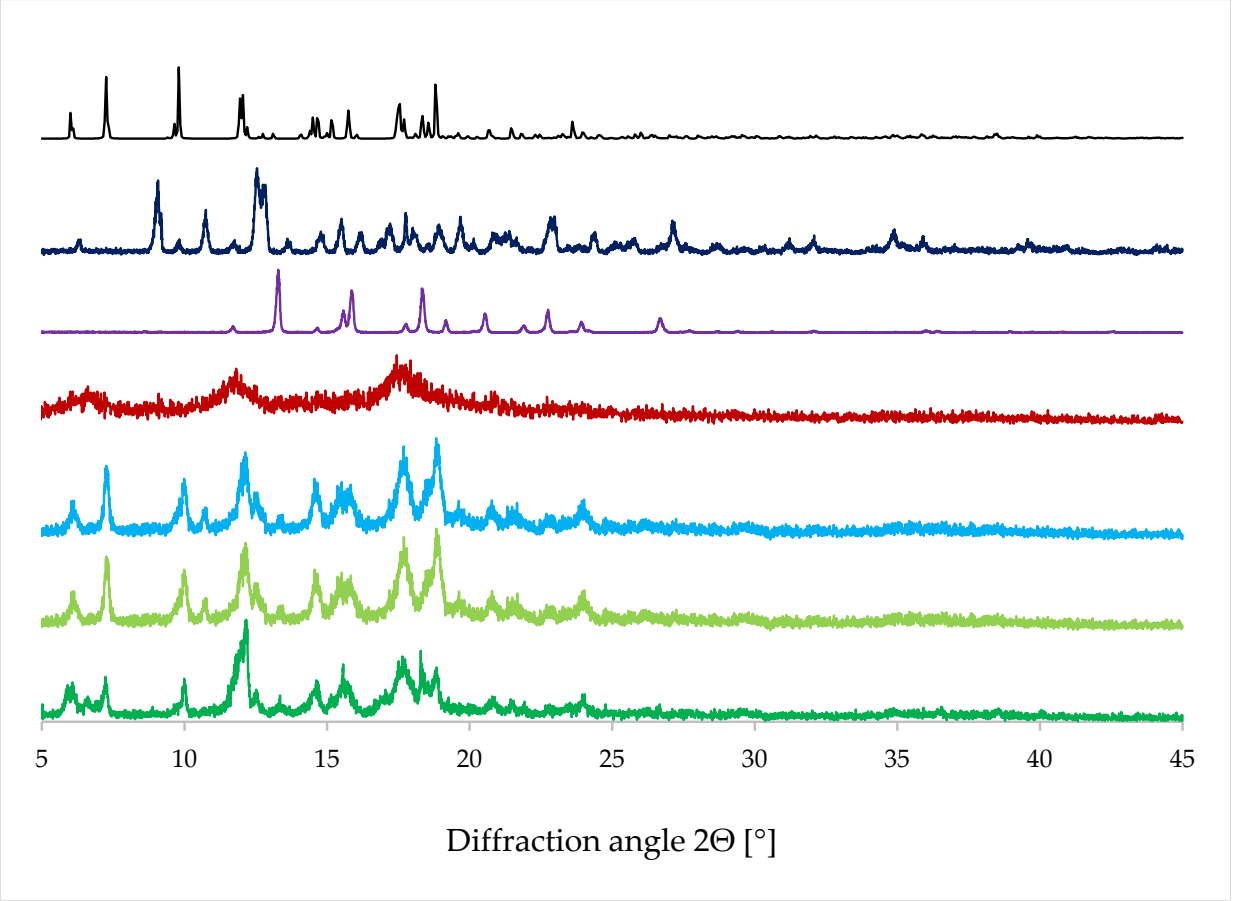


Figure 10. PXRD patterns: simulated for EST/ß-CD (black) and experimental PXRD prints for ß-CD (dark blue), EST (violet), LYS (red), STAND (blue), MECH (olive green), and STANDSHORT (green).

Comparison of the thermal analysis results (Table 4, Figures S2–S7) revealed additional differences between the complexes obtained using different preparation methods. The differences in mass loss during heating can be explained by the various ratios between the phases (EST:ß-CD:EST/ß-CD) in the analyzed samples. According to the SCXRD results, the total amount of water in EST and EST/ß-CD is 6.20% and 12.95%, respectively. The amount of water in BCD is variable [35], but usually within the 12.5–16% range. The TGA results for EST and ß-CD are in agreement with their corresponding crystal structures, indicating the total water loss in the analyzed temperature range. Lower than theoretically calculated values obtained for the EST/ß-CD complexes may indicate that some of the water molecules in the structure of EST/ß-CD are characterized by lower crystallographic occupancies and that the amount of crystal water in those complexes is variable. Moreover, the presence of non-complexed EST, with 6.20% water content, additionally decreases the anticipated values of water loss during the heating of the complexes. Significantly lower mass loss has been observed for the LYS sample, which can be explained by the final step of this method, lyophilization, which is used to decrease the amount of water in the sample. The DSC analysis of the complexes revealed two endothermic peaks in STANDSHORT and MECH, the first associated with the dehydration and the second with the decomposition of the sample. The higher enthalpy of dehydration was found in the sample with larger water loss, MECH. In the DSC thermograms of STAND and LYS, no clear endothermic peaks were observed.

Table 4. DSC-TGA analysis results.

STAND	DSC	(1) Onset temp. 29.26 °C Peak temp. 61.83 °C Enthalpy 83.82 J/g	(3) Onset temp. 152.0 °C Peak temp. 177.96 °C Enthalpy 10.74 J/g
		(2) Onset temp. 101.1 °C Peak temp. 111.27 °C Enthalpy 18.53 J/g	(4) Onset temp. 202.0 °C Peak temp. 212.69 °C Enthalpy 6.434 J/g
	TGA	Temp. range of dehydration: 27–200 °C Associated mass loss: 6.911%	
STANDSHORT	DSC	Onset temp. 26.38 °C Peak temp. 45.33 °C Enthalpy 185.1 J/g	
	TGA	Temp. range of dehydration: 25–200 °C Associated mass loss: 7.034%	
MECH	DSC	Onset temp. 30.55 °C Peak temp. 69.66 °C Enthalpy 209.8 J/g	
	TGA	Temp. range of dehydration: 20–200 °C Associated mass loss: 9.408%	
LYS	DSC	Onset temp. 41.0 °C Peak temp. 55.01 °C Enthalpy 47.53 J/g	
	TGA	Temp. range of dehydration: 25–200 °C Associated mass loss: 4.065%	
EST	DSC	(1) Onset temp. 81.88 °C Peak temp. 104.81 °C Enthalpy 19.70 J/g	
		(2) Onset temp. 175.91 °C Peak temp. 178.20 °C Enthalpy 91.06 J/g	
	TGA	Temp. range of dehydration: 20–200 °C Associated mass loss: 6.20%	
ß-CD	DSC	Onset temp. 56.31 °C Peak temp. 91.48 °C Enthalpy 380.8 J/g	
	TGA	Temp. range of dehydration: 20–100 °C Associated mass loss: 12.95%	

The surface morphology of EST/ß-CD complexes were assessed by SEM and the images are provided in Figure 11. As shown in Figure 11a, the SEM picture of STANDSHORT demonstrates a crystalline structure of this sample, dominated by cuboid-like crystals with average dimensions of about 100 µm. The SEM picture of MECH in Figure 11b presents an irregularly shaped crystalline structure. This sample is composed of crystals of different size, ranging from a few µm to a few hundred µm. Additionally, the crystals found in this sample are irregularly shaped. Meanwhile, the image of STAND (Figure 11c) shows a lot of large, prism-like crystals, larger even than those found in STANDSHORT, but with more irregular shapes. Finally, the LYS picture (Figure 11d) reveals the amorphous character of this sample, dominated by small particles of irregular shape, which is also in agreement with the PXRD results.

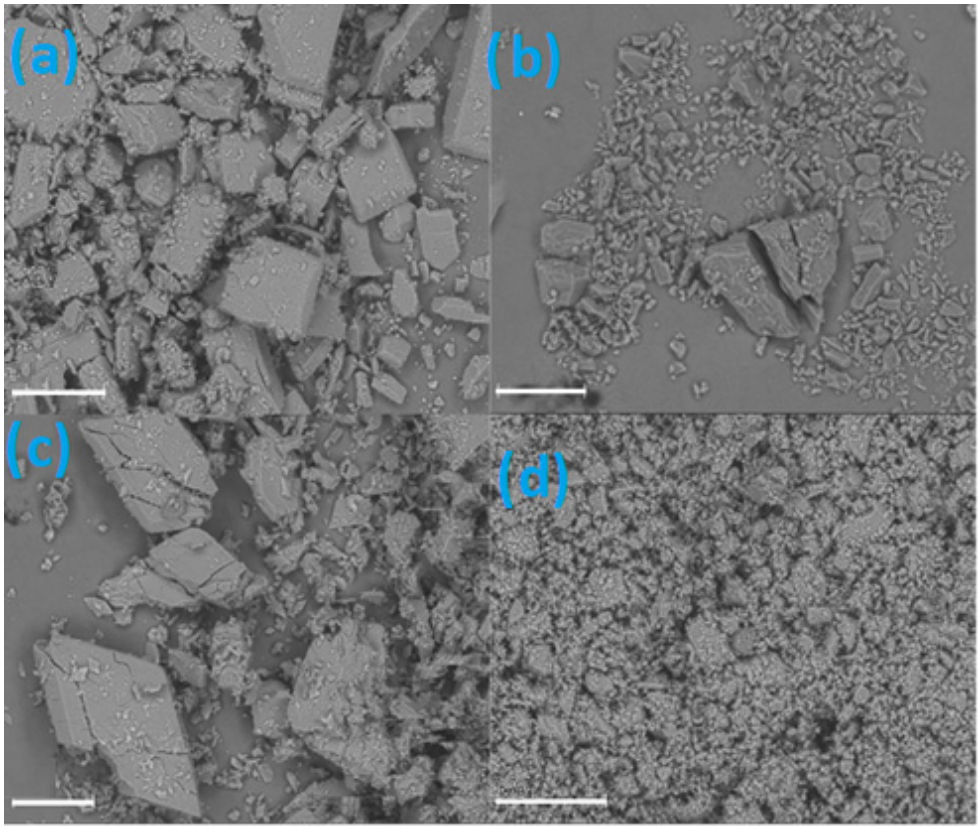


Figure 11. SEM images of STANDSHORT (**a**), MECH (**b**), STAND (**c**), and LYS (**d**). The length of the scale (white horizontal line) in each case is the same: 200 µm.

3. Conclusions

In this study, the inclusion complex of 17-β-estradiol and β-cyclodextrin has been prepared by four different methods. We have found that the method of complex preparation influences the final composition of the obtained sample as it affects the yield of complexation. However, regardless of the applied method, only one crystal form of the complex has been obtained, with an exception for a method involving lyophilization that resulted in the formation of an amorphous sample.

It should be noted that EST is an API with a long history of application in the treatment of various conditions, such as hormonal contraception, hormone replacement therapy (HRT), and treatment of menopausal and postmenopausal symptoms. ß-CD, on the other hand, is an excipient used in multiple original drugs currently on the worldwide market. It has also been shown that the complexation of EST with this particular cyclodextrin decreased toxicity of the studied hormone [23]. Moreover, the combination of EST and ß-CD was found to improve the bioavailability of the API by increasing its solubility [28].

The crystal structure of the 17-β-estradiol/β-cyclodextrin complex has been obtained for the first time by means of SCXRD. The 2:1 stoichiometry of host:guest has been determined. It has also been found that the highly disordered encapsulated 17-β-estradiol molecule has a unique orientation inside the dimeric host cavity. Using the solved crystal structure, periodic DFT calculations have been conducted to assess the energy differences between the two modeled structures. The GIPAW NMR calculations for the optimized structures facilitated peak assignment in the ^{13}C CP MAS NMR spectra. The ssNMR results confirmed that in the crystal structure of the studied complex, two orientations of 17-β-estradiol exist, as for the C3, two signals in the spectra were observed, indicating that only some of the hydroxyl groups of this carbon atom form hydrogen bonds.

SEM and thermal (TGA/DSC) analysis revealed noticeable differences between the complexes obtained using various methods. PXRD analysis confirmed the formation of the

complex in each case, with the exception of LYS, as this sample was proven to be amorphous. No major differences in the FT-IR spectra of the complexes obtained by different methods were observed.

The fact that, despite using numerous methods to obtain the studied complex, only one form has been received may indicate that only one stable crystal form is present in normal conditions. This, however, does not exclude the possibility of polymorphism at other temperature or pressure conditions, especially since in the studied structure significant disorder has been detected, which usually indicates the possibility of polymorphic phase transition.

This work highlights the importance of a combined ssNMR/SCXRD approach to studying the structure of the inclusion complexes formed by cyclodextrins, especially those characterized by a high level of structural disorder.

4. Materials and Methods

4.1. Materials

17-β-estradiol hemihydrate and β-cyclodextrin were purchased from BIOSYNTH; Biosynth AG, Staad, Switzerland) and used as received, without any further purification. For the mechanochemistry (MECH) method, freshly prepared, twice-distilled Milli-Q (Mq) water (Milli-Q water purification system, Millipore Corp., Waltham, MA, USA), with a conductivity of ~1 μS/cm, was used in the grinding process.

4.2. Methods of Complex Preparation

In this study, four different EST/ß-CD complex preparation methods were applied.

The first one was the **STANDARD [STAND]** method, commonly used in cases of CD inclusion complexes, when crystals of sufficient quality for SCXRD measurement are required. The STANDARD method is a slow-cooling crystallization technique: 60 mg of ß-CD was mixed in a flask with 1 mL distilled water and put into 70 °C water for 20 s to obtain a clear solution. Then the contents of the flask were poured into a beaker. In accordance with the molar mass of ß-CD and EST, the respective amount of EST was added to the beaker to maintain the 1:1 molar ratio. The beaker was put on a magnetic stirrer and left at room temperature for 15–20 min until a clear solution was obtained. Afterwards, the contents of the beaker were poured into a glass tube. The beaker was poured along with 0.5–1.0 mL water, which was also added to the glass tube. The glass tube was held in 70 °C water for 20 s to obtain a clear solution. Later, the tube was closed and put into a 70 °C water bath. A slow, gradual cooling process was performed over 10 days, reaching a temperature of 24 °C on the 10th day. At the end, a rotary evaporator was used.

The second one was the **STANDARD SHORT [STANDSHORT]** method: ß-CD and EST in a molar ratio of 1:1 were mixed with distilled water and put into a round bottom flask. The flask was closed and left on a magnetic stirrer for 24 h at a temperature of 44 °C. To obtain crystals, a rotary evaporator was used.

The third one was the **MECHANOCHEMICAL [MECH]** method: ß-CD and EST in a molar ratio of 1:1 were mixed in a mortar with 3–5 drops of Mq water and knitted for 5 min every day over 5 consecutive days.

The fourth one was the **LYOPHILIZATION [LYS]** method: the **STANDARD SHORT** crystallization method with the application of lyophilization instead of slow evaporation. A solution, which was obtained in accordance with the **STANDSHORT** method, was poured into smaller containers in which the lyophilization process took place. Firstly, the probes were frozen with the application of liquid nitrogen. Secondly, the probes were put into a lyophilizator for 48 h.

4.3. Powder X-ray Diffraction (PXRD)

For the PXRD measurements, a Panalytical Empyrean (Malvern, UK) diffractometer was used. The samples were analyzed in Bragg–Brentano reflection mode, using Cu-Kα radiation (λ = 1.54187 Å), a 2Θ range of 3–45°, and a 0.006565° step size. For the incident

beam a fixed divergence slit of 1/16°, an anti-scatter slit of 1/4°, and a fixed mask of 4 mm were used, and the diffracted beam path was equipped with a 7.5 mm anti-scatter slit.

4.4. Single-Crystal X-ray Diffraction (SCXRD)

A clear, light, colorless prism-like specimen, with dimensions of about 0.130 mm × 0.270 mm × 0.400 mm and coated with paraffin oil as cryo-protectant, was used for the X-ray crystallographic analysis. The X-ray intensity data were measured at 100 (2) K with a Bruker (Billerica, MA, USA) D8-VENTURE diffractometer using Cu *Kα* radiation (λ = 1.54178 Å). A low-temperature device (Oxford Cryosystems Ltd., Long Handorough, UK) provided a continuous stream of nitrogen vapor on the specimen during the data collection, while diffraction patterns were recorded using a CMOS-PHOTON III detector.

The total exposure time was 29.33 h. The frames were integrated with the Bruker SAINT software package [36] using a narrow-frame algorithm. The integration of the data using a monoclinic unit cell yielded a total of 133,843 reflections to a maximum θ angle of 76.27 (0.79 Å resolution). The final cell constants of a = 19.1245(15) Å, b = 24.4180(18) Å, c = 15.6004(11) Å, β = 109.500(5)°, and volume = 6867.2(8) Å^3 were based upon the refinement of the XYZ-centroids of 9073 reflections above 20 σ(I) with 6.094 < 2θ < 148.5°. Data were corrected for absorption effects using the Multi-Scan method (SADABS) [36].

The structure was solved by the intrinsic phasing method with SHELXT [37] and refined by full-matrix least squares against F^2 using SHELXL-2014/7 [38] through the SHELXLE GUI [39]. H-atoms were placed geometrically and refined in riding mode with isotropic displacement parameters fixed by SHELXL. Due to the structural complexity and disorder of the final model, soft restraints on bond lengths and angles, generated from the PRODRG2 webserver [40], were applied on the host and guest molecules of the inclusion complexes. Anisotropic displacement parameters were refined using soft restraints (SIMU) [41] implemented in the SHELXL program, where necessary. The final anisotropic full-matrix least-squares refinement on F^2 with 950 variables converged at R_1 = 9.5% for the observed data and wR_2 = 26.9% for all data. The goodness of fit (GoF) was 1.04. The crystallographic data along with the structural refinement details are summarized in Table 1. The data can be obtained from the Cambridge Crystallographic Data Centre under the reference number 2250781.

Geometric features of the crystal structure, e.g., interatomic distances, angles, dihedral angles, centroid coordinates, and mean plane equations through various groups of atoms, along with their e.s.d. estimations, were calculated via the full covariance matrix using the Olex2 program [42]. The final 3D model was drawn with Mercury 4.3.1 [43] and PyMoL [44].

4.5. Fourier-Transform Infrared Spectroscopy (FT-IR)

The studies were performed using a Perkin-Elmer Spectrum 1000 FT-IR spectrometer equipped with an MTEC 300 detector (MTEC, Ames, IA, USA). The samples were packed in ring cups with a diameter of 10 mm. The detector's chamber was purged with helium to reduce the effect of moisture evaporating from the samples during measurement. A spectrum obtained from the background sample was subtracted from the spectrum of each sample to eliminate the residual peaks of CO_2 and moisture. For each sample, 1024 scans were recorded and averaged in the infrared region between 4000 and 500 cm^{-1} at a resolution of 4 cm^{-1}. All spectral plots were prepared using the GRAMS/AI 8.0 Spectroscopy Software.

4.6. Cryo-Scanning Electron Microscopy (Cryo-SEM)

The analysis was performed using low-temperature scanning electron microscope ZEISS AURIGA (Warsaw, Poland) 60 coupled with a focused ion beam. Cryo-SEM allows sample observation without chemical fixing or drying. The procedure consists of sample freezing by immersion in liquid nitrogen, breaking, and etching.

4.7. Differential Scanning Calorimetry and Thermogravimetry Analysis (DSC-TGA)

Analysis was performed using apparat SDT Q600 (TA Instruments, New Castle, DE, USA) under nitrogen flow. The heating rate was equal to 10 °C/min and the sample mass was approximately 6–8 mg. Pierced aluminum sample pans were used in the analysis and the temperature range was set to 0–500 °C.

4.8. ^{13}C CP MAS NMR

Solid-state ^{13}C CP/MAS NMR spectra were recorded on a Bruker Avance III 600 spectrometer (Bruker BioSpin, Rheinstetten, Germany) operating at 600.15 MHz (^{1}H) and 150.91 MHz (^{13}C), and powder samples were spun at 12 kHz in a 4 mm ZrO_2 rotor using a double air-bearing probe head. Acquisition was performed with a standard CP pulse sequence with ramped CP scheme, 2 ms CP contact time, 4 s recycle delay, and a swept-frequency two-pulse phase modulation decoupling scheme, using a 3.2 µs proton 90° pulse. The decoupling field strength was set to 78 kHz. A total of 256 scans were acquired, 10.00 exponential apodization, a receiver gain equal to 2050, and 2048 acquired points. After zero filling and LP application, the spectrum size was 4096. ^{13}C chemical shifts were referenced to adamantane CH_2 at 38.48 ppm.

4.9. Periodic DFT Calculations

The density functional theory (DFT) calculations of geometry optimization and NMR parameters, under periodic boundary conditions, were carried out with the CASTEP program [45] implemented in the Materials Studio 2020 software [46] using the plane wave pseudopotential formalism. On-the-fly-generated ultrasoft pseudopotentials were generated using a Koelling–Harmon scalar relativistic approach [47]. The Perdew–Burke–Ernzerhof (PBE) [48] exchange-correlation functional, defined within the generalized gradient approximation, with Tkatchenko–Scheffler (TS) [49] dispersion correction, was used in the calculations.

4.9.1. Geometry Optimization

Geometry optimization was carried out using the limited memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS) [50] optimization scheme and smart method for finite basis set correction. The kinetic energy cutoff for the plane waves (E_{cut}) was set to 630.0 eV. The number of Monkhorst–Pack k-points during sampling for a primitive cell Brillouin zone integration [51] was set to 2 × 2 × 1 (for EST-ß-CD) and 1 × 1 × 2 (for 17-β-estradiol hemihydrate, refcode ESTDOL10), respectively.

The experimental X-ray structure of EST/ß-CD was used to create two initial periodic structures for calculations, containing two EST and four ß-CD molecules in the unit cell each. Details on the structural preparation can be found in Section 2.2.

During geometry optimization, all atom positions and cell parameters were optimized, with no constraints. The convergence criteria were set at 1×10^{-5} eV/atom for the energy, 3×10^{-2} eV/Å for the interatomic forces, 5×10^{-2} GPa for the stresses, and 1×10^{-3} Å for the maximum displacement. A fixed-basis set quality method for the cell optimization calculations and a 1×10^{-6} eV/atom tolerance for SCF were used.

4.9.2. NMR Parameter Calculations

The computation of shielding tensors was performed using the Gauge Including Projector Augmented Wave Density Functional Theory (GIPAW) method of Pickard et al. [52]. To compare the theoretical and experimental data, the calculated chemical shielding constants (σiso) were converted to chemical shifts (δiso) using the following equation: δiso = (σGly + δGly)—σiso, where σGly and δGly stand for the shielding constant and the experimental chemical shift, respectively, of the glycine carbonyl carbon atom (176.50 ppm).

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/molecules28093747/s1: SI.pdf file containing Table S1: Geometric

parameters of the ß-CD host molecule in the EST/ß-CD crystal structure; Table S2: Main H-bonds in the crystal structure of EST/ß-CD; Figure S1: Radar plots of some geometrical parameters of Table S1; Figure S2: DSC/TGA curve of LYS; Figure S3: DSC/TGA curve of EST; Figure S4: DSC/TGA curve of β-CD; Figure S5: DSC/TGA curve of STAND; Figure S6: DSC/TGA curve of MECH; Figure S7: DSC/TGA curve of STANDSHORT; DAAD.cif; ADAD.cif; DAADopt.cif; ADADopt.cif; estradiol_beta-CD.cif; and checkcif-1.pdf.

Author Contributions: Conceptualization, A.H.M., Ł.S., K.B. and D.M.P.; methodology, A.H.M., Ł.S., K.B., E.C., M.K.D., M.Z.-P. and D.M.P.; software, A.H.M., Ł.S., K.B., E.C. and D.M.P.; validation, A.H.M., Ł.S., K.B., M.K.D., M.Z.-P. and D.M.P.; formal analysis, A.H.M., Ł.S., K.B., E.C., M.K.D., M.Z.-P. and D.M.P.; investigation, A.H.M., Ł.S., K.B., E.C. and D.M.P.; resources, A.H.M., Ł.S. and K.B.; data curation, A.H.M., Ł.S. and K.B.; writing—original draft preparation, A.H.M., Ł.S. and K.B.; writing—review and editing, A.H.M., Ł.S., K.B., M.K.D. and D.M.P.; visualization, A.H.M., Ł.S., K.B. and E.C.; supervision, A.H.M. and Ł.S.; project administration, A.H.M. and Ł.S.; funding acquisition, A.H.M. and Ł.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Medical University of Warsaw, Poland, grant number F/MB/02/22.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data can be obtained from the corresponding author (Ł.S.) by email. Crystallographic data have been deposited with the Cambridge Structural Database (CSD) under deposition number CCDC: 2250781.

Acknowledgments: The authors acknowledge the support of the National and Kapodistrian University of Athens (NKUA) Core Facilities and Nikolaos Tsoureas (Department of Chemistry, NKUA) for the SCXRD data collection.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Sample Availability: Samples of the compounds are available from the corresponding author (Ł.S.) by email.

References

1. Thomas, M.P.; Potter, B.V. The structural biology of oestrogen metabolism. *J. Steroid Biochem. Mol. Biol.* **2013**, *137*, 27–49. [CrossRef] [PubMed]
2. Grandi, G.; Napolitano, A.; Cagnacci, A. Metabolic impact of combined hormonal contraceptives containing estradiol. *Expert Opin. Drug Metab. Toxicol.* **2016**, *12*, 779–787. [CrossRef]
3. Santoro, N.; Epperson, C.N.; Mathews, S.B. Menopausal Symptoms and Their Management. *Endocrinol. Metab. Clin. N. Am.* **2015**, *44*, 497–515. [CrossRef] [PubMed]
4. Daulbayev, C.; Kaidar, B.; Sultanov, F.; Bakbolat, B.; Smagulova, G.; Mansurov, Z. The recent progress in pitch derived carbon fibers applications. A Review. *S. Afr. J. Chem. Eng.* **2021**, *38*, 9–20. [CrossRef]
5. Guo, A.; Gong, X.; He, J.; Guo, Y.; Ning, L.; Chen, X.; Xu, J.; Guo, Y.; Wang, H. Study on Co-crystals of Estradiol. *Her. Med.* **2021**, *40*, 1716–1723.
6. Wang, J.R.; Wang, X.; Yang, Y.; Chen, X.; Mei, X. Solid-state characterization of 17β-estradiol co-crystals presenting improved dissolution and bioavailability. *CrystEngComm* **2016**, *18*, 3498–3505. [CrossRef]
7. Ning, L.; Gong, X.; Li, P.; Chen, X.; Wang, H.; Xu, J. Measurement and correlation of the solubility of estradiol and estradiol-urea co-crystal in fourteen pure solvents at temperatures from 273.15 K to 318.15 K. *J. Mol. Liq.* **2020**, *304*, 112599. [CrossRef]
8. Kovacs, T.; Nagy, P.; Panyi, G.; Szente, L.; Varga, Z.; Zakany, F. Cyclodextrins: Only Pharmaceutical Excipients or Full-Fledged Drug Candidates? *Pharmaceutics* **2022**, *14*, 2559. [CrossRef] [PubMed]
9. European Medicines Agency. Available online: https://www.ema.europa.eu/en (accessed on 21 February 2023).
10. U.S. Food & Drug Administration (FDA). Available online: https://www.fda.gov (accessed on 21 February 2023).
11. Pharmaceutical and Medical Devices Agencty. Available online: https://www.pmda.go.jp/english/index.html (accessed on 21 February 2023).
12. Sun, J.; Hong, H.; Zhu, N.; Han, L.; Suo, Q. Effect of preparation methods on tosufloxacin tosylate/hydroxypropyl-β-cyclodextrin inclusion complex. *Braz. J. Pharm. Sci.* **2022**, *58*, e18650. [CrossRef]

13. Cid-Samamed, A.; Rakmai, J.; Mejuto, J.C.; Simal-Gandara, J.; Astray, G. Cyclodextrins inclusion complex: Preparation methods, analytical techniques and food industry applications. *Food Chem.* **2022**, *384*, 132467. [CrossRef]
14. Deckmann Nicoletti, C.; de Sá Haddad Queiroz, M.; de Souza Lima, C.G.; de Carvalho da Silva, F.; Futuro, D.O.; Ferreira, V.F. An improved method for the preparation of β-lapachone:2-hydroxypropyl-β-cyclodextrin inclusion complexes. *J. Drug Deliv. Sci. Technol.* **2020**, *58*, 101777. [CrossRef]
15. Wdowiak, K.; Rosiak, N.; Tykarska, E.; Żarowski, M.; Płazińska, A.; Płaziński, W.; Cielecka-Piontek, J. Amorphous Inclusion Complexes: Molecular Interactions of Hesperidin and Hesperetin with HP-B-CD and Their Biological Effects. *Int. J. Mol. Sci.* **2022**, *23*, 4000. [CrossRef] [PubMed]
16. Mazurek, A.H.; Szeleszczuk, Ł. A Review of Applications of Solid-State Nuclear Magnetic Resonance (ssNMR) for the Analysis of Cyclodextrin-Including Systems. *Int. J. Mol. Sci.* **2023**, *24*, 3648. [CrossRef] [PubMed]
17. Mazurek, A.H.; Szeleszczuk, Ł. Current Status of Quantum Chemical Studies of Cyclodextrin Host–Guest Complexes. *Molecules* **2022**, *27*, 3874. [CrossRef] [PubMed]
18. Mazurek, A.H.; Szeleszczuk, Ł.; Gubica, T. Application of Molecular Dynamics Simulations in the Analysis of Cyclodextrin Complexes. *Int. J. Mol. Sci.* **2021**, *22*, 9422. [CrossRef]
19. Mazurek, A.H.; Szeleszczuk, Ł.; Pisklak, D.M. Periodic DFT Calculations—Review of Applications in the Pharmaceutical Sciences. *Pharmaceutics* **2020**, *12*, 415. [CrossRef]
20. Gallez, A.; Palazzo, C.; Blacher, S.; Tskitishvili, E.; Noël, A.; Foidart, J.M.; Evrard, B.; Pequeux, C.; Piel, G. Liposomes and drug-in-cyclodextrin-in-liposomes formulations encapsulating 17β-estradiol: An innovative drug delivery system that prevents the activation of the membrane-initiated steroid signaling (MISS) of estrogen receptor α. *Int. J. Pharm.* **2020**, *573*, 118861. [CrossRef]
21. Schwarz, D.H.; Engelke, A.; Wenz, G. Solubilizing steroidal drugs by β-cyclodextrin derivatives. *Int. J. Pharm.* **2017**, *531*, 559–567. [CrossRef]
22. Cai, W.; Yao, X.; Shao, X.; Pan, Z. Bimodal Complexations of Steroids with Cyclodextrins by a Flexible Docking Algorithm. *J. Incl. Phenom. Macrocycl. Chem.* **2005**, *51*, 41–51. [CrossRef]
23. Silva, M.C.G.D.; Silva, J.F.D.; Santos, T.P.; Silva, N.P.C.D.; Santos, A.R.D.; Andrade, A.L.C.; Souza, E.H.L.D.S.; Sales Cadena, M.R.; Sá, F.B.; Silva Junior, V.A.D.; et al. The complexation of steroid hormones into cyclodextrin alters the toxic effects on the biological parameters of zebrafish (Danio rerio). *Chemosphere* **2019**, *214*, 330–340. [CrossRef]
24. Haimhoffer, Á.; Vas, A.; Árvai, G.; Fenyvesi, É.; Jicsinszky, L.; Budai, I.; Bényei, A.; Regdon, G., Jr.; Rusznyák, Á.; Vasvári, G.; et al. Investigation of the Drug Carrier Properties of Insoluble Cyclodextrin Polymer Microspheres. *Biomolecules* **2022**, *12*, 931. [CrossRef]
25. Lin, Z.H.; Wang, X.X.; Kou, S.B.; Shi, J.H. Exploring the inclusion interaction of estradiol with β-CD and HP-β-CD with the help of molecular dynamics simulation as well as multi-spectroscopic approaches. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2022**, *269*, 120764. [CrossRef] [PubMed]
26. Sadlej-Sosnowska, N. Fluorometric determination of association constants of three estrogens with cyclodextrins. *J. Fluoresc.* **1997**, *7*, 195–200. [CrossRef]
27. van Uden, W.; Woerdenbag, H.J.; Pras, N. Cyclodextrins as a useful tool for bioconversions in plant cell biotechnology. *Plant Cell Tiss Organ Cult.* **1994**, *38*, 103–113. [CrossRef]
28. Vicatos, A.I.; Hoossen, Z.; Caira, M.R. Inclusion complexes of the steroid hormones 17β-estradiol and progesterone with β- and γ-cyclodextrin hosts: Syntheses, X-ray structures, thermal analyses and API solubility enhancements. *Beilstein J. Org. Chem.* **2022**, *18*, 1749–1762. [CrossRef]
29. Mentzafos, D.; Mavridis, I.M.; Le Bas, G.; Tsoucaris, G. Structure of the 4-It Tert-Butylbenzyl Alcohol–β-Cyclodextrin Complex. Common Features in the Geometry of β-Cyclodextrin Dimeric Complexes. *Acta Crystallogr. Sect. B* **1991**, *47*, 746–757. [CrossRef]
30. Commodari, F.; Sclavos, G.; Ibrahimi, S.; Khiat, A.; Boulanger, Y. Comparison of 17beta-estradiol structures from X-ray diffraction and solution NMR. *Magn. Reson. Chem.* **2005**, *43*, 444–450. [CrossRef]
31. Rekik, N.; Issaoui, N.; Ghalla, H.; Oujia, B.; Wójcik, M.J. Infrared spectral density of H-bonds within the strong anharmonic coupling theory: Indirect relaxation effect. *J. Mol. Struct.* **2007**, *844–845*, 21–31. [CrossRef]
32. Brela, M.Z.; Klimas, O.; Surmiak, E.; Boczar, M.; Nakajima, T.; Wójcik, M.J. Comparison of the Hydrogen Bond Interaction Dynamics in the Guanine and Cytosine Crystals: Ab Initio Molecular Dynamics and Spectroscopic Study. *J. Phys. Chem. A* **2019**, *123*, 10757–10763. [CrossRef]
33. Rekik, N.; Issaoui, N.; Ghalla, H.; Oujia, B.; Wójcik, M.J. IR spectral density of H-bonds. Both intrinsic anharmonicity of the fast mode and the H-bond bridge. Part I: Anharmonic coupling parameter and temperature effects. *J. Mol. Struct. Theochem.* **2007**, *821*, 9–21. [CrossRef]
34. Wójcik, J.M. Theoretical Modeling of Vibrational Spectra and Proton Tunneling in Hydrogen-Bonded Systems. *Adv. Chem. Phys.* **2016**, *160*. [CrossRef]
35. Pereva, S.; Nikolova, V.; Angelova, S.; Spassov, T.; Dudev, T. Water inside β-cyclodextrin cavity: Amount, stability and mechanism of binding. *Beilstein J. Org. Chem.* **2019**, *15*, 1592–1600. [CrossRef]
36. Bruker. *APEX 3, SAINT, SADABS*; Bruker AXS Inc.: Madison, WI, USA, 2012.
37. Sheldrick, G.M. It SHELXT—Integrated Space-Group and Crystal-Structure Determination. *Acta Crystallogr. Sect. A* **2015**, *71*, 3–8. [CrossRef] [PubMed]

38. Sheldrick, G.M. Crystal Structure Refinement with It SHELXL. *Acta Crystallogr. Sect. C* **2015**, *71*, 3–8. [CrossRef] [PubMed]
39. Hübschle, C.B.; Sheldrick, G.M.; Dittrich, B. It ShelXle: A Qt Graphical User Interface for It SHELXL. *J. Appl. Crystallogr.* **2011**, *44*, 1281–1284. [CrossRef]
40. Schüttelkopf, A.W.; van Aalten, D.M.F. PRODRG: A Tool for High-Throughput Crystallography of Protein–Ligand Complexes. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 1355–1363. [CrossRef]
41. Thorn, A.; Dittrich, B.; Sheldrick, G.M. Enhanced Rigid-Bond Restraints. *Acta Crystallogr. Sect. A Found. Crystallogr.* **2012**, *68*, 448–451. [CrossRef]
42. Dolomanov, O.V.; Bourhis, L.J.; Gildea, R.J.; Howard, J.A.K.; Puschmann, H. It OLEX2: A Complete Structure Solution, Refinement and Analysis Program. *J. Appl. Crystallogr.* **2009**, *42*, 339–341. [CrossRef]
43. Macrae, C.F.; Bruno, I.J.; Chisholm, J.A.; Edgington, P.R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P.A. It Mercury CSD 2.0—New Features for the Visualization and Investigation of Crystal Structures. *J. Appl. Crystallogr.* **2008**, *41*, 466–470. [CrossRef]
44. *The Pymol Molecular Graphics System*, Version 1.8; Schrödinger, Inc.: New York, NY, USA, 2015.
45. Clark, S.J.; Segall, M.D.; Pickard, C.J.; Hasnip, P.J.; Probert, M.J.; Refson, K.; Payne, M.C. First principles methods using CASTEP. *Z. Krist.-Cryst. Mater.* **2005**, *220*, 567–570. [CrossRef]
46. BIOVIA. Materials Studio. Available online: http://accelrys.com/products/collaborative-science/biovia-materials-studio (accessed on 23 February 2023).
47. Koelling, D.D.; Harmon, B.N. Technique for relativistic spin-polarized calculations. *J. Phys. C Solid State Phys.* **1977**, *10*, 3107–3114. [CrossRef]
48. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868. [CrossRef] [PubMed]
49. Tkatchenko, A.; Scheffler, M. Accurate Molecular van der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.* **2009**, *102*, 073005. [CrossRef]
50. Packwood, D.; Kermode, J.; Mones, L.; Bernstein, N.; Woolley, J.; Gould, N.; Ortner, C.; Csányi, G. A universal preconditioner for simulating condensed phase materials. *J. Chem. Phys.* **2016**, *144*, 164109. [CrossRef] [PubMed]
51. Monkhorst, H.J.; Pack, J.D. Special points for Brillouin-zone integrations—A reply. *Phys. Rev. B* **1977**, *16*, 1748–1749.
52. Pickard, C.J.; Mauri, F. All-electron magnetic response with pseudopotentials: NMR chemical shifts. *Phys. Rev. B* **2001**, *63*, 63–77. [CrossRef]

Journal of Molecular Structure 1313 (2024) 138710

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: www.elsevier.com/locate/molstr

17-β-Estradiol—β-Cyclodextrin complex as an aqueous solution: Structural and physicochemical characterization supported by MM and QM calculations

Anna Helena Mazurek [a,b], Łukasz Szeleszczuk [a,*], Kostas Bethanis [c], Elias Christoforides [c], Marta Katarzyna Dudek [d], Ewelina Wielgus [d], Dariusz Maciej Pisklak [a]

[a] Department of Organic and Physical Chemistry, Faculty of Pharmacy, Medical University of Warsaw, Banacha 1 Str., 02-093 Warsaw, Poland
[b] Doctoral School, Medical University of Warsaw, Żwirki i Wigury 81 Str., 02-093 Warsaw, Poland
[c] Laboratory of Physics, Department of Biotechnology, Agricultural University of Athens, 11855 Athens, Greece
[d] Structural Studies Department, Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, Sienkiewicza 112 Str., 90-363 Łódź, Poland

ARTICLE INFO

Keywords:
Estradiol
Cyclodextrin
DFT calculations
Complex stability constant

ABSTRACT

17-β-estradiol (EST) is an Active Pharmaceutical Ingredient characterized by a low water solubility. Complexation with β-cyclodextrin (βCD) enhances its bioavailability, hence such complex is an interesting research object from pharmaceutical point of view. However, basic facts like description of complex's structure and definition of its molar ratio, were debatable already for decades. This work for the first time justifies the EST:βCD molar ratio as 1:2 using the HRMS (high-resolution mass spectrometry) and phase solubility studies. The latter are used to define complex stability constant, as well. The structure and stability is analyzed using a variety of computational approaches: Quantum Mechanics (QM) based methods (DFT, semiempirical approaches) and MD/MMGBSA approach. In case of the QM, for the first time in the computational analysis of cyclodextrin complexes, a thorough benchmarking test is presented. Different computational parameters (solvent model, presence/absence of dispersion correction etc.) are used. Obtained results are compared with the experimental data.

1. Introduction

Cyclodextrins (CDs) are oligosaccharides of a donut-like structure, which enables them to form inclusion complexes with non-polar substances. This characteristic is used by the pharmaceutical industry but at the same time it often poses a non-trivial questions about the structure and stability of the created complexes.

More precisely, CDs are cyclic structures composed of glucose subunits, joined by α−1,4 glycosidic bonds [1]. Because of the cyclic character of CDs, various chemical compounds can enter CD's void and this way inclusion complexes are created. Depending on the size of a chemical guest, different types of CDs are preferred. However, the most common one is a medium size beta-CD (βCD) which consists of 7 glucose units [1,2] (Fig. 1).

Due to the presence of hydroxyl groups, the external fragments of CDs are polar. When a non-polar substance enters the molecular hole of a CD, the formed host–guest complex is polar and more water soluble than a separate non-complexed guest molecule [2,3]. Therefore, knowing also that CDs are non-toxic for a human organism [4], CDs are commonly used in the pharmaceutical industry in order to increase the solubility of a complexed Active Pharmaceutical Ingredient (API) or protect it from external factors like light, humidity or heat [5-7]. An important API group characterized by a poor solubility in water are hormonal steroids. Encapsulation in CDs enhance their solubility in water and as a result also their bioavailability.

An example of the steroid hormones used as a medication is estradiol (EST) (Fig. 1). EST belongs to the estrogens group and is the most potent estrogen naturally produced by a human body [8]. Therefore, the first complexes between EST and a well-known βCD have been obtained already decades ago [9]. However, the structural analysis of such complexes happens not to be as straightforward as one would assume.

In 1997 in the article entitled ''Fluorometric Determination of Association Constants of Three Estrogens with Cyclodextrins'' the EST-βCD complex molar ratio has been defined as 1:1 [10]. This information has been used as a reference for instance in the following articles [11,12] by other scientists. Only two decades afterwards, this pre-defined EST-βCD

* Corresponding author.
E-mail address: lukasz.szeleszczuk@wum.edu.pl (Ł. Szeleszczuk).

https://doi.org/10.1016/j.molstruc.2024.138710
Received 19 February 2024; Received in revised form 3 May 2024; Accepted 21 May 2024
Available online 24 May 2024

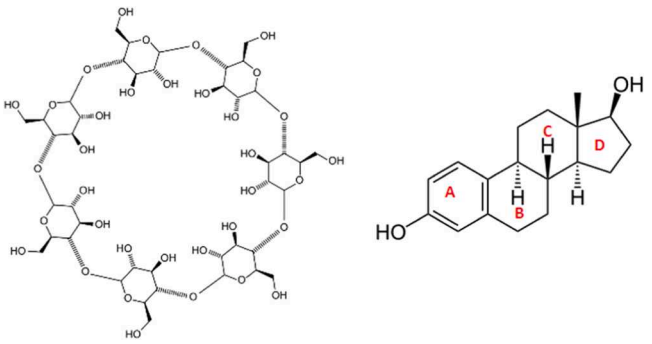


Fig. 1. Structures of β-CD and EST.

complex molar ratio and its stability constant have been revisited [13, 14]. However, again only either fluorescence or UV spectroscopy have been used for that purpose. Despite quite a few studies in this area, the 1:1 molar ratio of the complex in question has been accepted as a status quo and has never been questioned or verified by application of any other experimental methods such as high-resolution mass spectrometry (HRMS).

Recently, we have analyzed the EST-βCD complex in a solid state. For the first time, we have obtained and analyzed the structure of this complex's crystal structure. At the same time, it was one of the first ever obtained crystal structures of a steroid hormone complexed with any CD. The results have been published as [15] and the crystal structure has been deposited in the Cambridge Crystallographic Data Centre (CCDC). The results of the structural analysis in a solid state clearly indicate the 1:2 EST-βCD molar ratio, however they have also revealed significant structural disorder of this complex. Having obtained such results for a solid state, we were curious what is the molar ratio in the aqueous solution. Knowing the solid state structure of the complex, its 1:2 molar ratio after dissolution in water seems to be a scientifically sound hypothesis. Nevertheless, as we have gathered data on the previously performed experimental analyses repeatedly pointing out the 1:1 molar ratio, we decided against performing fluorescence or UV measurements and decided to concentrate on the molecular modelling techniques and HRMS spectrometry.

Complexes of various compounds with CDs have been analyzed using different molecular modelling approaches. At the beginning, computational methods used to predict the CD complexes structures and properties were the same as ones used for the analysis of much larger systems like receptors with ligands. Those were mainly Molecular Mechanics (MM) based methods and among them the most popular was molecular docking, sometimes followed by atomistic molecular dynamics (MD) simulations, at the same level of theory. However, in the recent years, with the increase of the computational power, such medium size systems like CD complexes started to be commonly analyzed using techniques based on the Quantum Mechanics (QM).

The conclusions of our recent review article [16] clearly show that the QM calculations era in the analysis of the CD complexes has already started more than a decade ago. The results show that the most commonly used are semi-empirical and Density Functional Theory (DFT) approaches. Among the former, PM6 and PM7 seem to be the most widely spread, also suggesting that they deliver the most appreciated results. When it comes to the DFT, a wider variety of computational options is used. The most frequent is application of B3LYP or M062X functionals both of the coming with or without dispersion correction. In terms of a solvent representation both possible approaches are practiced: either no solvent model is used or an implicit solvent model is applied.

As opposed to QM calculations, the Molecular Mechanics (MM) approach allows analysis of much larger systems but at the same time it delivers results of a significantly lower accuracy. However, for years CD complexes were too big for conducting QM calculations and Molecular Dynamics (MD) was the only option to obtain any information about the complexes' inner structure. Therefore, in the literature there are numerous examples of MD's application in the analysis of CD inclusion complexes. Even now this method is still in use. Therefore, we wanted to apply this well-described type of simulations in our study and compare them with the QM approaches.

Having gathered all the data on the previously conducted experimental analyses of the EST-βCD complex and having in mind our results concerning the solid state, we decided to try to determine the structure and the molar ratio of this complex. The aim of this study was to perform a benchmark studies on this topic using semi-empirical and DFT computational approaches. Additionally, we have managed to experimentally reveal the true complex's molecular ratio using a technique which has been never used in this particular case.

2. Materials and methods

2.1. Sample preparation

The EST-βCD complex was obtained by a method which is commonly used to obtain the CD inclusion complexes, a slow-cooling crystallization technique. This approach has already been used by us in a previous study describing the SCXRD analysis [15]. 60 mg of βCD was mixed in a flask with 1 mL distilled water and put into 70 °C water for 20 s to obtain a clear solution. Then the contents of the flask were poured into a beaker. In accordance with the molar mass of βCD and EST, the respective amount of EST was added to the beaker to maintain the 1:1 molar ratio. The beaker was put on a magnetic stirrer and left at room temperature for 15–20 min until a clear solution was obtained. Afterwards, the contents of the beaker were poured into a glass tube. The beaker was poured along with 0.5–1.0 mL water, which was also added to the glass tube. The glass tube was held in 70 °C water for 20 s to obtain a clear solution. Later, the tube was closed and put into a 70 °C water bath. A slow, gradual cooling process was performed over 10 days, reaching a temperature of 24 °C on the 10th day. At the end, a rotary evaporator was used.

2.2. HRMS

HRMS measurements were performed using Synapt G2-Si mass spectrometer (Waters) equipped with an ESI source and quadrupole-time-of-flight mass analyser. The mass spectrometer was operated in the positive ion detection mode. The optimized source parameters were: capillary voltage 3.0 kV, cone voltage 50 V, source temperature 110 °C, desolvation gas (nitrogen) flow rate 650 L/h with the temperature 450 °C, nebulizer gas pressure 6.5 bar. All samples were dissolved in water–methanol solution (1:1) and infused through a standard electrospray ion source into the instrument. The scan range was m/z 500–4000 and the acquisition method run time was 2 min. Mass calibration was performed using a cesium iodide solution. To ensure accurate mass measurements, data were collected in centroid mode and mass was corrected during acquisition using leucine enkephalin solution as an external reference (Lock-SprayTM) which generated reference ion at m/z 556.2771 Da ($[M + H]+$) in positive ESI mode. The results of the measurements were processed using the MassLynx 4.1 software (Waters) incorporated with the instrument.

2.3. Phase solubility studies

The UV–visible (UV–Vis) spectrophotometer (BioBase BK-S380, China) was utilized to assess the properties of EST and its inclusion complex. The EST showed a visible absorption peak at 280 nm. To generate a calibration curve, five standard solutions of EST in methanol were measured, each replicated three times. The concentrations of the standards used were 0.1, 0.25, 0.50, 0.75, and 1.00 mM.

Phase solubility studies were carried out following the procedure

outlined by Higuchi & Connors (1965) [17]. An excess amount of EST (50 mg) was added to 10 mL of deionized water containing various concentrations ranging from 0.20 to 20.00 mM for β-CD. The mixtures were then subjected to agitation using an orbital shaker (PHOENIX Instrument Laboratory Shaker RS-OS 5; Berlin, Germany) at 25 °C for 48 h to reach equilibrium. Afterwards, the collected samples were filtered through a 0.45 μm filter and assayed using a UV–Vis spectrophotometer at 280 nm.

2.4. QM calculations

All of the QM calculations were performed using the Gaussian 16 software [18]. All electron DFT computations were done employing the 6–311G(d,p) basis set and B3LYP or M062X functional, while the semiempirical calculations have been done using PM6 and PM7 approaches. Those four methods were used either with or without Grimme's dispersion force corrections (D3). All of the calculations were performed either in vacuo or using one of the implicit solvation models: Polarizable Continuum Model (PCM) [19] or SMD (Solvation Model Density) [20], each time choosing water as the solvent with dielectric constant 78.540. For the details of computational models, please see Table 1 and Table 2. At the review stage, additional calculations have been performed at the ωB97X-D/6–31G(d,p)-PCM-Water level [21].

Vibrational frequencies were calculated to estimate thermodynamic parameters, including Zero Point Vibrational Energy (ZPVE) and Gibbs free energy (ΔG) at 298.15 K and 101.325 kPa.

According to our recent review, different types of QM approaches are commonly used when modeling CDs inclusion complexes. In this work we decided to use those methods which are the most commonly encountered in the recent literature. The goal was to compare the geometrically optimized 1:2 and 1:1 systems from a quantitative and qualitative perspective. For this reason, we have analyzed the structural aspects as well as values of the energy ΔE and Gibbs free energy ΔG defined as

$$\Delta E = E_{com} - (E_{EST} + nE_{CD}) \quad (1)$$

where E_{com} is energy of the EST-βCD complex, E_{EST} is the energy of EST, E_{CD} is the energy of βCD and n is the number of βCD molecules forming the complex, in this case $n = 1$ or $n = 2$

$$\Delta G = G_{com} - (G_{EST} + nG) \quad (2)$$

where G_{com} is free enthalpy of the EST-βCD complex, G_{EST} is the free enthalpy of EST, G_{CD} is the free enthalpy of βCD and n is the number of βCD molecules forming the complex, in this case $n = 1$ or $n = 2$

2.5. Molecular dynamics (MD) simulations

In addition to the crystallographically determined structure of the EST/β-CD inclusion complex, which is characterized by a host: guest ratio of 2:1 and a head-to-head inclusion mode, two additional models of the same complex assuming host: guest ratio of 1:1 but two different inclusion modes (down and up) were investigated through MD simulations. The initial coordinates for these latter two complexes were generated using molecular docking with Autodock Vina [22], where the simulation boxes were defined around the coordinates of the CD centers with size 40 Å in each direction and a grid spacing of 0.375 Å. The Lamarckian genetic algorithm in AutoDockTools was utilized for this purpose, enabling effective management of numerous degrees of freedom. The 3D structure of estradiol was retrieved from the PubChem database (Compound CID: 5757) [23], while the coordinates of β-CD correspond to its crystal structure in complex with EST [15]. The docking runs were set to 10 and the produced models with the most favorable binding energies for each one of the two inclusion modes were chosen for subsequent analysis.

The AMBER 12 software package [24] was used for the simulation of the EST/β-CD inclusion complexes in a aqueous environment. Three simulations were performed. In the first case, the stating 3D model was provided by the crystallographically determined atomic coordinates of a β-CD dimer[15], which includes one EST guest molecule (site A) inside the formed dimeric cavity. Thus, the host: guest stoichiometry of the entire system in the simulation was 2: 1. In the other two cases, monomers of EST/β-CD inclusion complexes with different inclusion modes from docking analysis were used as the starting models.

The geometry of EST was optimised following the AM1BCC methodology with the program Antechamber [25]. xLeaP, the GUI version of AMBER's LeaP program, was utilized for system preparation. The GLYCAM-06j [26] force field, which is suitable for β-CD atoms' treatment and the generalized AMBER (GAFF) (for the guest molecule) were applied for the simulation. Additionally, the TIP3P water model [27] was used to solvate the CD dimer in a periodic, octahedral box forming a 12 Å thick water shell around the structure.

Minimization and MD calculations that resulted in a single trajectory of the hydrated inclusion complex system were performed with Sander. The particle mesh Ewald summation approach [28] was followed in order to handle the long-range electrostatic interactions with a 10 Å cut-off limit for the direct space sum. Hydrogen bonds were handled using the SHAKE algorithm [29]. The simulation protocol was as following: Energy minimization for hydrogens and waters using 1000 steps of steepest descent (SD) followed by 500 steps of conjugated gradient (CG) methods, while the rest non-hydrogen atoms were fixed with positional restraints of 50 kcal mol 1 Å 2. Heating equilibration up to 300 K of the water in the canonical (NVT) ensemble for 50 ps using positional restraints and the Berendsen thermostat algorithm with coupling constants of 0.5 ps to control temperature and pressure. Energy minimization of all system atoms with weak positional restraints (10 kcal mol 1 Å 2), gradual temperature increase from 5 to 300 K with 10 kcal mol-1 Å-2 restraints on the atoms of the system followed by gradual release of the restraints in successive steps at 300 K in NVT ensemble and finally density equilibration in the isobaric-isothermal (NPT) ensemble for 250 ps. Subsequently, production runs of the system under the NPT ensemble using a Berendsen-type algorithm with coupling constants of 1.0 ps were carried out under physiological conditions until reaching 12 ns.

The MD outputs were processed through the cpptraj module [30] of AMBER 12, to calculate the structural analyses (RMSD, distances, H-bonding). Moreover, the total guest binding energy ΔGbind, including the entropic term (ΔS) (calculated with the nmode module of AMBER 12) and the analysis of its components ΔEMM (changes in the gas-phase molecular mechanics (MM) energy) and ΔGsolv (solvation free energy) were computed with the aid of MM/PBSA.py script [31] implemented in AMBER 12. VMD [32] was also used for visualization and structural analyses of the MD trajectories.

Table 1
DFT computational approaches.

1	2	3	4	5	6	7	8	9	10	11	12	13
B3LYP In vacuo	M062X In vacuo	B3LYP-D3 In vacuo	M062X-D3 In vacuo	B3LYP PCM	M062 PCM	B3LYP-D3 PCM	M062X-D3 PCM	B3LYP SMD	M062X SMD	B3LYP-D3 SMD	M062X-D3 SMD	ωB97X-D PCM

Table 2
Semi-empirical computational approaches.

A	B	C	D	E	F	G	H	I	J	K	L
PM7	PM7-D3	PM6	PM6-D3	PM7	PM7-D3	PM6	PM6-D3	PM7	PM7-D3	PM6	PM6 -D3
In vacuo	In vacuo	In vacuo	In vacuo	PCM	PCM	PCM	PCM	SMD	SMD	SMD	SMD

3. Results and discussion

3.1. HRMS

As it was already mentioned, in our previous article concerning the EST-βCD complex in the solid state, we reported that the guest:host molar ratio in a crystal form is 1:2 [15]. However, we were interested whether 1:2 is the only form present in the aqueous solution of the complex, as well. Hence, we have defined 3 possible scenarios. The first option was presence of just 1:1 molar ratio complex, as it was stated many times in the literature. The second option was presence of both 1:1 and 1:2 complexes, where the 1:2 ratio would be the effect of the increasing concentration happening due to the solvent evaporation. According to that hypothesis, the amount of 1:2 complex should increase with the decrease of water content, resulting in the solely 1:2 complex in the solid state. And finally, the third possibility was presence of the identical stoichiometry as observed in a solid state which is 1:2 molar ratio.

To verify which scenario is correct, we have used the high-resolution mass spectrometry (HRMS). So far, this method has not been used often to define the molar ratio of the cyclodextrin complexes, however there are already some examples of its successful application for this purpose, published in the recent years [33,34]. The advantage of this method over the UV or fluorescence spectroscopy is that HRMS delivers a direct answer about the complex's molar ratio.

The results of the HRMS measurement of EST-βCD complex are presented in Fig. 2 and Table 3. EST forms stable associate with two βCD which can be detected in the mass spectra as ions corresponding to the protonated complex $[EST\text{-}2\beta CD{+}H]^+$ at *m/z* 2541.92 and the sodium adduct of complex $[EST\text{-}2\beta CD{+}Na]^+$ at *m/z* 2563.91. These observations suggest an interaction with 1:2 stoichiometry. No peaks corresponding to the association of one EST and one βCD were detected revealing the absence of 1:1 stable noncovalent complexes.

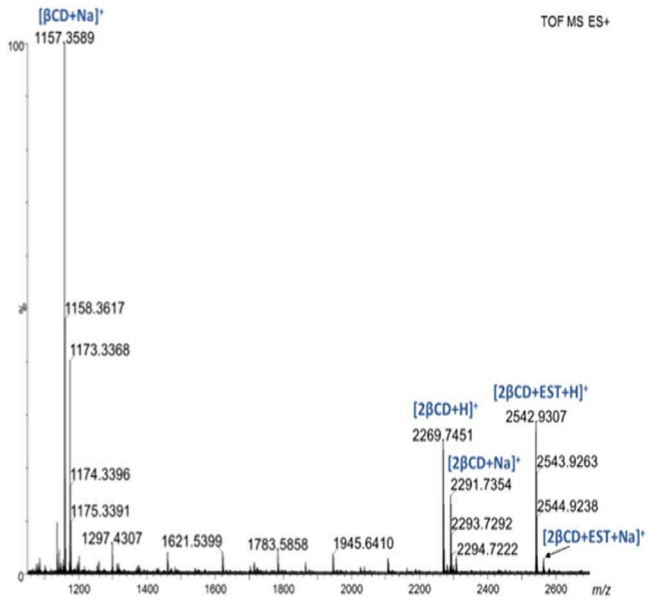


Fig. 2. HRMS spectrum of EST-βCD complex.

Table 3
The major peaks (*m/z* values and molecular formulas) from the HRMS measurement of the EST-βCD complex.

	Elemental composition	*m/z* $[M+H]^+$ or $[M{+}Na]^+$	
		Calculated	Found
$[\beta CD{+}Na]^+$	$C_{42}H_{71}O_{35}$	1157.3595	1157.3589
$[2\beta CD{+}H]^+$	$C_{84}H_{141}O_{71}$	2269.7474	2269.7451
$[2\beta CD{+}Na]^+$	$C_{84}H_{140}O_{70}Na_1$	2291.7293	2291.7354
$[2\beta CD\ {+}EST{+}H]^+$	$C_{102}H_{165}O_{72}$	2541.9250	2541.9248
$[2\beta CD\ {+}EST{+}Na]^+$	$C_{102}H_{164}O_{72}Na_1$	2563.9069	–

3.2. Phase solubility studies

The phase-solubility profile of EST in aqueous solution of successively increased β-CD concentrations at 25 °C (Fig. 3) indicates a B_S-type system which is usually observed with natural CDs, especially β-CD [34]. EST solubility increases linearly with increasing β-CD concentration in the range of 0.2 - 1.4 mM due to the formation of 1:1 EST:β-CD molar ratio inclusion complexes at the first stage. At the end of this linear portion, the maximum solubility S_{max} of EST is achieved. The solubility of the 1:1 EST:β-CD complex $S_{1:1}$ can be calculated as: $S_{1:1} = S_{max} - S_0$, where S_0 is the solubility of EST in pure water determined by the intercept of the phase solubility diagram.

Additional CD does not further increase the EST solubility and a first plateau is observed at β-CD concentration of 2 to 10 mM. This is due to the limited 1:1 complex solubility and/or the formation of 1:2 EST:β-CD molar ratio complexes. As the Gibbs phase rule indicates [35] only one discrete complex may precipitate at the plateau segment of the diagram. Thus, at even higher β-CD concentrations (12 – 20 mM), where a second plateau is observed, the solubility approximates that of the pure 1:2 complex $S_{1:2}$. The scheme of the probable complex formation is given below as eq. 4.

For the two-step association process of EST complexation with β-CD, the apparent stability constants, K_1, K_2 of the following equilibria:

$$EST + CD \overset{K_1}{\leftrightarrow} EST/CD \tag{3a}$$

$$EST/CD + CD \overset{K_2}{\leftrightarrow} EST/CD2 \tag{4a}$$

and the $K_{overall} = K_1 \cdot K_2$where estimated according to Liu et al. [35]

$$K_1 = \frac{slope}{S_0(1 - slope)} = 21,599\ M^{-1} \tag{3b}$$

$$K_2 = \frac{S_{1:2} \cdot slope}{{S_{1:1}}^2(1 - slope)} = 757\ M^{-1} \tag{4b}$$

where the slope was obtained from the linear part of the diagram and S_0, $S_{1:1}$, $S_{1:2}$ as described above. Thus, $K_{overall} = 1.6 \pm 0.4 \cdot 10^7\ M^{-2}$

This value is in the same order of magnitude to those estimated for progesterone, testosterone and cortisone by Liu et al. [36]. In that work the formation of inclusion complexes of steroids with β-CD at stoichiometric ratio of 1:2 was shown and its dependence on the steroid structure was discussed.

The *K* can be used to calculate Gibbs free energy (ΔG) according to the Eq. (5):

$$\Delta G = -RTlnK \tag{5}$$

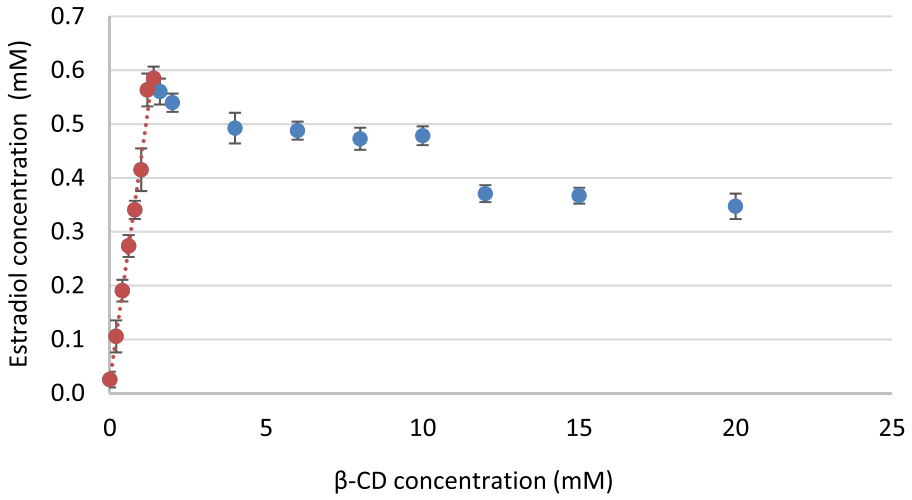


Fig. 3. Phase solubility diagram of EST/β-CD system in water at 25 °C ($n = 3$). The linear portion of the diagram (red line) was used for the calculation of K_1.

where R is gas constant (8.314 J mol^{-1} K^{-1}) and T is temperature (298 K).

This calculation reveals that the obtained here ΔG is equal to −9.92 kcal/mol.

3.3. QM calculations

As described above, the determination of this complex's stoichiometry is experimentally difficult and not straightforward task. Therefore, having in mind a huge applicability of the molecular modelling in the analysis of structure and properties of the CD complexes reviewed by us in the last year [16] we decided to apply these techniques also for this purpose. Therefore, the next step of our work was assessment of the QM approaches to check if they can properly foresee the host-guest molecular ratio and the complex association constant.

Knowledge about the crystal structure of a studied system significantly facilitates the calculations as it can be used to set the initial geometry of the complex. Thankfully, in our previous work we have determined the crystal structure of the EST-βCD complex [15]. In the current study it has been used as a starting point for all of the calculations.

Since we wanted to validate whether QM calculations can be used to predict the molar ratio and structure of the most stable complex, we have prepared 3 types of systems. The first one was 1:2 molar ratio complex and the experimental crystallographic structure of the hydrate of EST-βCD complex, after removing the water molecules, was used as an input for the computations. As EST is a molecule with quite limited conformational space, but also it is characterized by a structural anisotropy, to represent 1:1 molar ratio we needed two models called ''head up'' and ''head down'', as presented in Fig. 4. ''Head up'' is the case when EST's 5-carbon ring (steroidal D ring) goes through the wider CD's rim and the ''head down'' option is the case when it is the EST's 6-carbon ring (steroidal A ring) that protrudes through CD's wider rim. Those structures were based directly on the 1:2 experimental crystallographic data. We obtained them by removing each time different CD molecule from the 1:2 system.

3.4. DFT calculations

We have decided to apply Density Functional Theory (DFT) and semi-empirical methods. The chosen computational approaches were already presented in Table 1 and Table 2. In contrast to the majority of previously published cases of the CD complexes analysis by the means of DFT [16] where usually either 6–31 G or 6–31G(d) basis sets have been used, in this study for all DFT calculations we have used a relatively

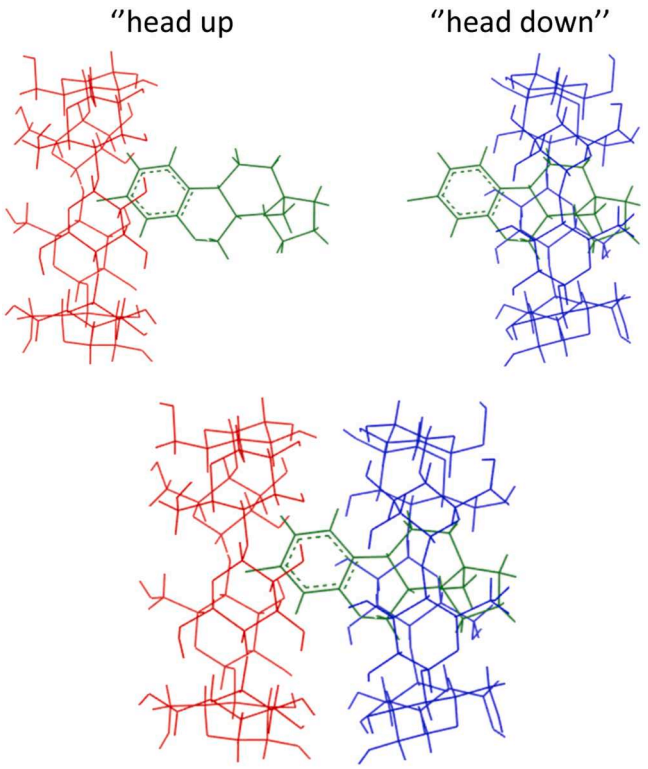


Fig. 4. Initial, non-optimized structures of 1:1 and 1:2 complexes, based on the experimental crystallographic data.

large 6–311G(d,p) basis set. Even if such approach has significantly elongated the computational time, it was a conscious choice. This way we have eliminated a risk of the influence of the basis set of the insufficient size on the obtained results.

Except of the choice of the functional in DFT (here: B3LYP and M062X) and type of the semi-empirical method (here: PM6 and PM7), one of the crucial decisions was about the type of the solvent representation. We have tested all 3 most common options: lack of solvent representation, PCM and SMD implicit solvent models.

PCM is the most often used solvent model in the computational analysis of CD complexes [37]. A different approach is presented by SMD. This model defines the free energy of solvation via two components: the one is electrostatic contribution arising from the

self-consistent reaction field, the other comes from the short-range interactions between the solute and solvent molecules [20].

Second choice of the calculations parameters was the implementation or lack of the dispersion correction. Noncovalent forces like London and van der Waals interactions are crucial for the formation and stability of the CD inclusion complexes. This aspect is not included in the calculations using exchange-correlation functionals, as the long-range electron correlation effect, known as the London part of the dispersion energy term, is not included in the Kohn–Sham DFT equation. For years this was a real issue influencing the accuracy of DFT calculations. Nowadays, different dispersion corrections are available. According to our literature research [16], in almost all studies concerning CD complexes only Grimme dispersion correction (here: D3) was used. Hence, in this study we also apply only this type of dispersion correction and compare the results to the calculations where no dispersion correction was implemented.

The results of the DFT calculations are presented in Table 4. It is clearly visible that each of three variables used in this study and described above: type of functional, solvation scheme, dispersion correction had an influence on the results.

From the energetic point of view, in 6 out of 13 cases clearly the preferred structure is the '1:1, either "head up'' (or "head down". As we know from the HRMS studies, this is not consistent with the reality. The approaches which properly predict the 1:2 complex stoichiometry are models numbered 3, 5–8, 11 and 13. On this example we can see that applying dispersion correction improve the accuracy of results when the system is treated as the in vacuo one (models 1 and 3). Simultaneously, PCM solvent model works well also when no dispersion correction has been used (Models 5 and 6) [38].

Among them there is one scheme, the scheme number 5 (B3LYP-PCM, without dispersion correction), that most significantly favours the stability of the 1:2 molar ratio complex over the 1:1 ones, both in terms of energy and free enthalpy of complexation.

3.5. Semi-empirical calculations

From the energetic point of view, the results of the semi-empirical calculations are diversified, Table 5. In 10 out 12 cases, the 1:2 molecular ratio complex has been defined as the most stable one, what stays in accordance with the experimental results. However, in those two cases where the 1:1 complex stoichiometry has been favoured, difference between the 1:1 most preferred complex and 1:2 complex is small and has the value of 0.59 kcal/mol and 3.15 kcal/mol for C and F schemes, respectively. In other words, in general, when taking into account the energetic aspect, the applied semi-empirical approaches properly predicted which molar ratio describes a complex of the highest stability.

11 out 12 models show that all three options: 1:1 head up, 1:1 head down and 1:2, are energetically stable. Only in model K the value of ΔE is positive for 1:1 head down complex. In some computational schemes differences between these three possible options are almost neglectable, for instance in model C. Whereas in other cases, the differences are much bigger, like in model L, where this difference reached up to around 37 kcal/mol.

On the contrary, from thermodynamic point of view, the majority of models pointed out the 1:1 complex as the most stable. The exceptions are models A, B, I and K. However, similarly to the ΔE results, in almost all cases, all three structural options have been defined as probable. Only in model K some of the ΔG values are positive. Model K is also the only computational scheme which shows a distant difference between both 1:1 stoichiometries and 1:2 stoichiometry and distinctively favours the 1:2 complex molar ratio. Hence, we can assume that the K computational model (PM6-SMD, without dispersion correction) predicted the experimental results in a most accurate way.

Except for this one K model, in rest of the cases the differences in values between 1:1 head up and 1:1 head down options within one computational method are similar. This suggests that there is a similar probability of creation of 1:1 head up and 1:1 head down complexes.

All those findings may be taken as a guide to create the hypothesis on the 1:2 molar ratio complex formation path. Most probably, the complexation happens in two steps. The first one, is creation of the 1:1 complex. The second step is association of the second βCD to the already existing 1:1 EST-βCD system. Which complex out of those two, 1:1 head up or 1:1 head down, is formed at the beginning can be deducted from the ΔG values. The complex characterized by a lower value should be created as the first one. According to both the most accurate DFT approach, 5 (B3LYP PCM), as well as to the MD MMGBSA results, the 1:1 head up complex is the more stable one. Therefore, it is highly probable that the complex formation occurs according to the scheme (Eq. (6)) presented below.

$$\boldsymbol{EST} + \boldsymbol{\beta CD} \rightleftharpoons \mathbf{1:1\ up} \xrightarrow{\beta CD} \mathbf{1:2} \tag{6}$$

With the first step being reversible and second irreversible reaction. Such predictions about the mechanism of the complex's formation are possible only thanks to application of the molecular modelling approach. The structural analysis of the results can be find in the Supporting Information.

Table 4
DFT calculations results. Yellow colour indicates the lowest value of ΔE within the given method (within the column). Blue colour indicates the lowest value of ΔG within the given method (within the column). ΔG and ΔE "1:2 from 1:1 up" – the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "up" is used as an initial structure for the second step; ΔG and ΔE "1:2 from 1:1 down" - the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "down" is used as an initial structure for the second step.

		B3LYP In vacuo	M062X In vacuo	B3LYP -D3 In vacuo	M062X -D3 In vacuo	B3LYP PCM	M062X PCM	B3LYP -D3 PCM	M062X -D3 PCM	B3LYP SMD	M062X SMD	B3LYP -D3 SMD	M062X -D3 SMD	ωB97X-D PCM
model	kcal/mol	1	2	3	4	5	6	7	8	9	10	11	12	13
1:1 up	ΔE	-6.47	-37.35	-54.67	-49.14	-6.93	-31.49	-30.41	-45.89	-5.06	-30.35	-48.12	-44.56	-36.80
	ΔG	8.00	-12.28	-31.73	-34.28	6.12	-8.18	-9.64	-19.64	7.44	-10.08	-24.81	-21.35	-17.91
1:1 down	ΔE	0.44	-27.64	-42.20	-47.06	1.87	-30.82	-30.88	--43.37	1.99	-30.88	-49.20	-44.17	-41.70
	ΔG	15.40	-12.10	-21.32	-23.36	16.93	-8.22	-10.9	-19.09	14.24	-11.01	-26.97	-19.61	-19.90
1:2	ΔE	2.12	-26.46	-64.81	-47.99	-61.42	-81.56	-41.75	-118.54	5.59	-20.69	-65.69	-36.78	-120.68
	ΔG	13.33	-9.18	-38.67	-26.28	-47.69	-28.08	-49.66	-34.58	15.18	-0.75	-44.76	-19.26	-65.87
1:2 from 1:1 up	ΔE	8.59	10.89	-10.14	1.15	-54.49	-50.07	-11.34	-72.65	10.65	9.66	-17.57	7.78	-83.88
	ΔG	5.33	3.1	-6.94	8	-53.81	-19.9	-40.02	-14.94	7.74	9.33	-19.95	2.09	-47.96
1:2 from 1:1 down	ΔE	1.68	1.18	-22.61	-0.93	-63.29	-50.74	-10.87	-161.91	3.6	10.19	-16.49	7.39	-78.98
	ΔG	-2.07	2.92	-17.35	-2.92	-64.62	-19.86	-38.76	-15.49	0.94	10.26	-17.79	0.35	-45.97

Table 5
Semi-empirical calculations results. Yellow colour indicates the lowest value of ΔE within the given method (within the column). Blue colour indicates the lowest value of ΔG within the given method (within the column). ΔG and ΔE "1:2 from 1:1 up" – the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "up" is used as an initial structure for the second step; ΔG and ΔE "1:2 from 1:1 down" - the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "down" is used as an initial structure for the second step.

		PM7 In vacuo	PM7 -D3 In vacuo	PM6 In vacuo	PM6 -D3 In vacuo	PM7 PCM	PM7 -D3 PCM	PM6 PCM	PM6 -D3 PCM	PM7 SMD	PM7 -D3 SMD	PM6 SMD	PM6 -D3 SMD
model	kcal/mol	A	B	C	D	E	F	G	H	I	J	K	L
1:1 up	ΔE	-59.50	-40.53	-11.68	-37.84	-54.53	-35.59	-7.45	-34.91	-51.21	-32.38	-5.08	-24.27
	ΔG	-37.13	-27.15	-38.64	-54.75	-30.07	-30.93	-10.23	-43.58	-34.98	-17.14	7.82	-24.52
1:1 down	ΔE	-54.76	-38.20	-11.05	-35.41	-43.27	-38.76	-8.96	-35.52	-47.85	-34.58	86.58	-27.98
	ΔG	-37.14	-43.81	-25.83	-31.87	-25.31	-42.85	-5.93	-53.45	-41.29	-18.23	9.60	-54.32
1:2	ΔE	-84.25	-61.17	-11.09	-61.45	-60.84	-35.61	-18.59	-54.94	-61.54	-40.77	-19.91	-64.91
	ΔG	-71.81	-48.48	5.34	-41.83	-43.04	-9.13	0.62	-34.81	-48.49	-23.39	-3.33	-39.68
1:2 from 1:1 up	ΔE	-24.75	-20.64	0.59	-23.61	-6.31	-0.02	-11.14	-20.03	-10.33	-8.39	-14.83	-40.64
	ΔG	-34.68	-21.33	43.98	12.92	-12.97	21.8	10.85	8.77	-13.51	-6.25	-11.15	-15.16
1:2 from 1:1 down	ΔE	-29.49	-22.97	-0.04	-26.04	-17.57	3.15	-9.63	-19.42	-13.69	-6.19	-106.49	-36.93
	ΔG	-34.67	-4.67	31.17	-9.96	-17.73	33.72	6.55	18.64	-7.2	-5.16	-12.93	14.64

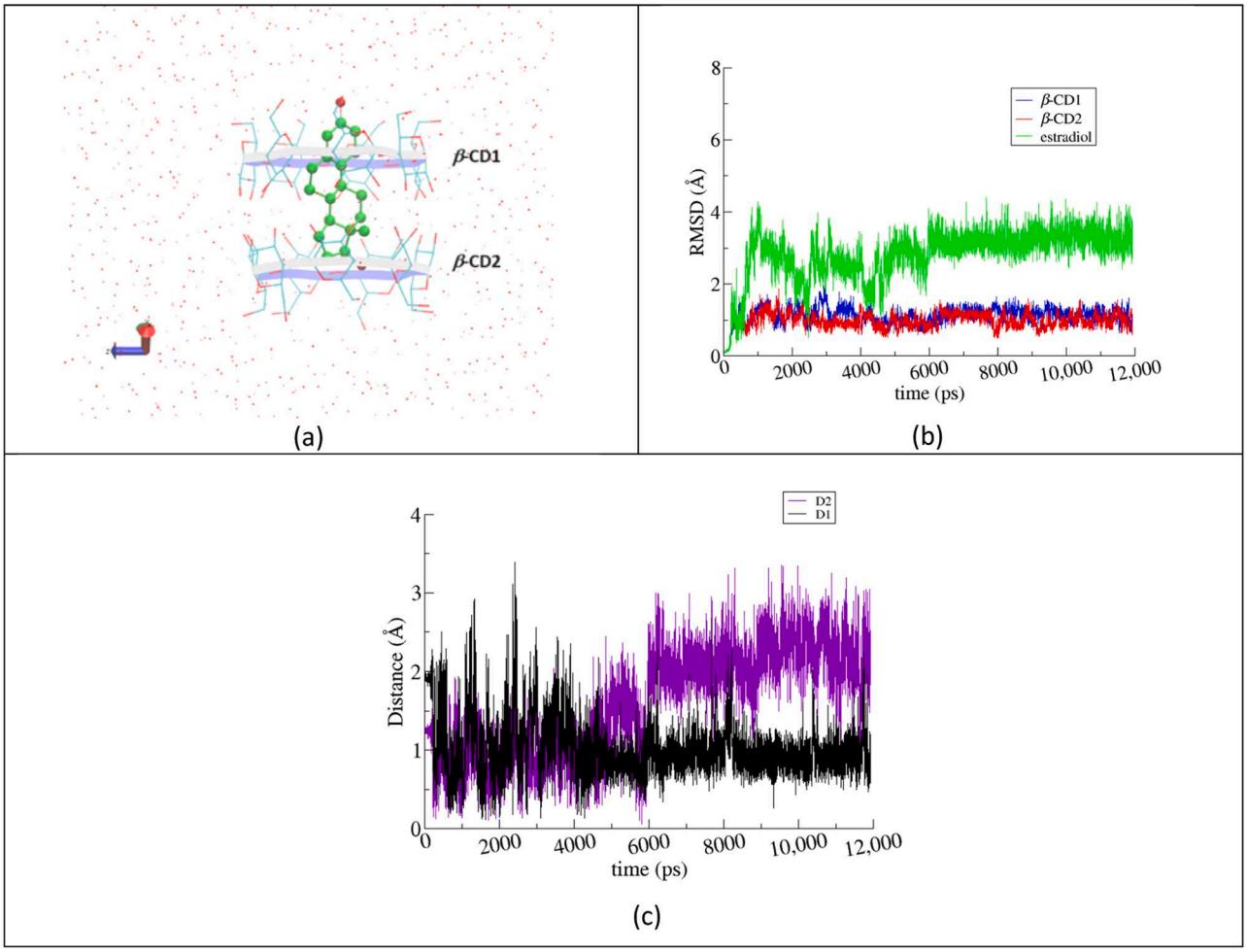


Fig. 5. (a) The starting model of the 1:2 inclusion complex based on the crystallographically determined atomic coordinates (CSD Refcode: OFANUI). (b) RMSD evolution of the host and guest molecules of the complex between the simulated states and the first frame of the simulations. (c) Distance D1 between the center of mass (COM) of the steroidal A-ring of EST and the O4*n* atom mean plane of the host β-CD1; distance D2 between COM of the steroidal d-ring of EST and the O4*n* atom mean plane of the other host of the dimeric cavity, β-CD2.

3.6. Computational thermodynamic results vs experimental data

As described in the previous section, obtention of the experimental complex stability constant allowed to define ΔG, which in case of EST-βCD complex is equal to −9.92 kcal/mol. Among the tested QM approaches, the ones which favour the most the 1:2 molar ratio are those in which the PCM correction has been applied However, the calculated ΔG overestimate the experimental ones Interestingly, the values close to the experimental one (−9.92 kcal/mol) have been obtained using M062X in vacuo (−9.18 kcal/mol) and PM7-D3 PCM approaches (−9.13 kcal/mol). Unfortunately, at the same time, both of them suggest that the 1:1 is more stable ratio than 1:2. Here, it should also be noted that β-CD is surrounded by hydration waters, which could be included in the equation of the complex formation. It was shown previously that an agreement between theoretical and experimental entropy data for inclusion complex formation was only attained when explicit water molecules were included [39].

3.7. MD simulations

MD simulations in aqueous media were carried out for the 3 types of systems, i.e. the crystallographically determined 1:2 guest:host complex and the two 1:1 monomeric complexes of opposite EST accommodation in the β-CD cavity (noted as "head up" and "head down" in Fig. 7). By monitoring the frames during the time interval of the simulations the following observations were made:

In the case of the 1:2 complex, where the starting model was retrieved from the crystal structure (Fig. 5a), the β-CD dimer encapsulating an EST molecule is preserved in the time frame of the simulation. In the absence of crystal contacts and in the presence of the surrounding water molecules, the guest EST rotates around and moves along the 7-fold molecular β-CD axis. However, it was observed a clear tendency of the guest's steroidal A-ring to be accommodated near the narrow rim of β-CD and its d-ring closer to the dimeric interface region. The measured distance between the center of mass (COM) of the A-ring and the O4n atoms mean plane of the host β-CD1, in whose cavity the A-ring is located, fluctuates around 1 Å, whereas that between the d-ring COM and the O4n plane of the other host (β-CD2), fluctuates in the range of 1 to 3 Å (Fig. 5b and 5c).

For the 1:1 molar ratio complex, both "head up" and "head down" models were examined. In the case of the "head down" binding mode, the A-ring of the guest cannot be stabilized in the host's wide rim with the rest part of EST protruding from the narrow rim of the host. Thus, EST is swiftly displaced from its initial location, exposing its A-ring to the solvent and accommodating the d-ring in the cavity (Fig. 6a). This behavior is also reflected in the high EST mobility displayed in the respective RMSD plot (Fig 6b).

On the other hand, in the case of the "head up" binding mode, the A-ring of EST which is initially exposed to the solvent by protruding from the narrow β-CD rim, it is accommodated quickly in the narrow β-CD rim where it remains relatively stable in the time frame of the simulation (Fig. 7a). The respective RMSD plot for the molecules of the system

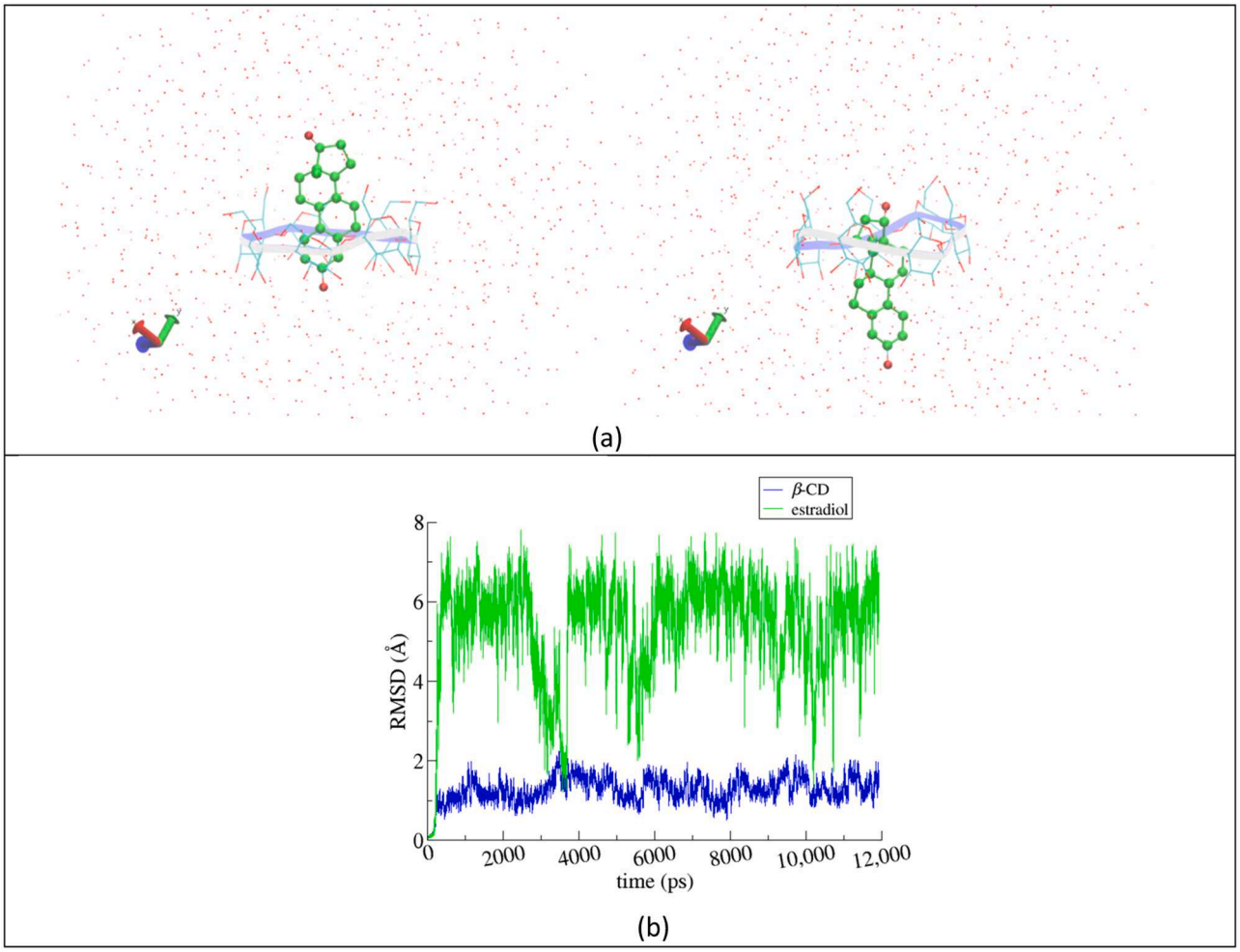


Fig. 6. (a) Two representative snapshots, at 0 and the 10th ns of the "head down" system simulation.
(b) RMSD plot for the host and guest molecule of the complex. The guest cannot be tightly stabilized in this orientation, thus exhibiting high mobility.

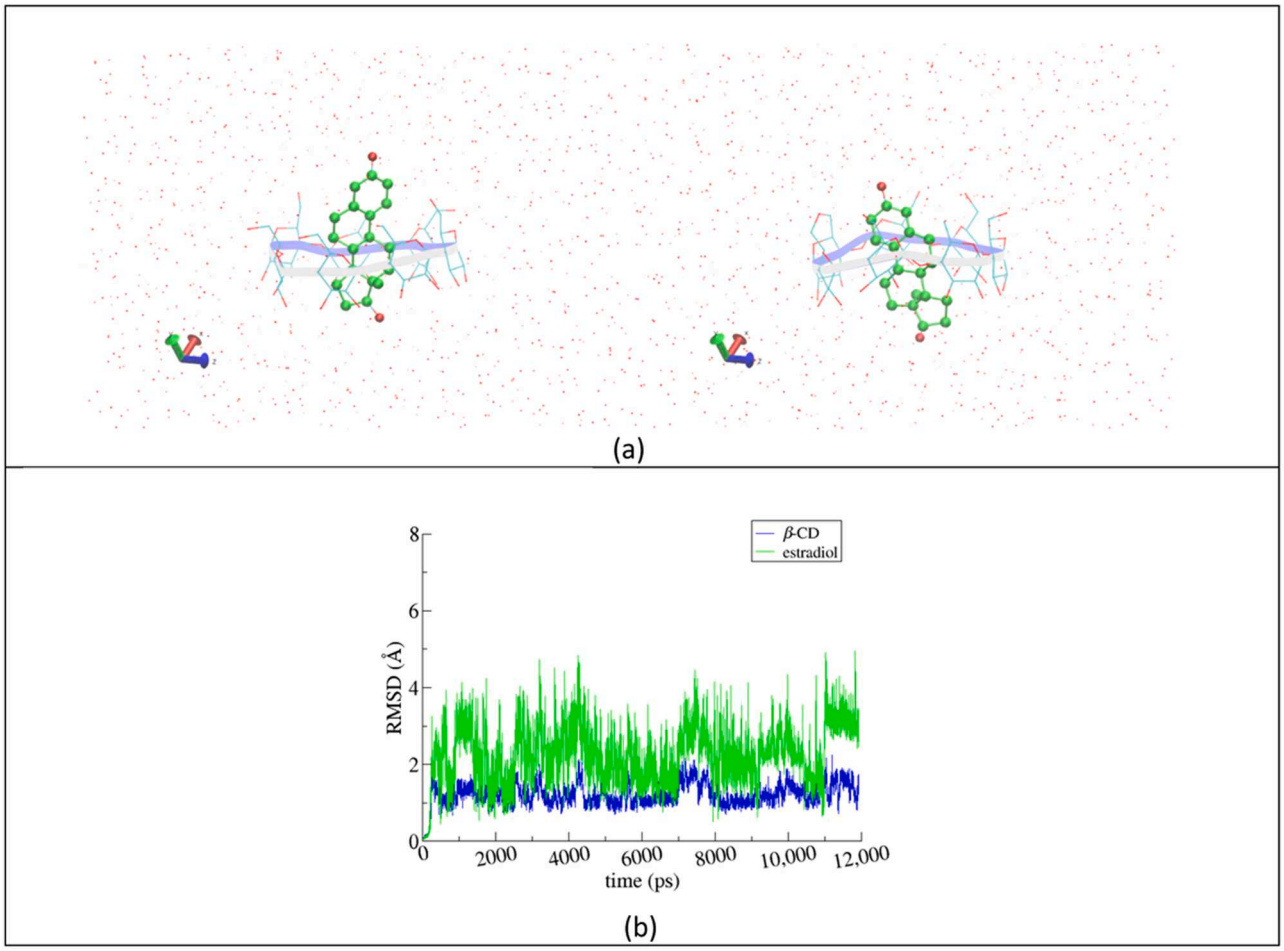


Fig. 7. (a) Two representative snapshots, at 0 and the 10th ns of the "head up" system simulation.
(b) RMSD plot the for host and guest molecule of the complex. The mobility of the guest is clearly lower than that of EST in the "head down" system simulation.

clearly shows a lower mobility of the guest compared to that of the "head down" system.

From all the above, it is concluded that the formation of a 1:1 inclusion complex of the "head up" binding mode is favored over the "head down" binding mode. Moreover, the dynamic behavior of the examined 1:2 complex, showed a very stable complex that tends to retain the accommodation of the guest (with its A-ring near the narrow rim and the d-ring near the interface of the hosts' dimer) according to that of the "head up" binding mode. These findings support the proposed complex formation described in the scheme of eq. 4.

Molecular Mechanics/Generalized Bohrn surface area (MM/GBSA) calculations [40], performed for the 3 examined types of complexes, further verify the above conclusions. The estimated host–guest binding affinities, as listed in Table 6, were extracted from 10,000 snapshots over the last 10-ns of the MD simulations. As expected, the lower ΔGbind value is estimated for the 1:2 complex mainly due to the extended van der Waals interactions between the guest and the dimeric host, that significantly decrease the averaged change of van der Waals energies (ΔEvdW) upon EST inclusion in the host dimeric cavity. By comparing the ΔGbind values estimated for the systems of common 1:1 guest:host stoichiometry ("head down" and "head up"), the considerably lower ΔGbind value of the "head up" system indicates a more stable inclusion complex.

Table 6
Binding free energies and their standard deviations (kcal/mole) resulting from MM/GBSA analysis of the inclusion compounds of EST in β-CD with guest: host ratios of 1:2, 1:1 ("head down" mode) and 1:1 ("head up" mode), respectively.

	EST/β-CD (1:2)	EST/β-CD (1:1 "head down")	EST/β-CD (1:1 "head up")
ΔE_{vdW}	−47.31 ± 2.41	−29.10±2.31	−30.46 ± 1.68
ΔE_{ele}	−4.15 ± 2.78	−2.12±2.13	−2.09 ± 1.72
ΔE_{MM} [a]	−51.46 ± 3.58	−31.23±3.21	−32.55 ± 2.38
ΔG_{GB}	23.22 ± 2.98	15.63±2.80	13.45 ± 1.99
$\Delta G_{nonpolar}$	−4.45 ± 0.22	−2.94±0.16	−3.03 ± 0.12
$\Delta G_{solvation}$ [b]	18.77 ± 2.94	12.69±2.72	10.42 ± 1.94
ΔH [c]	−32.69 ± 4.63	−18.54±2.06	−22.13 ± 1.97
$T{\cdot}\Delta S$ [d]	−19.03 ± 3.00	−16.26±1.21	−15.88 ± 1.36
ΔG_{bind} [e]	−13.66 ± 4.20	−2.27±2.39	−6.25 ± 2.40

ΔE_{vdW} = van der Waals contribution from molecular mechanics; ΔE_{ele} = electrostatic energy as calculated by the molecular mechanics force field; ΔG_{GB} = the electrostatic solvation energy (polar contribution) calculated using the GB model; $\Delta G_{nonpolar}$ = nonpolar contribution to the solvation free energy, calculated by the solvent-accessible surface area (SASA) method;.

[a] $\Delta E_{MM} = \Delta E_{vdW} + \Delta E_{ele}$;.
[b] $\Delta G_{solvation} = \Delta G_{GB} + \Delta G_{nonpolar}$;.
[c] $\Delta H = \Delta G_{solvation} + \Delta E_{MM}$;.
[d] $T{\cdot}\Delta S$ entropic term calculated by normal mode analysis;.
[e] $\Delta G_{binding} = \Delta H - T{\cdot}\Delta S$.

3.8. Computational stability results vs experimental data

As described in the previous section, acquisition of the experimental complex stability constant allowed to define ΔG, which in case of EST-βCD complex is equal to −9.92 kcal/mol. This value is within the limits of uncertainty of ΔGbind for 1:2 complex, −13.66 ± 4.20 kcal/mol (Table 6). However, the values of ΔGbind obtained for 1:1 complex,

either head-up or head-down orientations, −2.27±2.39 and −6.25 ± 2.40 respectively, although not strictly within the uncertainty limits, are also close to the experimentally determined one. Therefore, while the MD/MMGBSA method allows to properly indicate the order of the magnitude of ΔGbind, it should be supported by the experimental analysis such as HRMS to confirm the complex ratio.

4. Conclusions

Even though the existence of the EST-βCD complex in water solution has been known for decades, and despite quite some studies performed to define its molar ratio, this complex's structure remained not properly determined until now. The hypothesis that the knowledge in this context might be not complete, has occurred after determination of the EST-βCD crystal structure where the molar ratio was found to be 1:2 (EST:βCD).

In this work, thanks to application of the HRMS approach, it has been indisputably proven that the EST-βCD complex molar ratio is 1:2 and not as previously assumed 1:1. Moreover, the phase solubility studies confirmed these results. This type of experiments has been performed before, however, never the 1:2 molar ratio has been taken into account as a possible description of this system. In other words, the indisputable HRMS measurement results prompted the revision of the phase solubility studies. This allowed to properly define complex stability constant K, what in turn delivered information about the Gibbs free energy value of the complex formation.

The structure and thermodynamics of the complex were further analyzed using various QM (DFT, semi-empirical) and MM (MD/MMGBSA) approaches. Tests on the application of different computation parameters such as presence/absence of dispersion correction, choice of implicit solvent model or DFT functional, have been performed.

Possession of the credible experimental data allowed to assess the computational approaches. While some of the "static" QM methods properly indicated the correct host: guest ratio at the same time they failed to accurately predict the Gibbs free energy of complexation. On the other hand, QM methods that properly described the value of ΔG of 1:2 complex formation, such as M062X in vacuo, favored the 1:1 stoichiometry, which was experimentally excluded. The MD/MMGBSA method, although performed at the lower level of theory, accurately predicted the stability constant of the complexed but was not conclusive to indicate the formation of either 1:1 or 1:2 complex.

This leads to the conclusion that among tested computational approaches, there are some which are able to properly predict the composition of such complex and some that can assess the stability of the studied system. However, there is not single method that would allow to reproduce both the stoichiometry and the thermodynamic stability of the complex at the same time.This study finally describes the structure and thermodynamics of the EST-βCD complex in aqueous solution and delivers experiment-based information about the formation of this complex.

CRediT authorship contribution statement

Anna Helena Mazurek: Writing – original draft, Software, Investigation, Funding acquisition, Data curation, Conceptualization. **Łukasz Szeleszczuk:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Conceptualization. **Kostas Bethanis:** Writing – original draft, Methodology, Investigation, Conceptualization. **Elias Christoforides:** Software, Resources, Project administration, Methodology. **Marta Katarzyna Dudek:** Software, Methodology, Conceptualization. **Ewelina Wielgus:** Writing – review & editing, Validation, Project administration. **Dariusz Maciej Pisklak:** Writing – review & editing, Writing – original draft, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the Medical University of Warsaw, Poland [grant number WF7/1/F/MB/N/23].

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.molstruc.2024.138710.

References

[1] J. Szejtli, Introduction and general overview of cyclodextrin chemistry, Chem. Rev. 98 (1998) 1743–1754.
[2] B.G. Poulson, Q.A. Alsulami, A. Sharfalddin, E.F. El Agammy, F. Mouffouk, A.-H. Emwas, L. Jaremko, M. Jaremko, Cyclodextrins: structural, chemical, and physical properties, and applications, Polysaccharides 3 (2022) 1.
[3] G. Crini, Review: a history of cyclodextrins, Chem. Rev. 114 (2014) 10940–10975.
[4] T. Irie, K. Uekama, Pharmaceutical applications of cyclodextrins. III. Toxicological issues and safety evaluation, J. Pharm. Sci. 86 (1997) 147–162.
[5] V. Aiassa, C. Garnero, M.R. Longhi, A. Zoppi, Cyclodextrin multicomponent complexes: pharmaceutical applications, Pharmaceutics 13 (2021) 1099.
[6] S.S. Jambhekar, P. Breen, Cyclodextrins in pharmaceutical formulations I: structure and physicochemical properties, formation of complexes, and types of complex, Drug Discov. Today. 21 (2016) 356–362.
[7] S.S. Jambhekar, P. Breen, Cyclodextrins in pharmaceutical formulations II: solubilization, binding constant, and complexation efficiency, Drug Discov. Today. 21 (2016) 363–368.
[8] M.P. Thomas, B.V. Potter, The structural biology of oestrogen metabolism, J. Steroid Biochem. Mol. Biol. 137 (2013) 27–49.
[9] Salole E.G. Estradiol, Analytical profiles of drug substances, 1986, 15, 283–318.
[10] N. Sadlej-Sosnowska, Fluorometric determination of association constants of three estrogens with cyclodextrins, J Fluoresc 7 (1997) 195–200.
[11] C. Yañez, J. Basualdo, P.J. Jara-Ulloa, A. Squella, Inclusion complexes of estrone and estradiol with β-cyclodextrin: voltammetric and HPLC studies, J. Phys. Org. Chem. 20 (2007) 499–505.
[12] R.L. Pérez, G.M. Escandar, Spectrofluorimetric study of estrogen-cyclodextrin inclusion complexes in aqueous systems, Analyst 138 (2013) 1239–1248.
[13] D.H. Schwarz, A. Engelke, G. Wenz, Solubilizing steroidal drugs by β-cyclodextrin derivatives, Int J Pharm 531 (2017) 559–567.
[14] Z.Y. Lin, X.X. Wang, S.B. Kou, J.H. Shi, Exploring the inclusion interaction of estradiol with β-CD and HP-β-CD with the help of molecular dynamics simulation as well as multi-spectroscopic approaches, Spectrochim Acta A Mol Biomol Spectrosc 269 (2022) 120764.
[15] A.H. Mazurek, Ł. Szeleszczuk, K. Bethanis, E. Christoforides, M.K. Dudek, M. Zielińska-Pisklak, D.M. Pisklak, 17-β-Estradiol—B-Cyclodextrin complex as solid: synthesis, structural and physicochemical characterization, Molecules 28 (2023) 3747.
[16] A.H. Mazurek, Ł. Szeleszczuk, Current status of quantum chemical studies of cyclodextrin host–guest complexes, Molecules 27 (2022) 3874.
[17] T. Higuchi, K.A. Connors, Phase solubility techniques, Adv. Anal. Chem. Instrument. 4 (1965) 117–212. Open Access Library Available online: http://www.oalib.com/references/7163685 (accessed on 26 January 2024).
[18] Software https://gaussian.com/products/ accessed on 24th January 2024.
[19] M. Cossi, N. Rega, G. Scalmani, V. Barone, Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model, J. Comput. Chem. 24 (2003) 669–681.
[20] A.V. Marenich, C.J. Cramer, D.G Truhlar, Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions, J. Phys. Chem. B 113 (2009) 6378–6396.
[21] J.-D. Chai, M. Head-Gordon, Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections, Phys. Chem. Chem. Phys. 10 (2008) 6615.
[22] J. Eberhardt, D. Santos-Martins, A.F. Tillack, S. Forli, AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings, J. Chem. Inf. Model. 61 (2021) 3891–3898.
[23] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B.A. Shoemaker, P. A. Thiessen, B. Yu, et al., PubChem 2023 update, Nucleic Acids Res. 51 (2023) D1373–D1380.
[24] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the amber biomolecular simulation package, Wiley Interdiscip. Rev.: Comput. Mol. Sci. 3 (2012) 198–210.

[25] J. Wang, W. Wang, P. Kollman, D. Case, ANTECHAMBER: an accessory software package for molecular mechanical calculations, J. Chem. Inf. Comput. Sci. - JCISD (2000) 222.
[26] K.N. Kirschner, A.B. Yongye, S.M. Tschampel, J. Gonzalez-Outeirino, C.R. Daniels, B.L. Foley, R.J. Woods, GLYCAM06: a generalizable biomolecular force field. Carbohydrates, J. Comput. Chem. 29 (2008) 622–655.
[27] P. Mark, L. Nilsson, Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K, J. Phys. Chem. A 105 (2001) 9954–9960.
[28] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[29] V. Kräutler, W.F. van Gunsteren, P.H. Hünenberger, A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations, J. Comput. Chem. 22 (2001) 501–508.
[30] D.R. Roe, T.E. Cheatham 3rd, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theory Comput. 9 (2013) 3084–3095.
[31] B.R. Miller 3rd, T.D.J. McGee, J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBSA.Py: an efficient program for end-state free energy calculations, J. Chem. Theory Comput. 8 (2012) 3314–3321.
[32] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (33–38) (1996) 27–28.
[33] G. Zengin, Yildiztugay E. Nilofar, A. Bouyahya, H. Cavusoglu, R. Gevrenova, D. Zheleva-Dimitrova, A comparative study on UHPLC-HRMS profiles and biological activities of inula sarana different extracts and its beta-cyclodextrin complex: effective insights for novel applications, Antioxidants (Basel) 12 (2023) 1842.
[34] T.K. Špehar, M. Pocrnić, D. Klarić, B. Bertoša, A. Čikoš, M. Jug, J. Padovan, S. Dragojević, N. Galić, Investigation of praziquantel/cyclodextrin inclusion complexation by NMR and LC-HRMS/MS: mechanism, solubility, chemical stability, and degradation products, Mol. Pharm 18 (2021) 4210–4223.
[35] M.E. Brewster, T. Loftsson, Cyclodextrins as pharmaceutical solubilizers, Adv. Drug Deliv. Rev. 59 (2007) 645–666.
[36] F.Y. Liu, D.O. Kildsig, A.K. Mitra, Beta-cyclodextrin/steroid complexation: effect of steroid structure on association equilibria, Pharm. Res 7 (1990) 869–873.
[37] M. Cossi, N. Rega, G. Scalmani, V. Barone, Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model, J. Comput. Chem. 24 (2003) 669–681.
[38] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132 (2010) 154104.
[39] J.F. Lopes, Nascimento C.S. Jr, C.P.A. Anconi, H.F.D. Santos, W.B Almeida, Inclusion complex thermodynamics: the β-cyclodextrin and sertraline complex example, J. Mol. Graph. Model 62 (2015) 11–17.
[40] S. Genheden, U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, Expert Opin. Drug Discov. 10 (2015) 449–461.

International Journal of
Molecular Sciences

MDPI

Review

Application of Various Molecular Modelling Methods in the Study of Estrogens and Xenoestrogens

Anna Helena Mazurek [1], Łukasz Szeleszczuk [1,*], Thomas Simonson [2] and Dariusz Maciej Pisklak [1]

[1] Chair and Department of Physical Pharmacy and Bioanalysis, Department of Physical Chemistry, Medical Faculty of Pharmacy, University of Warsaw, Banacha 1 str., 02-093 Warsaw Poland; annamazurek21@gmail.com (A.H.M.); dpisklak@wum.edu.pl (D.M.P.)
[2] Laboratoire de Biochimie (CNRS UMR7654), Ecole Polytechnique, 91-120 Palaiseau, France; thomas.simonson@polytechnique.edu
* Correspondence: lszeleszczuk@wum.edu.pl; Tel.: +48-501-255-121

Received: 21 July 2020; Accepted: 1 September 2020; Published: 3 September 2020

Abstract: In this review, applications of various molecular modelling methods in the study of estrogens and xenoestrogens are summarized. Selected biomolecules that are the most commonly chosen as molecular modelling objects in this field are presented. In most of the reviewed works, ligand docking using solely force field methods was performed, employing various molecular targets involved in metabolism and action of estrogens. Other molecular modelling methods such as molecular dynamics and combined quantum mechanics with molecular mechanics have also been successfully used to predict the properties of estrogens and xenoestrogens. Among published works, a great number also focused on the application of different types of quantitative structure–activity relationship (QSAR) analyses to examine estrogen's structures and activities. Although the interactions between estrogens and xenoestrogens with various proteins are the most commonly studied, other aspects such as penetration of estrogens through lipid bilayers or their ability to adsorb on different materials are also explored using theoretical calculations. Apart from molecular mechanics and statistical methods, quantum mechanics calculations are also employed in the studies of estrogens and xenoestrogens. Their applications include computation of spectroscopic properties, both vibrational and Nuclear Magnetic Resonance (NMR), and also in quantum molecular dynamics simulations and crystal structure prediction. The main aim of this review is to present the great potential and versatility of various molecular modelling methods in the studies on estrogens and xenoestrogens.

Keywords: molecular modelling; estrogens; xenoestrogens; estradiol; docking; Density Functional Theory (DFT)

1. Introduction

Successful applications of molecular modelling methods can be found in almost every branch of modern physics, chemistry, and biology. This versatility and popularity results from the constantly increasing computing power of both personal computers and specialized servers as well as the availability of molecular modelling software. The number of properties that can be accurately predicted and phenomena that can be explained as well as problems that can be solved using such calculations are enormous. Therefore, in this review, recent advances in molecular modelling applications to study the chemistry and biochemistry of estrogens [1] and xenoestrogens [2] will be presented. The aim of this article is not only to present the large volume of information concerning title compounds that have been obtained in recent years using in silico methods but also to present to

researchers who are not specialized in molecular modelling methods the possible applications of these compounds and, in that way, encourage them to use such calculations in their own studies.

This review is organized as follows: first the title compounds, estrogens, and xenoestrogens are briefly summarized, with particular focus on those that have already been objects of computational studies. Then, the most important biomolecules that are involved in the metabolism and action of estrogens and xenoestrogens are described. In the main part of this review, the computational methods that have been employed in the studies on estrogens and xenoestrogens are presented. Each of the methods is briefly described, without too many details, as there are many very good reviews referenced in this article focusing on the basics of particular methods. Wherever possible, the computational results were compared to the corresponding experimental ones; however, in many cases, such comparison was impossible either due to the lack of experimental results or the purely theoretical character of the published work. In the next section, a critical analysis of the reviewed methods is presented, supported by the presentation of some technical aspects of the reviewed studies. This was done to facilitate the choice of a certain method or its properties. From the authors' perspective, the number of studies on estrogens and xenoestrogens is constantly increasing; however, only in some of them are the experiments supported by suitable theoretical studies. Therefore, our aim was to present to researchers working with estrogens and xenoestrogens selected computational tools that can be used to facilitate and improve their studies.

1.1. Estrogens and Xenoestrogens: Types, Main Representatives, and Their Toxicity

Estrogens are a group of natural steroid sex hormones. There are four of them: estrone (E1), estradiol (E2), estriol (E3), and estretrol (E4) [1] (see the Supplementary Materials: Figure S1). The last one is produced only during pregnancy by the fetus liver [3]. Among these four compounds, E2 plays the most important role in the human organism and, therefore, is of high importance in breast or ovarian cancer progression. With regard to their relative binding affinity (RBA) to Estrogen Receptor (ER), with the exclusion of estretrol, estrogens are ranked in the following order: estradiol > estrone > estriol. In comparison to estradiol, the activity and potency of estrone and estriol are, respectively, 10 and 100 times lower than that of E2 [4]. E1 mainly functions as estradiol's metabolite and, at the same time, serves as its precursor (the estrone-to-estradiol transformation is reversible) [5]. On the other hand, in non-pregnant women, estriol levels in the blood are hardly detectable, whereas during pregnancy its amount distinctly grows because it is produced by the placenta as well [6]. All estrogens are used as medication in menopausal hormone therapy, although estradiol is the most applied [7,8]. In such external applications, they should be considered as xenoestrogens.

Apart from estrogens, other non-endogenous substances, called endocrine-disrupting chemicals (EDCs), can bind to the ER. According to the European Commission Regulation from 2018 [9], EDCs are substances that have adverse effects in non-target organisms, have an endocrine mode of action, and exhibit these adverse effects as a consequence of an endocrine mode of action. In the Guidance on the identification and studies regarding ECDs, published by the European Chemical Agency (ECA) and the European Food Safety Authority (EFSA) [10], the importance of in silico studies in the process of EDC research is clearly pointed out.

More detailed information on in silico examination regarding EDCs is provided in OECD (Organization for Economic Co-operation and Development) Guidance (Update v3, 2017) [11], where Quantitative Structure–Activity Relationship (QSAR) methodologies, prediction of the metabolic transformation (ADME), and CYP450 metabolism investigations are listed as Level 1 methods in the identification and study process. It is emphasized that applications of such methods can be used to identify the groups of chemicals and structural characteristics that are responsible for the observed in vivo effects and can serve as a tool to explain possible differences between in vitro and in vivo results, clarifying EDCs' mechanism of action. In the Conceptual Framework for Testing and Assessment of EDCs [12], at investigation Level 1, not only QSAR and ADME but also 'other in silico model predictions' have been listed. Additively, special QSAR Guidance on the topic has been published.

A large group of EDCs that are well-known and have been investigated for decades are xenoestrogens, which are xenobiotics that cause either an estrogenic or an antiestrogenic effect [13]. Nowadays, they are widespread and originate from different sources (see the Supplementary Materials: Figure S2). In the in silico studies, the main emphasis has been put, so far, on the investigation of pharmaceuticals, phytoestrogens, bisphenol A, and phthalates.

1.2. Estrogen-Related Biomolecules as the Molecular Modelling Study Objects

Numerous studies involving various in silico methods, but mainly ligand docking, have been performed to look for new inhibitors of the enzymes metabolizing estradiol. These proteins are a possible target for new drugs and are depicted in Table 1. Most of the research deals with aromatase [14–22], as its inhibitors, like letrozole, are already in medicinal use [23]. However, those inhibitors exhibit some detrimental side effects, which is the reason why research continues on this topic. 17β-Hydroxysteroid dehydrogenase (17β-HSD) [24,25], sulfatase (STS) [26], as well as sulfotransferase (SULT) [27] have also been taken under close consideration using molecular modelling in order to contribute to anti-cancer drug development. The first of the two enzymes transforms estrone into estradiol. The second inactivates estradiol by transforming it into a sulfated form.

Table 1. Reagents of the main metabolic processes regarding estradiol that have been investigated with in silico methods.

Substrate	Enzyme	Product	Ref.
Testosterone	Aromatase (CYP219A1)	Estradiol	[16]
Estradiol	17β-OH-dehydrogenase (17β-HSD)	Estrone	[24,25]
Estradiol	CYP1B1	4-OH-hydroxylated estradiol	[30]
Estradiol	CYP1A1, CYP1A2	2-OH-hydroxylated estradiol	[28,29]
Estradiol	Sulfotransferase (SULT)	Inactivated estradiol (sulfated)	[27]
Inactivated (sulfated) estradiol	Sulfatase (STS)	Activated estradiol	[26]

Even if conceptually estradiol is central to the above-mentioned studies, none of them used it as a target ligand. Nevertheless, in order to model estradiol activity metabolism in molecular studies, the most common objects are estradiol itself, its natural receptor—the estrogen receptor (ER)—and its main hydroxylating enzymes: CYP1A1 [28], CYP1A2 [29], CYP1B1 [30], and SHBG (sex hormone-binding globulin). The last one is a protein that can bond estradiol (E2); thus, it has direct impact on the amount of free E2 in plasma and, consequently, on the hormone's bioavailability [31]. Taking into account the interest of in silico research for the enzymes listed in this paragraph, only topics associated with these systems are covered in the main part of this review.

Two subtypes of the estrogen receptor are known, ERα and ERβ, with tissue-specific expression [32]. ERα is found mainly in the mammary gland, uterus, ovary (thecal cells), male reproductive organs, prostate, liver, and adipose tissue. ERβ is present in the prostate, bladder, ovary (granulosa cells), colon, adipose tissue, and immune system [33]. Despite being encoded by different genes, both estrogen receptors show high homology, and in both of them the E domain contains the ligand-binding domain (LBD) and C domain, which is the DNA binding domain (DBD) [34]. The homology between ERα and ERβ in DBD is more than 95% [35]. In LBD the homology is about 55% [35]. In ER structures, two transactivation functions are present, called AF-1 (located in N-terminal domain) and AF-2 (located in LBD) [36]. They contain the nuclear location signals and, after proper exposure of their surface, are

responsible for incorporation of the co-activators, which is a necessary step to induce activation of the intercellular signaling pathways.

Binding of 17β-estradiol and any agonist to the ER requires creating a hydrogen bond with His524 (in ERα) [37] or His475 (in ERβ) [38]. This leads to a unique agonist-bound conformation of the receptor's LBD, characterized by a specific repositioning of the H12 helix, which is the most C-terminal helix of the LBD (molecular switch) [39]. On the contrary, selective ER modulators (SERMs) such as raloxifene or tamoxifen induce relocation of H12 into the co-activator binding cleft, which blocks AF-2 activity. Finally, pure antagonists completely destabilize H12 [40].

All the above actions concerning estrogens' binding to ER are genomic actions. This means that they require translocation of the estrogen–ER complex to the nucleus and interaction with chromatin at specific sequences, known as estrogen response elements [41]. There are other estrogen signaling paths that are non-genomic and involve indirect regulation of gene expression [42]. They include activation of various protein-kinase cascades after binding of an estrogen molecule to a membrane receptor, usually GPER1 (G protein-coupled ER1) [43]. It has been proven that estrogen binding to GPER1 shows similarity to estrogen–ER binding; however, estrogen's affinity for GPER1 is significantly lower than for ER [43]. Nevertheless, all other steroid hormones are characterized by even lower affinities towards GPER1.

Experimental analysis of the GPER1 structure has been limited, as the protein is rather refractory to both X-ray crystallography and NMR due to its relatively high lipophilicity [44]. This explains the importance of computational homology modelling [45]. Homology models are later used to simulate binding with estrogens. There are only a few studies dealing directly with this subject [44,46]. Application of homology modelling, molecular docking, and molecular dynamics provides insight into the process of estrogen binding and helps to explain the induced non-genomic effects. As non-genomic estrogen actions through GPER1 and kinase cascades alter the cell membrane shape, further research on cell membranes and estrogens is also performed (see Section 2.1.3).

2. Application of Molecular Modelling Methods in the Study of Estrogens and Xenoestrogens

Molecular modelling could be of a great help to experimentalists. Ligand docking directs research to the most probable hit molecules; QSAR, an officially accepted OECD method, predicts toxicity and often can facilitate research; and MD allows one to observe the time-dependent mechanisms (e.g., in membranes) and, therefore, helps to explain the experimentally observed phenomena. What is more, two types of theories on which calculations are based deliver two levels of accuracy. These are QM and MM. The former is more accurate, but as a consequence, calculations are more time-consuming. The latter, on the contrary, allows one to obtain a general view on the topic in a shorter time. The diversity of methods highly contributes to more successful experimental studies as it saves time and points out these research approaches that have a high probability of success.

2.1. Application to Estrogens

2.1.1. Ligand Docking Using Force Field Methods

- Principles of docking and re-docking

Protein–ligand docking is a technique used to predict the orientation and active conformation of a molecule in an active center. Based on this prediction, the binding energy between protein and ligand is calculated [47]. Ligand docking can identify the chemical bonds crucial for activity and the specific atoms/residues that are responsible for them. Target proteins are usually acquired from the RCSB Protein Data Bank (PDB) [48] (Table 2). Presence of a ligand in the crystallographic receptor structure simplifies and speeds up researchers' work, as a proper region for docking is already plainly indicated.

Table 2. Selected proteins involved in the metabolism of estrogens present in the RCSB PDB (Protein Data Bank) [48]. All structures were obtained using X-ray diffraction.

Protein	RCSB PDB Reference Code	Resolution (Å)	Incorporated Ligands
βER	5TOA	2.5	Estradiol
βER-LBD	1QKM	1.8	Genistein
Phosphorylated βER-LBD	3OLL	1.5	Estradiol, N-peptide linking
αER-LBD	3UUC	2.1	Bisphenol C
αER-LBD mutant	4Q50	3.07	4-hydroxytamoxifen
αER-LBD mutant	2QXS	1.7	Raloxifene
αER-LBD	2R6Y	2.0	SERM
17β-HSD	1IOL	2.30	17β-estradiol
17β-HSD	6MNC	2.40	Estrone
17β-HSD	6MNE	1.86	Estrone, NADP+
17β-HSD	3DHE	2.30	Dehydroepiandrosterone (DHEA)
17β-HSD	4FJ0	2.2	3,7-dihydroxy flavone
17β-HSD	4FJ1	2.3	Genistein
SULT1E1	1AQU	1.6	Estradiol, PAP cofactor
SULT1E1	4JVM	1.994	Flame retardant, PAP cofactor
Placental E1/DHEA STS	1P49	2.6	L-octylglucoside, *N*-acetylo-D-glucosamine, Ca^{2+}, PO_4^{3-}
CYP1B1	6OyV	3.101	Estradiol

Accessibility of the X-ray protein structures with docked ligands gives also possibility to prove the quality of the chosen simulating docking method and parameters [49]. This is performed via extracting the ligand from the available in the PDB structure and re-docking it to this crystallographic measurements-based protein. The results are obtained in a form of the root mean square deviation (RMSD) of atom positions [49]. It is calculated as differences between the original experimental atom positions in the crystallized structure and the theoretically obtained molecular docking results [50]. It is acknowledged that the used modelling tools are able to identify the correct ligand pose, are repeatable and reliable when RMSD is smaller than 2Å [51].

Molecular modelling performed on estrogen-related receptors is not an exception. As the studies are performed mostly to search for new agonists or antagonists, the ligands re-docked into the ERs or SHBG are mainly E2 [52–57] and 4-hydroxytamoxifen [52,58,59]. RMSD value in most of the studies varies from 0.26 to 1.4Å, which proves the correctness of the docking methods in finding the proper orientation of the ligand in the active site.

- Enzymes and receptors used as targets in estrogen-related docking studies

A great many proteins have been used in molecular docking studies of estradiol. They include 17β-hydroxysteroid dehydrogenase [60], progesterone receptor [61], protein disulfide isomerase [62], SHBG [63], CYP1B1 [64], steroid sulfatase [65], and even the voltage- and Ca^{2+}-activated K^+ channel β1 subunit [66]. In all studies, an emphasis is put on the importance of OH-hydrogen binding [67,68]. The relevance of this emphasis is confirmed by experimental results. One example is a study where the estrogen analog with the highest affinity, measured in a fluorescence polarization displacement assay, appears to have the second highest predicted affinity [67].

Docking serves either to investigate the binding of estradiol to the ER or to dock other molecules, potentially new drugs, such as potent and highly selective estriol analogues [69]. In the second case, information from previous studies on estradiol behavior in an active site serves as reference data [70–73]. This is a starting point for the commonly applied research sequence: investigation of E2 binding mode, comparison with the calculation results for potential drugs, and confirmation of the hypothesis by analytical techniques. Such a three-step process has been performed, for example,

for estrogen-dependent MCF-7 cancer cells [70]. Eighteen compounds with antiproliferative activity have been docked into the cavity where E2 normally binds. It was predicted that, compared to E2, an additional aromatic ring is involved in the binding mechanism. The prediction was confirmed by site-directed mutagenesis.

In some cases, allosteric modulators have been studied. This means that estradiol must be present in the binding site during the docking process [74,75]. To test the prediction of allosteric activity, in vivo experiments are often performed. For example, in one study [74] it was shown that one particular compound can properly fit into the region of the binding pocket, along with E2. Afterwards, this ligand was investigated in vivo and was demonstrated, indeed, to be a new ER-beta-selective, negative allosteric modulator of E2 binding.

Including an estradiol molecule in the docking studies helps to properly score the resulting data [76] and rank the tested molecules according to their binding affinity [77]. For large ligand sets, high-throughput screening with a pure agonist (estradiol), an antagonist (tamoxifen), and decoys (known non-binders) can be performed. This can aid in setting a laboratory's experimental testing priority, reducing the cost and time of its research, and boosting its effectivity [77]. This explains why in vivo and in vitro investigations are often combined with in silico ones [78,79]. Altogether, it enables the discovery of new possible drug molecules that could influence estradiol's signaling pathways [80].

- Docking studies of plant-derived potential xenoestrogens

Docking plant-derived substances into ER and comparing their binding energies and interactions (above all hydrogen bonding) with those of estradiol is a common practice. Such studies deliver information on conformational flexibility [81], the ability of the investigated molecules to selectively modulate ERα/β ability [82,83], their possible reductive influence on breast cancer risk [84], or their applicability to reduce menopausal symptoms [85]. Most importantly, data derived in this way very often show good consistency with experiments [86]. For the obtained data, a correlation with estradiol, but sometimes also with tamoxifen [75] (ER antagonist) or genistein [87,88], is found. This last substance serves as an important reference, as genistein has a higher affinity for ERβ than 17β-estradiol.

Even if docking into ER is most widely used, modulation of CYP450 activity by plant-derived substances with regard to estrogenic effects has also been examined [89]. Indeed, the same signaling pathways are regulated by CYP450–estradiol interactions, and there are plenty of data available on this topic. Moreover, other calculation methods are also applied, including ADMET and molecular dynamics (MD) [90,91] (for MD description and examples, see Section 2.1.3). Nowadays, in silico methods are a standard tool in plant xenoestrogens studies. Thus, molecular modelling is often performed in parallel to either in vitro [92] or in vivo toxicity studies [93], and in most cases an agreement between data obtained from both sources is found. This was the case in the docking-based binding affinities of compounds derived from *C. elegans* and their measured reproductive toxicity [93]. Wide examination of many substances of a natural origin is possible thanks to the thorough structural knowledge on the estradiol molecule bound to the ER (Table 2).

2.1.2. Quantitative Structure–Activity Relationship (QSAR)

QSAR methods use mathematical models to correlate structural characteristics with the biological activity of a set of compounds that have closely related structures [94]. Empirical and theoretical molecular descriptors are used. Conceptually, three main types of QSARs are known: ones based on fragment analysis of a system (here, a pharmacophore [95] concept is used), ones that consider the given system as a whole (descriptors are computed from scalar quantities), and 3D-QSAR. In this last type, descriptors are obtained by application of a force field (3D approach). To achieve a 3D target structure, either software-based alignment or manual superimposition on the crystallographic data must be performed [96]. This is a necessary step, as different ligand-binding modes and bioactive conformations are possible. Examples of 3D-QSAR are Comparative Molecular Field Analysis (CoMFA) [97] and

Comparative Molecular Similarity Indices Analysis (CoMSIA) [98]. From the created QSAR models, predictions on the bio-activity and toxicity of other molecules are made.

A smaller group of structure–activity studies are ones that are non-quantitative, namely SAR ones (structure–activity relationship). They include virtual screening with a pharmacophore and a large set of molecules. Afterwards, in vitro evaluation of the data is performed. This methodology has been used to determine potential 17β-HSD inhibitors [99,100] that could be applied in the treatment of osteoporosis provoked by estradiol deficiency.

Among all molecular modelling approaches, QSAR is one of the most commonly used to examine estradiol's structure and activity. Often, data on estradiol serve only as a reference for information gathered on new possible drugs [14,101] such as raloxifene derivatives [102] or estradiol metabolites. Through comparison with experiments, it has been shown that QSAR models have a high sensitivity and specificity in providing relative binding affinities (RBAs), where estradiol's RBA equals 100% [103]. Among the most important descriptors are ones calculated with quantum mechanics at the DFT level [104,105].

While QSAR is mostly applied to ER-binding, other proteins like CYP1A2 [106] and CYP1B1 [107] have also been targeted. One study used the CoMFA approach to model estriol's influence on CYP1A1 [106]. It focused on the inhibition of estradiol to mutagenic 4-OH estradiol transformation. Out of 90 steroid candidates, thioestrone was selected and shown to have the desired inhibitory ability. Its mechanism of action was revealed by molecular modelling. It is desired because thioestrone's -SH group is closer to the iron atom in the CYP1B1 heme than the -OH group in natural ligands of the enzyme, namely estradiol and estrone.

OECD Guidance identifies QSAR as an important element in toxicity evaluations. As a result, QSAR plays an important role in substance risk assessment [108]. It has been applied in studies dealing with mutations in enzymes that metabolize estradiol [109] and to study estradiol oxidation and emerging contaminants [110]. In the latter study, the most accurate DFT descriptors were used. This enabled a better understanding of the degradation mechanisms.

2.1.3. Advanced Docking Using Combined Quantum Mechanics/Molecular Mechanics (QM/MM) or Molecular Dynamics (MD) Methods

Molecular modelling with QM is much more accurate than MM methods. However, the direct use QM approaches in drug design is limited due to the cost and size of protein structures. In recent years, QM/MM has been gaining attention, as it allows one to consider a whole protein–ligand complex and not only a binding site [111,112]. Due to computational limitations, MM calculations on the outer part of a receptor deliver only approximate data. Nevertheless, QM/MM provides more knowledge on the protein's influence than is obtained when only the LBD region is considered.

More commonly applied methodology is MD. It simulates time-dependent processes and provides data that are otherwise unavailable [52,54]. For example, in case of ERs or SHBG ligands MD can point out whether the examined substance is receptor's agonist or antagonist. This assumption is based on the RMSF (root mean square fluctuation) value which is extracted from the MD trajectories. RMSF represents the flexibility of the amino acid residues [50]. Both RMSF and RMSD depend on the interactions between the protein and the ligand and are the result of the ligand movements in the active site trying to achieve the appropriate position [63]. If the same ligand undergoes the MD process in both ERα and ERβ and its RMSF and RMSD values significantly differ for these two receptors, it indicates that the investigated molecule probably occupies more favorably one of the investigated receptors. Such data leads to the hypothesis of the ligand's selectivity.

Moreover, in terms of the ERs, RMSF and RMSD values suggest whether the analyzed molecule is the receptor's agonist or antagonist [50,52–54,59]. As already mentioned in Section 1.2, the positioning of the H12 helix is differently influenced by agonists and antagonists. If the estrogenicity of the compound is known, comparison of RMSF values with the known molecule's estrogenic effect, can serve as an evaluation of the applied docking parameters.

Nevertheless, MD requires a lot of computation time, which increases with the system's size and simulation length. In return, it describes the dynamics of the system as well as entropic effects associated with the protein–ligand interaction.

MD can also provide quantitative estimates of relative binding affinities through techniques known as "free energy calculations". The most rigorous is the "free energy perturbation," or FEP family of methods [113–117]. To compare two ligands, say A and B, one introduces a model where both ligands are present, each with a partial occupancy. This is closely analogous to a crystallographic refinement where a particular group (ligand or side chain, say) has two possible conformations. Each ligand is assigned a weight between 0 and 1, say w_A and $w_B = 1 - w_A$. These multiply terms in the energy function involving either ligand. By varying the weights gradually, one can effectively remove one ligand and introduce the other. Thus, when $w_B = 0$, ligand A is fully weighted while interactions of ligand B with its surroundings have a zero weight: ligand B is "invisible" to its environment. Usually, interactions within the ligand are not weighted. As w_B changes from zero to one, B is introduced and A is removed. Intermediate weight values correspond to an "alchemical" mixture, where both ligands are partially present. MD simulations (or Monte Carlo simulations) are performed for a few w_A values, typically around 10. This series of simulations mimics a gradual, reversible replacement of A by B. Energy statistics are collected from all simulations. From these, a free energy difference between A and B can be obtained. The same process is carried out for the unbound ligands, solvated by a box of water. Subtracting the bound and unbound free energy changes yields the binding free energy difference. The method requires force field parameters for each ligand but has no other adjustable parameters. The tradeoff is that not one, but several MD simulations are required, and these should be sufficiently long. Indeed, FEP accuracy is limited by the MM force field, but also the amount of conformational sampling that is carried out. The method can also be used to compute the binding free energy changes due to point mutations of the protein. In this case, a particular residue is modeled with two side chains, each having a partial occupancy. Nowadays, FEP can be applied to one ligand or mutation per day on a medium-sized computer cluster.

A simplified version of FEP is to use only two simulations per ligand: one bound to the protein and one in solution. From these, binding free energies can also be obtained, if one is willing to extrapolate from a w_A value of unity to a value of zero. To counter the use of such a large extrapolation, one introduces empirical weighting factors that multiply the interactions between the ligand and its surroundings (protein or solution). Usually one is applied to the electrostatic interactions and one to the Lennard–Jones interactions. The extrapolation and use of interaction energies have led to the name Linear Interaction Energy, or LIE method [118–120]. For a thorough review of its theoretical basis, see [121]. The tradeoff for its speed is that experimental data are needed to adjust the values of the empirical weights, which are not very transferable between different proteins and classes of ligands. Once the weights are optimized for the molecules of interest, predictions can be made.

In many studies, estradiol has been simulated in a complex with ER. The QM/MM approach seems to be the most accurate. It uses a QM description of the ligand and its binding pocket, but thanks to the MM description of more distant protein regions, it preserves information on the whole enzyme's impact [122]. Very recently, QM/MM elucidated the important role of estradiol's D-ring in the active site of ERα [123]. What is more, it helped to understand the influence of each enzyme segment on the ERα-agonist (estradiol, diethylsilbestrol) binding [124]. Most importantly, the binding energies of E2 and DES correlated well with experimental agonist binding affinities for the ER.

To analyze the changes in the investigated complexes upon DNA binding, not only MM, as with estradiol and DNA [125], but also MD has been applied [126]. The latter study identified specific bases within the aptamer (short-stranded DNA/RNA, binds only specific molecules [127]) and demonstrated the importance of water-mediated hydrogen bonds in the aptamer–estradiol complex.

MD is also often used to describe the effects of enzyme mutations on ligand binding. This methodology has been applied in estradiol studies, to compare wildtype and mutated CYP1B1 [128], which is mostly responsible for the 2-OH-hydroxylation of estradiol. MD serves also as a tool to explain

the results of molecular docking. For example, it has been stated that the interactions between human α-fetoprotein and agonists (estradiol, estrone, diethylsilbestrol) were caused by van der Waals forces, whereas binding of antagonists (tamoxifen and its analogues) was equally based on hydrophobic and electrostatic interactions [129]. Moreover, information from MD simulations can be used to construct a pharmacophore in order to screen protein databases for a desired type of ligand. This methodology was used to search for substrates and inhibitors of the estrone-SULT [130]. Nine selected molecules were consistent with the ones indicated by the experiment.

Another nuclear receptor, the farnesoid X receptor (FXR), was used as a test case for free energy calculations in 2018 [116,117,120]. FXR is involved in regulating bile acid, lipid, and glucose homeostasis [131], and it has been linked to hepatocarcinogenesis [132]. Its hormone binding site is hydrophobic with few conserved interaction motifs and strong induced fit effects. With FEP, mean errors were about 1.5 kcal/mol for relative binding free energies of around 30 ligands, and the largest errors were about 2.5 kcal/mol. In one of the studies [117], ligands were first docked to the receptor, then compared using FEP. LIE gave similar errors for the same protein and 47 ligands [120], at a lower computational cost, but required optimization of the two adjustable LIE parameters using a subset of the ligands.

An earlier LIE study [119] considered the ER binding of estradiol and a series of xenoestrogens. A training set of 19 ligands was used to optimize the LIE parameters. A mean unsigned error of 0.6 kcal/mol was then obtained for a test set of 13 ligands. Several binding poses (3–4) were considered for each ligand; this was not too expensive because only short MD simulations were run.

2.1.4. Other MD-Based Studies of Estrogens

- Membranes

As ERs are located in either the cytoplasm or the lipid bilayer (mER, membrane ER) [27,133], and estradiol itself is a steroid hormone, closer insight into the ligand's interaction with the cell membrane is an obvious research target. Since transfer through this cellular barrier is not a stable state, but a dynamic process, MD calculations could be seen as a method of choice. Recent studies revealed [134] that, regarding estradiol's long axis and the lipid acyl chains, E2 adopts a perpendicular position in the membrane. By having four rings located near the membrane interface, participation of the hormone's 3-OH and 17β-OH groups in hydrogen bonds and electrostatic interactions with the lipids are possible.

Combining MD and QM enables further research into E2 membrane crossing. It provides information not only about the E2 orientation in the membrane (MD) but also about the strength of the electrostatic potential mapped on the electron density surface (QM) [135]. These data, derived from a HDL disc model, enable deeper insight into the mechanism of E2 incorporation into lipid membranes and is an important step forward to develop tissue-specific discs encircled by a membrane, which would serve as transporters for E2 or its derivatives.

Similar methodology (QM, MD) was used to study the removal of hormonal pollutants from water. It has been shown that by using high levels of salinity, which increases the strength of hydrogen bonding and hydrophobic interactions, one can perform a membrane-based sorption of 17α-ethinyl estradiol on the polyethersulfone membrane [136].

- Nanotubes

Nowadays, estradiol is a relatively common water pollutant, and numerous studies have been performed to find a reliable tool for its removal. One possibility is to use nanotubes. For that purpose, within the molecular modelling approach, mostly single-walled carbon nanotubes (SWNTs) [137] are used. Free energies of adsorption have been calculated with a QM approach. In some cases, MD was also implemented [138]. The target ligands for these studies were 17β-estradiol and its medically useful derivative 17α-ethinyl-estradiol [139–141]. These simulations revealed the adsorption energy,

a preferential sorption among different nanotubes and estradiol derivatives, and provided a molecular explanation for the observed results [142].

2.1.5. Density Functional Theory (DFT) Calculations in the Study of Estrogens

DFT is a QM approach that determines the ground-state properties of a many-body system by applying the electron density concept. The underlying concept is the Hohenberg–Kohn theorem [143], later developed into the Kohn–Sham theory (KS-DFT) [144]. Firstly, the energy of the system for a non-degenerate stationary state is uniquely determined by its electron density, which depends on three spatial coordinates. For this reason, the energy is expressed as a functional of the (scalar) electron density function. Secondly, according to the H-K theorems, the minimum energy occurs for a unique, precise electron density in the ground state. KS-DFT includes the Coulomb interactions between electrons and considers the energy of the exchange and correlation interactions. For a long time, the dispersion energy (the energy of the long-ranged electron correlation) [145] represented a difficult problem, as it is a time-dependent phenomenon, and it was not included in KS-DFT. However, nowadays, dispersion corrections are available and can be included in DFT functions [146]. Therefore, application of DFT leads to the highest obtainable accuracy in calculations. The only existing drawbacks are the risk of underestimating the energy [147] and the time needed to acquire the results. DFT-based calculations are especially widely used in solid-state studies.

Although application of DFT is mostly concentrated on the investigation of single molecules (see the paragraphs below), it is also used to determine total binding energies between a ligand and protein. This is the case for systems composed of ER, SERMs, and two widely used ER antagonists: 4-hydroxy tamoxifen (4OH-T) and raloxifene (RAL) [148]. The results show that the 4OH-T-SERMs set binds more strongly to the ER than the RAL-SERMs set. This is fully in agreement with the experimental data and, once more, as many other studies, confirms the high accuracy of the DFT calculations.

- Crystal structure prediction

DFT is the theoretical basis for periodic calculations performed on solids, often pharmaceuticals, in order to find and depict new polymorphic forms of drugs or potentially bioactive molecules [149]. It can also be a part of the Crystal Structure Prediction (CSP) approach. Such methodology has been recently used to study crystal structures of 17β-estradiol. As a result, an estradiol hemihydrate has been computationally determined [150]. DFT calculations are also often necessary to refine a crystal structure obtained from powder X-ray diffraction (PXRD) experiments. These calculations are mostly consistent with experimental data, as in the case of estradiol ethinyl cocrystals [151].

DFT-based methodology has also been applied to examine the dissolution process in a study of estradiol cocrystals [152] and to calculate the free energy of solvation in the estradiol–ER complex [153].

- NMR and vibrational properties calculations

One of the most common DFT applications is to predict NMR and vibrational data. For NMR properties, GIAO [154] and GIPAW [155] methods are implemented. Estradiol being relatively complex (many atoms, presence of both aromatic and non-aromatic rings) could cause computational problems. For this reason, it has been part of a many-ligand study to prove the applicability of the GIAO method as well as its ability to calculate J_{HH}, J_{HC}, and J_{CC} NMR coupling constants [156,157]. Generally, the results obtained were satisfying. However, nowadays it is more common to study the solid state with the second method. GIPAW NMR calculations were also done recently for a new polymorph of 17β-estradiol [158]. The calculations helped to improve the assignment accuracy of chemical shifts obtained from the experiment and, therefore, to elucidate the structure of a new anhydrous estradiol form Figure 1.

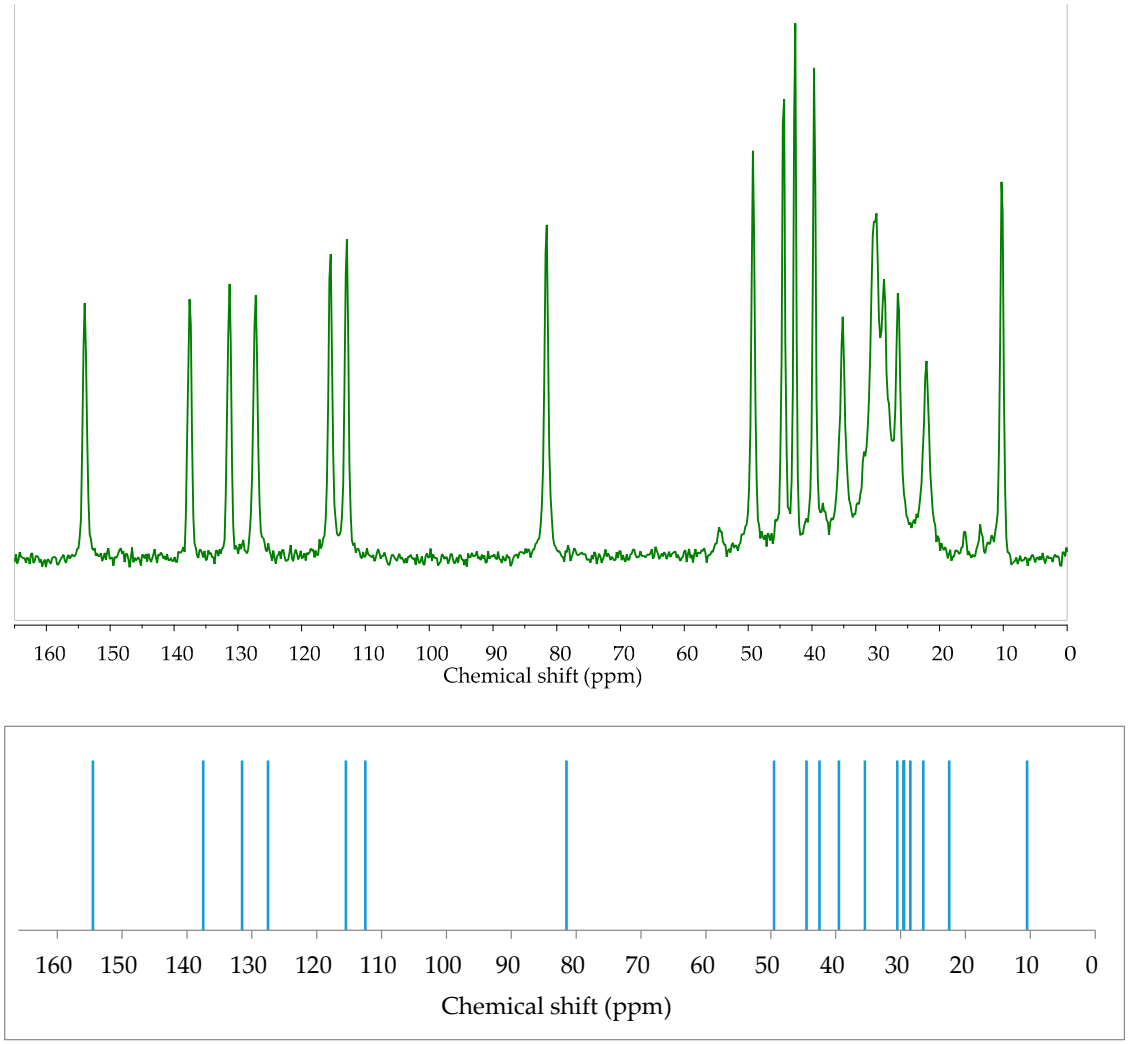


Figure 1. Experimental (top, green) Cross Polarization Magic Angle Spinning (CP MAS) and calculated (bottom, blue) GIPAW ^{13}C solid-state NMR spectra of E2. Very good agreement between calculated and experimental values proves the usefulness of DFT calculations in solid-state analysis of estrogens. More details in [158]. Source: Author's archive.

Another study showed the applicability of DFT-NMR to explain the nature of a more complex system: the transversal distribution of 17β-estradiol in lipid membranes [159]. NOESY 2D NMR spectra contained cross-peaks between the hormone and lipids. Here, too, DFT calculations helped to properly interpret the experimental data and, as a consequence, greatly aided in describing the position of the estradiol aromatic ring in the membrane. An implication of such study is to increase our understanding of estradiol's transfer through a lipid bilayer [159,160]. A second area where DFT methodology is widely applied is in calculating vibrational properties. The direct usefulness of the method is manifested through its contribution to accurate assignment of the vibrational modes in IR or Raman studies. This has been implemented, for example, in investigations of estrogens [160] and estradiol-17 valerate [161]. In most reported cases, computationally generated spectra were in very good agreement with the experimental data.

From a wider perspective, DFT calculations enable one to properly describe the examined subject. Additionally, in case of estrogens and estrogen derivatives DFT-derived spectra (both vibrational [162,163] and NMR [158]) stay in a good agreement with the experimental data and are often the only way to properly assign bands (Figure 2). In order to obtain theoretical spectra which ideally meet the experimental ones, scaling factors must be implemented [164]. Thanks to such

combination of the theoretical and experimental approach, the first full interpretation of estriol IR spectrum could have been published [164]. The analogical situation has been reported for E1, E2, E3, and ethynylestradiol Raman spectra [163]. Simulation was necessary to identify unique marker peaks in the finger-print region what was useful to differentiate between very similar estrogen structures.

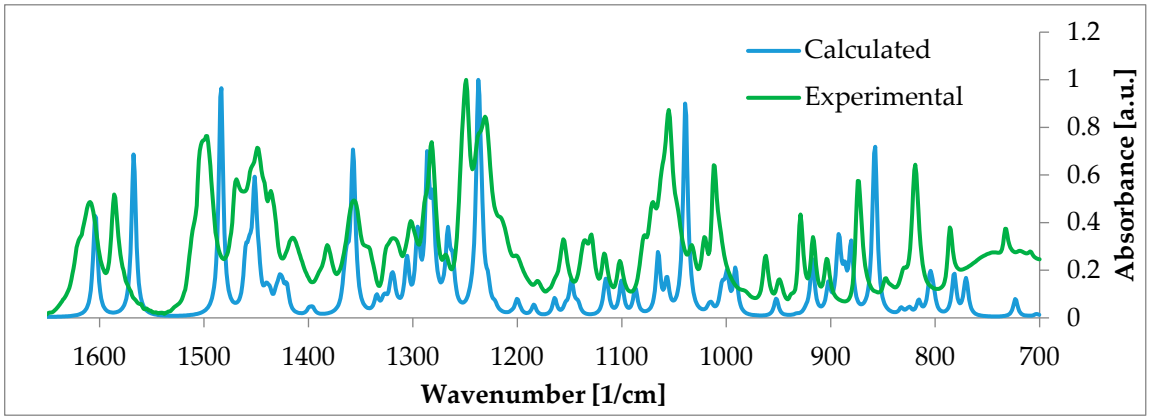


Figure 2. Experimental (green) and calculated (blue) IR spectra of β-estradiol hemihydrate, selected range 700–1650 cm^{-1}. Such calculations enable proper band assignments and thus facilitate the spectrum analysis. More information can be found in [158]. Source: Author's archive.

Another example shows that the vibrational frequency calculations of estradiol alone and in monohydrated form [165] has given insight into hydrogen bond formation by estradiol's D-ring. This has been used to discuss the relationship between the stability of hydrated clusters and the estradiol conformation. In another study [166], DFT-based IR and Raman frequencies were used to investigate estradiol and tamoxifen structures. This helped to understand the intermolecular interactions made by these two molecules and to interpret opposite estrogenic effects. However, it should be mentioned that in the studies of estrogens, Raman spectroscopy is used less frequently than IR. This is due to the inherently weaker signals and common presence of fluorescence interference from the contaminants [163]. Apart from IR and Raman, low-frequency vibrations could also be calculated with the help of DFT. A good example is the assignment of vibrational modes in terahertz spectra for testosterone, estradiol, and estrone [167].

- Removal of estrogenic pollutants

With regard to estrogen removal from the environment, QM calculations are performed not only on nanotubes but also on other sorbents (e.g., lignocellulosic material) [168]. The adsorption energy shows that this adsorbent could be used to remove all three main estrogens, E1, E2, and E3, from a solution by means of liquid phase extraction. Another extended study revealed the applicability of reduced graphene oxide modified with silver nanoparticles in electrochemical detection of estriol [169]. Firstly, MD simulations at 1000 K were performed in order to obtain 100 conformers of estriol. Later, these structures were used in a semi-empirical Hartree–Fock geometry that was pre-optimized with solvent simulated via the conductor-like Screening Model COSMO [170]. Afterwards, the most stable conformers were fully optimized by the DFT-based software. In this step, the solvent was included via application of the Polarizable Continuum Model [171]. Then, on the structure with the lowest energy, MD at 300 K was performed. Electric properties of the newly obtained conformers were determined with DFT calculations. This study is a good example of the wide range of in silico methods that can be applied to molecularly model the investigated subject (i.e., to develop a method to detect estriol in tap water and urine samples).

2.2. Application of Various Molecular Modelling Methods in the Study of Xenoestrogens

2.2.1. Various Molecular Modelling Methods Applied in Xenoestrogen Studies

The same methods as for estradiol, above, are also used to investigate xenoestrogens. They include QM/MM [172] and QSAR approaches [173–175] and confirm the importance of the hydrogen bond between xenoestrogen molecules and the His524 residue in the active site. A variety of xenoestrogens, such as bisphenol A and C, butylparaben, 4-octylphenol, DDE, phthalate, zaeranalol, estradiol, 4-OH-tamoxifen [176], and several proteins (ER [177], SHBG [178]), have been studied. These studies either looked for structural similarities between different ligands or focused on one specific molecule, such as zearalenone [69]. The latter study suggested a lack of agonistic activity against the ER due to the lack of any stable, functionally active conformation of the tested molecule in the LBD. On the other hand, according to QSAR analyses performed on zearalenone analogues [179] and metabolites [180], these zearalenone-related compounds show some estrogenicity due to the presence of a keto/hydroxyl group, a trans double bond in the macrolide ring, and two hydroxyl groups in the aromatic ring, which participate in binding to ER.

As endocrine-disrupting chemicals are present in the environment, research into their removal is constantly being performed. For example, to study dioxin adsorption on graphene, DFT [181,182] and MD [183] calculations have been undertaken.

2.2.2. Bisphenol A

- Bisphenol A (BPA)–ER complex studies

Bisphenol analogues are among the most studied xenoestrogens. Molecular modelling helps show how these closely related structures adopt agonist/antagonist orientations in the estradiol binding pocket [184] and delineate the binding modes of each bisphenol molecule [185]. A deeper study, concentrating not only on the pure ligand–receptor binding but also including the allosteric effects and application of MD calculations, showed that BPA causes changes in a full-size receptor, and its effect is not limited to the separate domains [186].

A separate set of studies investigated BPAs with halogen substituents on the phenolic rings. All studies showed that a hydrogen bond with His524 for an agonist and with Thr347 for an antagonist was created, exactly as in the case of estradiol and 4-OH-tamoxifen, respectively [187,188]. What is more, just as for estradiol, the stability of helix H12 is crucial in halogenated BPAs–ER complexes [188,189]. The in silico results have been confirmed by experimentally measured affinities.

Although BPA mimics estradiol's action in the LBD, the QM/MM study revealed that, in comparison to other tested EDCs, it exhibited lower estrogenic activity, probably due to the lack of interaction with His524 [190]. In turn, application of MD helped to elucidate mechanisms driving BPA–ER binding. According to that study, direct hydrogen bonds and hydrophobic interactions are responsible for the binding [173]. Like the previously mentioned experiment, this one confirmed that the ER binding affinity is slightly lower for BPA than for estradiol. MD not only helps to elucidate bound conformations and binding energies between LBD and BPA in the ER [191], but it also gives insight into the influence of that binding on the whole protein, including the DBD. One of the studies [185] showed that the allosteric effects in the LBD due to BPA binding could cause relaxation of the DBD and, therefore, alter ER's function. Other researchers reported the influence of bisphenol compounds on the protein's allosteric modulation, altering the Helix12 stability and reducing the recruitment potency of co-activators [187]. This knowledge can be useful in the process of estimating the toxicity of compounds.

- Risk assessment and removal attempts

In 2012, a protocol for in silico risk assessment of BPA on the ER was proposed [189]. Later, many studies dealing with BPA removal from water were performed [192,193]. Applied molecular

modelling techniques include DFT and MD. Since BPA results from the depolymerization of, for example, polycarbonates, both BPA and the initial polymer have been studied under periodic boundary conditions (pbc) [194]. The same DFT-pbc approach has been used to model the photocatalytic degradation of BPA caused by cobalt-doped BiOCl nanosheets [195] and by the effect of humidity combined with UV irradiation [196]. Another example is the evaluation of BPA's binding to microextraction coatings [197]. In this case, dispersion-corrected DFT was applied.

2.2.3. Phthalates

The second group of xenoestrogens most widely examined by molecular modelling are the phthalates [198,199]. In silico investigations showed that estradiol has a lower ER binding affinity than phthalates. The highest RBA is exhibited by monophthalates [200]. In the case of SHBG binding, score values suggest that short-chain phthalates are more potent than long-chain ones [201]. This agrees with known experimental data.

To look for an effective method of phthalate removal from water, as with estradiol, DFT and MD calculations on SWNT–pollutant complexes have been performed [202]. The adsorption energy has been calculated, and the adsorbent's chemical groups responsible for the binding have been determined. MD has also been used to examine polymer–solvent interactions while looking for a new, more eco-friendly substitute for plasticizers [203,204].

2.2.4. Technical Aspects of Calculations Performed on (xeno)Estrogens

The computational method most commonly applied in the analysis of (xeno)estrogens is molecular docking. Available publications show that, for this purpose, the most common software packages are Maestro Schrödinger and AutoDockTools. The former is also widely applied in Virtual Screening Workflow (Table 3, N° 9–11), which includes high-throughput virtual screening (HTVS) [205] and molecular docking with either standard or extra precision (SP, XP). In both cases, the OPLS 2005 force field is used. The mentioned (N° 9) consensus score is an effective score that enables ranking of the investigated ligands. It is a combination of different scores, like DockingScore (GlideScore + state penalties for protonation) [206], MM/GBSA Score (binding free energy calculations based on the MD trajectories) [207], and QSAR Score.

Table 3. Selected technical computation data in terms of ER and (xeno)estrogens regarding the publications cited in this article.

N°	Code/Software Used	Force Field or DFT Functional and Basis Set	Type of Calculation	Ref. Method	Ref. in Article
1	GOLD		Molecular docking	[208]	[57,69]
2	-Ghemical 2.95 -Swiss Dock	-Tripos 5.2 -CHARMM	-Geometry optimization -Molecular docking	[209–212]	[70]
3	-Swiss model -Hex 8.0, HADDOCK	-OPLS	**-Homology of receptors** -Molecular docking	[212–215]	[87]
4	-Swiss model -SybylX	-Tripos 5.2	**-Homology of receptors** -Molecular docking	[209,214,216]	[84]
5	Maestro Schrödinger	OPLS 2005, Glide SP, XP	Molecular docking	[217,218]	[83,85,86]
6	Maestro Schrödinger	MMFF94	Geometry optimization, molecular docking	[217–219]	
7	-Gaussian09W -AutoDockTools	-B3LYP/6-31G(d) -AutoDockZN	-Geometry optimization -Molecular docking	[220–223]	[169,184]

Table 3. *Cont.*

N°	Code/Software Used	Force Field or DFT Functional and Basis Set	Type of Calculation	Ref. Method	Ref. in Article
8	Gaussian03	B3LYP/6-311++g**, PCM	Hydration enthalpy	[220,224]	[183]
9	Maestro Schrödinger	ZINC database OPLS 2005, Glide SP eHiTS docking module consensus score	Energy minimization **HTVS** rank	[205,217,218, 225]	[75]
10	Maestro Schrödinger	OPLS 2005, Glide SP, XP	**HTVS**	[205,217,218]	[76]
11	Maestro Schrödinger	OPLS 2005, Glide SP, XP	**Segregation**: agonists/antagonists	[217,218]	[77]
12	Maestro Schrödinger	-OPLS 2005, Grid (Glide) -Desmond OPLS 2005	-Molecular docking, MD -ADMET parameters	[217,218]	[90]
13	-Maestro Schrödinger -AMBER14 -AMBER14	-OPLS 2005, Glide -FF03 (protein) GAFF (ligand) -MMPBSA, MMGBSA	-Docking -MD -Binding free energy, decomposition energy	[207,217,218, 226–228]	[91]
14	-MOPAC2016 -Gaussian09 -Gabedit package	-PM6 in HF, COSMO model -B3LYP, PCM -Verlet algorithm	-Pre-optimization, solvent model -Optimization (DFT), solvent model -MD	[156,220,224, 229,230]	[169]
15	-GOLD -GROMACS -Swiss Param Tool	-CHARMM27 -CHARMM27	-Molecular docking -MD -Ligand parametrization	[209–211,231]	[180]
16	-SybylX -AMBER11 -AutoDock 4.0	-Tripos 5.2 -AMBER -AutoDockZN	-Geometry optimization -MD -Molecular docking	[216,221–223, 226–228]	[186]
17	-Gaussian09 -LeDock -AMBER12 -AmberTools14	-B3LYP/-cc-pVTZ -CHARMM -AMBER -MM/GBSA	-Geometry optimization -Molecular docking -MD -Binding free energy	[207,220–223, 226–228]	[188]
18	-Gaussian09 -Molegro Virtual Docker -AMBER Tools	-B3LYP/6-311++G(d,p) -AMBER -AMBER03	-Molecular electrostatic potential -Molecular docking -MD	[220,226–228]	[187]
19	-Maestro Schrödinger -Gaussian09 -AMBER10	-OPLS 2005 -HF, 6–31G* -GAFF (ligand), ff03 (protein)	-Molecular docking -Geometry optimization -MD	[217,218,220, 226–228]	[189]
20	-VASP -GROMACS	-PBE GGA (DFT-D3) -GROMOS96	-Geometry optimization -MD	[230,232]	[192]
21	-GROMACS -AutoDock Tools -AutoDock Vina, Hex8.0.0 GROMACS	-AutoDockZN -AutoDock Vina, GROMOS96	-Energy minimization -Molecular docking -MD	[221–223,231]	[199]

Table 3. *Cont.*

N°	Code/Software Used	Force Field or DFT Functional and Basis Set	Type of Calculation	Ref. Method	Ref. in Article
22	-NAMD -Spartan04	-Charm CMAP FF -HF 3-21G	-MD -QM	[233,234]	[125]
23	-Gaussian 03 -AutoDock -AMBBER	-B3LYP/6311**G -AutoDockZN -PM3/Amberff14SB FF	-Geometry optimization -Molecular docking -QM/MM	[220–223,226–228]	[114]
24	-GROMACS -Gaussian 09	-CHARMM (MM) -GGA-D2 (QM)	-Geometry optimization -DFT calculations	[210,211,220–223,229]	[115]
25	-Maestro Schrödinger -AMBER Tools	-OPLS 2005 Glide -B3LYP/Amberff14SB	-Protein, ligand preparation (geometry optimization), molecular docking -QM/MM	[217,218,226–228]	[124]
26	-Crystal Predictor -Crystal Optimizer (Gaussian) -DMACRYS	-PBE0/631G(d,p)	-Conformations -Geometry optimization **CSP** -Intermolecular lattice energies	[220,235–237]	[150]
27	-GULP, DFTB+ -VASP	-optB88 level	-Geometry pre-optimization **CSP** -Geometry re-optimization	[232,235–237]	[181]
28	DMol3	DNP basis set, PBE GGA	Geometry, energy optimization	[238,239]	[182]
29	CASTEP	GGA PBE	DFT, NMR	[239,240]	[157]
30	CASTEP	GGA PBE	DFT, structure parameters calculation	[239,240]	[195]
31	Gaussian09W	B3LYP/631G(d)	DFT, IR	[238,241]	[164]
32	Gaussian09	M05-2X/6-311++G**	DFT, IR	[238,241]	[165]
33	Gaussian09W	B3LYP/6-31G (d,p)	DFT, Raman	[238,241]	[166]
34	Gaussian	B3LYP/6-31G(d,p)	DFT, IR	[238,241]	[167]

The most widely used are commercial codes: Maestro Schrödinger [217], CASTEP [240], GOLD [208], Gaussian [220], AMBER [226], SybylX, VASP [220]; academic codes: AutoDock [209], CHARMM [210], GROMACS [232]. (AMBER and CHARMM are names of both the codes and the force fields.) AMBER (Assisted Model Building with Energy Refinement), CHARMM (Chemistry at HARvard Macromolecular Mechanics), CSP (Crystal Structure Prediction), GAFF (General AMBER Force Field), Glide (Grid-based Ligand Docking with Energetics), GOLD (Genetic Optimization for Ligand Docking), GROMACS (GROningen MAchine for Chemical Simulations), HADDOCK (High Ambiguity Driven protein-protein DOCKing), HTVS (high throughput virtual screening), MM/GBSA (Molecular Mechanics/Generalized Born Surface Area), PCM (polarizable continuum model), VASP (Vienna ab initio Simulation Package), Glide SP (Standard Precision), XP (Extra Precision). References in the last column refer to articles already cited in this review. These are examples of application of the listed methods in (xeno)estrogens research. References in the fourth column refer to articles that describe the theoretical basis of the listed software and calculation methods.

Virtual Screening Workflow makes it possible to screen large sets of ligands. It helps to differentiate between 'actives' (compounds active against the target protein), 'decoys' (compounds of known non-activity against the target protein), and 'inhibitors/activators' (potential bio-active substances). As a consequence, the Virtual Screening Workflow approach guides future ligand synthesis and helps in setting a priority for in vitro testing.

Often, a computational step following molecular docking is MD. Here, the most applied codes are GROMACS and AMBER using CHARMM and AMBER force fields, respectively. In most cases, the TIP3 (transferable intermolecular potential with three points) solvent model is applied [242].

Computational methods applied in (xeno)estrogens studies that deal not with a solvent environment but with a solid state are Crystal Structure Prediction (CSP) (N° 26, 27) [235–237] and DFT-based calculation of spectroscopic (IR, Raman, NMR) properties (N° 28–34). For the latter, the most commonly applied codes are CASTEP and Gaussian with GGA PBE [239] and B3LYP [241] functionals, respectively. These two functionals seem to be the most reasonable for the investigated subjects. GGA PBE establishes the non-homogeneity in the electron density and leads to more precise results, which in turn is of high importance in NMR spectra calculations. B3LYP is a hybrid functional based on combining DFT and Hartree–Fock theories and finds its application in spectroscopic spectra generation.

On the contrary, CSP [235–237] is a multi-step and much more complicated methodology, as it implements both MM and QM and sometimes even MD. Firstly, MM calculations are performed to generate and rank possible compound conformations. Afterwards, the selected conformers are subjected either to ab initio calculations on a molecule or to DFT-D (dispersion corrected) [146] calculations performed on the whole crystal structure. This enables one to observe conformational polymorphisms in the first case and packing polymorphisms in the second. The lattice energies obtained could be adjusted if kinetic factors (like temperature) are included. For that purpose, time- and computational power-consuming MD must be applied.

In contrast to molecular docking or MD simulations, one of the most important calculation methodologies for (xeno)estrogens, QSAR, is independent of the protein and based solely on the ligand structure. This explains why different codes must be applied for QSAR. Their application to (xeno)estrogens has already been described in detail, and the available codes have been compared in [14].

The above-mentioned codes and parameters cover the most common calculations. However, it is impossible to point out the best ones due to the insufficient number of studies. We can only observe that, out of the gathered data, some standard methodologies emerge.

However, a comparison of the applied methodologies in terms of their usage as well as their advantages and drawbacks is possible. The most important aspects have been gathered in Table 4.

Table 4. Comparison of the calculation methods used in (xeno)estrogen investigations.

Calculation Method	Pros and Capabilities	Cons and Limitations
Molecular docking	-Explanation of a molecular basis for protein–ligand binding -Relatively short calculation time -Enables virtual screening for active compounds	-Lower accuracy when compared to QM methods -Significant increase in time and complexity of calculations when combined with QM (QM/MM)
QSAR	-Evaluation of estrogenicity -No protein preparation needed	-No receptor–ligand binding data -Large set of high-quality experimental data needed to obtain accurate results
QM (DFT-D)	-High accuracy of calculations -Simulation of IR, Raman, NMR spectra -Thermodynamic calculations	-Long calculation time -A lot of computational power needed -Usually limited to small molecules and systems such as estrogen complexes, salts, co-crystals, etc.
QM/MM	-High accuracy of calculations in the binding area (QM) -Consideration of a whole complex (protein–ligand) with emphasis on the binding pocket	-Calculation time elongated due to QM -Limitation of the QM-calculated area
MD	-Simulation of dynamical processes -Possibility to perform DFT-MD	-Significantly longer time required

To predict the binding affinities or the interactions between the (xeno)estrogens and biomacromolecules, either simple molecular docking or more sophisticated methods such as QM/MM, MD/MM, or FEP can be used. Notably, the more sophisticated methods require not only more specialized software but more computational time and power. Since, to the best of our knowledge, no study has been reported comparing the accuracy of various ligand docking methods applied to the particular group of (xeno)estrogens, no specific indications can be provided.

When focusing on the structural and physicochemical properties of estrogens and xenoestrogens, DFT-based methods have proven their high accuracy and reasonable calculation time. Therefore, such computations can be performed to obtain structural, spectroscopic (IR, Raman, NMR), and thermodynamic data of estrogens, xenoestrogens, their complexes, and solid-state forms such as solvates, salts, and co-crystals. Standard DFT functions (B3LYP for isolated compounds and PBE for periodic structures) have been found to be accurate in multiple studies.

3. Conclusions

In this review, it was clearly shown that molecular modelling methods are valuable tools in studies on estrogens and xenoestrogens. Their relatively low cost, requiring only certain specialized software licenses and computing servers, their increased personal and environmental safety, and their reasonable accuracy make molecular modelling methods unique and modern tools for these studies. In this article, the most common biomolecules studied using molecular modelling were presented, together with appropriate references to the published results. This group of molecules is composed mostly of enzymes participating in the metabolism of estrogens, along with estrogen receptors and even specific nucleic acid domains. While most studies focused on predicting the affinities of small molecules (ligands) to the chosen receptors, the computational research is not limited to this aspect. Another important role for modelling is to explain the conformational changes resulting from binding. Such in silico studies would not be possible without the very large number of already deposited, high-quality crystal structures of estrogen-related proteins that can be easily accessed and used in molecular modelling studies. An overview of those structures was presented in this review. The oldest studies in which molecular modelling was used to study the biochemistry of estrogens focused on molecular docking with molecular mechanics. More recent studies have used more sophisticated methods such as molecular dynamics or combinations of quantum mechanics and molecular mechanics. Further, in this review, it was shown that computational studies concern not only interactions between estrogens and biomacromolecules, but they also can be used to describe phenomena such as migration of estrogens through lipid bilayers or their adsorption on various materials. This can help predict the most efficient way to remove them from the environment when treated as pollutants. Further, it has been shown by multiple examples that quantum molecular modelling methods, such as those based on density functional theory, can be successfully used in structural studies on new solid forms of estrogens such as salts, co-crystals, hydrates, and polymorphs as well as on the complexes of estrogens with other molecules (i.e., cyclodextrins). In addition, the possibility to accurately calculate vibrational and NMR properties can be very helpful to explain spectroscopic results. Finally, we presented similar molecular modelling studies on xenoestrogens such as Bisphenol A and phthalates. Therefore, taking into consideration the versatility and confirmed accuracy of molecular modelling methods, it is not surprising that they have been listed in the specific guidance for studies on EDC published by international organizations such as ECA, EFSA, and OECD.

Supplementary Materials: The following are available online at http://www.mdpi.com/1422-0067/21/17/6411/s1.

Author Contributions: Conceptualization, A.H.M. and Ł.S.; investigation, A.H.M. and Ł.S.; writing—original draft preparation, A.H.M. and T.S.; writing—review and editing, A.H.M., Ł.S., T.S., and D.M.P.; visualization, A.H.M., Ł.S., and D.M.P.; supervision, A.H.M. and Ł.S.; project administration, Ł.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

17β-HSD	17β-hydroxysteroid dehydrogenase
ADMET	Adsorption distribution metabolism elimination toxicity
BPA	Bisphenol A
CoMFA	Comparative molecular field analysis
DBD	DNA binding domain
DFT	Density functional theory
E1	Estrone
E2	Estradiol
E3	Estriol
E4	Estretrol
ER	Estrogen receptor
EDCs	Endocrine-disrupting chemicals
FEP	Free energy perturbation
H-K	Hohenberg–Kohn theorems
K-S	Kohn–Sham theorems
LBD	Ligand-binding domain
LIE	Linear interaction energy
MD	Molecular dynamics
MM	Molecular mechanics
PDB	Protein Data Bank
QM	Quantum mechanics
QSAR	Quantitative structure–activity relationship
RBA	Relative binding affinity
SERMs	Selective estrogen receptor modulators
SHBG	Sex hormone-binding globulin
STS	Sulfatase
SULT	Sulfotransferase

References

1. Bennink, H.C.; Verhoeven, C.; Zimmerman, Y.; Visser, M.; Foidart, J.-M.; Gemzell-Danielsson, K. Clinical effects of the fetal estrogen estetrol in a multiple-rising-dose study in postmenopausal women. *Matur. Eur. Menopause J.* **2016**, *91*, 93–100. [CrossRef]
2. Singleton, D.W.; David, W.S. Xenoestrogen exposure and mechanisms of endocrine disruption. *Front. Biosci.* **2003**, *8*, 110–118. [CrossRef] [PubMed]
3. Bennink, F.C.; Holinka, C.F.; Visser, M.; Bennink, H.J.T.C. Maternal and fetal estetrol levels during pregnancy. *Climacteric* **2008**, *11*, 69–72. [CrossRef]
4. Schreiner, W.E. *The Ovary in Labhart, A. Clinical Endocrinology: Theory and Practice*; Springer: Berlin/Heidelberg, Germany, 2012; p. 548.
5. Kuhl, H. Pharmacology of estrogens and progestogens: Influence of different routes of administration. *Climacteric* **2005**, *8* (Suppl. 1), 3–63. [CrossRef]
6. Blackburn, S. *Maternal, Fetal, & Neonatal Physiology*; Elsevier: Amsterdam, The Netherlands, 2014.
7. Fait, T. Menopause hormone therapy: Latest developments and clinical practice. *Drugs Context* **2019**, *8*, 1–9. [CrossRef]
8. Tofovic, S.P.; Jackson, E.K. Estradiol Metabolism: Crossroads in Pulmonary Arterial Hypertension. *Int. J. Mol. Sci.* **2019**, *21*, 116. [CrossRef]
9. EC. Corrigendum to Commission Regulation (EU) 2018/605 of 19 April 2018 Amending Annex II to Regulation (EC) No 1107/2009 by Setting Out Scientific Criteria for the Determination of Endocrine Disrupting Properties. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R0605&rid=210 (accessed on 25 August 2020).

10. Ciubotaru, R.M.; Oyedele, J.; Zancanaro, G. Annual assessment of Echinococcus multilocularis surveillance reports submitted in 2018 in the context of Commission Regulation (EU) No 1152/2011. *EFSA J.* **2018**, *16*, 33–36. [CrossRef]
11. OECD. *Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption, No. 150, Update v3, Series on Testing and Assessment, ENV/JM/MONO(2012)22*; OECD: Paris, France, 2017; 988p.
12. OECD. Conceptual Framework for Testing and Assessment of Endocrine Disrupters. Available online: https://www.oecd.org/env/ehs/testing/oecdworkrelatedtoendocrinedisrupters.htm (accessed on 25 August 2020).
13. Safe, S.; Khan, S.; Wu, F.; Li, X. Chemical-induced estrogenicity in Xenoestrogens and phytoestrogens as SERMs and implications for risk assessment (Chapter 65). In *Veterinary Toxicology Basic and Clinical Principles*; Elsevier: Amsterdam, The Netherlands, 2007; pp. 811–822.
14. Cotterill, J.; Palazzolo, L.; Ridgway, C.; Price, N.; Rorije, E.; Moretto, A.; Peijnenburg, A.; Eberini, I. Predicting estrogen receptor binding of chemicals using a suite of in silico methods—Complementary approaches of (Q)SAR, molecular docking and molecular dynamics. *Toxicol. Appl. Pharmacol.* **2019**, *378*, 114630. [CrossRef]
15. Kerdivel, G.; Habauzit, D.; Pakdel, F. Assessment and Molecular Actions of Endocrine-Disrupting Chemicals That Interfere with Estrogen Receptor Pathways. *Int. J. Endocrinol.* **2013**, *2013*, 501851. [CrossRef]
16. Park, J.; Czapla, L.; Amaro, R.E. Molecular Simulations of Aromatase Reveal New Insights into the Mechanism of Ligand Binding. *J. Chem. Inf. Model.* **2013**, *53*, 2047–2056. [CrossRef]
17. Lokhande, K.; Mehere, A.S.; Yadav, A.K.; Swamy, K.V. Molecular Modeling and Docking Studies of Aromatase Inhibitors with Aromatase for ERP Breast Cancer. In Proceedings of the 86th Conference of Society of Biological Chemists: Emerging Discoveries in Health and Agricultural Sciences, New Delhi, India, 16–19 November 2017. [CrossRef]
18. Kang, H.; Xiao, X.; Huang, C.; Yuan, Y.; Tang, D.; Dai, X.; Zeng, X. Potent aromatase inhibitors and molecular mechanism of inhibitory action. *Eur. J. Med. Chem.* **2018**, *143*, 426–437. [CrossRef] [PubMed]
19. Awasthi, M.; Singh, S.; Pandey, V.P.; Dwivedi, U.N. Molecular docking and 3D-QSAR-based virtual screening of flavonoids as potential aromatase inhibitors against estrogen-dependent breast cancer. *J. Biomol. Struct. Dyn.* **2014**, *33*, 804–819. [CrossRef] [PubMed]
20. Suvannang, N.; Nantasenamat, C.; Isarankura-Na-Ayudhya, C.; Prachayasittikul, V. Molecular Docking of Aromatase Inhibitors. *Molecules* **2011**, *16*, 3597–3617. [CrossRef]
21. Rampogu, S.; Son, M.; Park, C.; Kim, H.-H.; Suh, J.-K.; Lee, K. Sulfonanilide Derivatives in Identifying Novel Aromatase Inhibitors by Applying Docking, Virtual Screening, and MD Simulations Studies. *BioMed Res. Int.* **2017**, *2017*, 1–17. [CrossRef] [PubMed]
22. Narayana, B.L.; Kishore, D.P.; Balakumar, C.; Rao, K.V.; Kaur, R.; Rao, A.R.; Murthy, J.N.; Ravikumar, M. Molecular Modeling Evaluation of Non-Steroidal Aromatase Inhibitors†. *Chem. Boil. Drug Des.* **2012**, *79*, 674–682. [CrossRef] [PubMed]
23. Fischer, J.; Ganellin, C.R. *Analogue-Based Drug Discovery*; John Wiley & Sons: Hoboken, NJ, USA, 2006; p. 516.
24. Maltais, R.; Ayan, D.; Trottier, A.; Barbeau, X.; Lagüe, P.; Bouchard, J.-E.; Poirier, D. Discovery of a Non-Estrogenic Irreversible Inhibitor of 17β-Hydroxysteroid Dehydrogenase Type 1 from 3-Substituted-16β-(m-carbamoylbenzyl)-estradiol Derivatives. *J. Med. Chem.* **2013**, *57*, 204–222. [CrossRef]
25. Lespérance, M.; Barbeau, X.; Roy, J.; Maltais, R.; Lagüe, P.; Poirier, D. Chemical synthesis of C3-oxiranyl/oxiranylmethyl-estrane derivatives targeted by molecular modeling and tested as potential inhibitors of 17β-hydroxysteroid dehydrogenase type 1. *Steroids* **2018**, *140*, 104–113. [CrossRef]
26. Maltais, R.; Djiemeny, A.N.; Roy, J.; Barbeau, X.; Lambert, J.-P.; Poirier, D. Design and synthesis of dansyl-labeled inhibitors of steroid sulfatase for optical imaging. *Bioorg. Med. Chem.* **2020**, *28*, 115368. [CrossRef]
27. Rakers, C.; Schumacher, F.; Meinl, W.; Glatt, H.; Kleuser, B.; Wolber, G. In SilicoPrediction of Human Sulfotransferase 1E1 Activity Guided by Pharmacophores from Molecular Dynamics Simulations. *J. Boil. Chem.* **2015**, *291*, 58–71. [CrossRef]
28. Kisselev, P.; Schunck, W.-H.; Roots, I.; Schwarz, D. Association of CYP1A1 Polymorphisms with Differential Metabolic Activation of 17 -Estradiol and Estrone. *Cancer Res.* **2005**, *65*, 2972–2978. [CrossRef]
29. Hong, C.-C.; Tang, B.-K.; Hammond, G.L.; Tritchler, D.; Yaffe, M.; Boyd, N.F. Cytochrome P450 1A2 (CYP1A2) activity and risk factors for breast cancer: A cross-sectional study. *Breast Cancer Res.* **2004**, *6*, R352–R365. [CrossRef] [PubMed]

30. Nishida, C.R.; Everett, S.; De Montellano, P.R.O. Specificity Determinants of CYP1B1 Estradiol Hydroxylation. *Mol. Pharmacol.* **2013**, *84*, 451–458. [CrossRef] [PubMed]
31. Selby, C. Sex Hormone Binding Globulin: Origin, Function and Clinical Significance. *Ann. Clin. Biochem. Int. J. Lab. Med.* **1990**, *27*, 532–541. [CrossRef]
32. Dechering, K.; Boersma, C.; Mosselman, S. Estrogen receptors alpha and beta: Two receptors of a kind? *Curr. Med. Chem.* **2000**, *7*, 561–576. [CrossRef] [PubMed]
33. Paterni, I.; Granchi, C.; Katzenellenbogen, J.A.; Minutolo, F. Estrogen receptors alpha (ERα) and beta (ERβ): Subtype-selective ligands and clinical potential. *Steroids* **2014**, *90*, 13–29. [CrossRef]
34. Ruff, M.; Gangloff, M.; Wurtz, J.; Moras, D. Estrogen receptor transcription and transactivation Structure-function relationship in DNA- and ligand-binding domains of estrogen receptors. *Breast Cancer Res.* **2000**, *2*, 353–359. [CrossRef]
35. Kumar, R.; Zakharov, M.N.; Khan, S.H.; Miki, R.; Jang, H.; Toraldo, G.; Singh, R.; Bhasin, S.; Jasuja, R. The Dynamic Structure of the Estrogen Receptor. *J. Amino Acids* **2011**, *2011*, 1–7. [CrossRef]
36. Billon-Galés, A.; Krust, A.; Fontaine, C.; Abot, A.; Flouriot, G.; Toutain, C.; Berges, H.; Gadeau, A.-P.; Lenfant, F.; Gourdy, P.; et al. Activation function 2 (AF2) of estrogen receptor-α is required for the atheroprotective action of estradiol but not to accelerate endothelial healing. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 13311–13316. [CrossRef]
37. Gao, L.; Tu, Y.; Ågren, H.; Eriksson, L.A. Characterization of Agonist Binding to His524 in the Estrogen Receptor α Ligand Binding Domain. *J. Phys. Chem. B* **2012**, *116*, 4823–4830. [CrossRef]
38. Hu, G.; Wang, J. Ligand selectivity of estrogen receptors by a molecular dynamics study. *Eur. J. Med. Chem.* **2014**, *74*, 726–735. [CrossRef]
39. Gu, X. Helix 12 in the human estrogen receptor (hER) is essential for the hER function by overcoming nucleosome repression in yeast. *J. Cell. Biochem.* **2002**, *86*, 224–238. [CrossRef] [PubMed]
40. Shiau, A.K.; Barstad, D.; Loria, P.M.; Cheng, L.; Kushner, P.J.; Agard, D.A.; Greene, G.L. The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen. *Cell* **1998**, *95*, 927–937. [CrossRef]
41. Fuentes, N.; Silveyra, P. Estrogen receptor signaling mechanisms. *Adv. Protein Chem. Struct. Biol.* **2019**, *116*, 135–170. [CrossRef] [PubMed]
42. Stefkovich, M.L.; Arao, Y.; Hamilton, K.J.; Korach, K.S. Experimental models for evaluating non-genomic estrogen signaling. *Steroids* **2018**, *133*, 34–37. [CrossRef] [PubMed]
43. Vrtacnik, P.; Ostanek, B.; Mencej-Bedrac, S.; Marc, J. The many faces of estrogen signaling. *Biochem. Medica* **2014**, *24*, 329–342. [CrossRef]
44. Rosano, C.; Ponassi, M.; Santolla, M.F.; Pisano, A.; Felli, L.; Vivacqua, A.; Maggiolini, M.; Lappano, R. Macromolecular Modelling and Docking Simulations for the Discovery of Selective GPER Ligands. *AAPS J.* **2015**, *18*, 41–46. [CrossRef]
45. Bruno, A.; Aiello, F.; Costantino, G.; Radi, M. Homology Modeling, Validation and Dynamics of the G Protein-coupled Estrogen Receptor 1 (GPER-1). *Mol. Inform.* **2016**, *35*, 333–339. [CrossRef]
46. Méndez-Luna, D.; Martínez-Archundia, M.; Maroun, R.C.; Ceballos, G.; Fragoso-Vázquez, M.; González-Juárez, D.; Correa-Basurto, J. Deciphering the GPER/GPR30-agonist and antagonists interactions using molecular modeling studies, molecular dynamics, and docking simulations. *J. Biomol. Struct. Dyn.* **2015**, *33*, 1–12. [CrossRef]
47. Taylor, R.; Jewsbury, P.; Essex, J.W. A review of protein-small molecule docking methods. *J. Comput. Mol. Des.* **2002**, *16*, 151–166. [CrossRef]
48. RCSB PDB. Available online: https://www.rcsb.org/ (accessed on 25 August 2020).
49. Landeros-Martinez, L.-L.; Glossman-Mitnik, D.; Orrantia-Borunda, E.; Flores-Holguin, N. A Combined Molecular Docking and Electronic Structure Study for a Breast Cancer Drug Design. In *Molecular Docking*; IntechOpen: London, UK, 2018. [CrossRef]
50. Muchtaridi, M.; Megantara, S.; Dermawan, D.; Yusuf, M. Antagonistic mechanism of α-mangostin derivatives against human estrogen receptor α of breast cancer using molecular dynamics simulation. *Rasayan J. Chem.* **2019**, *12*, 1927–1934. [CrossRef]
51. Cavasotto, C.; Aucar, M.G. High-Throughput Docking Using Quantum Mechanical Scoring. *Front. Chem.* **2020**, *8*, 246. [CrossRef] [PubMed]

52. Spiriti, J.; Subramanian, S.R.; Palli, R.; Wu, M.; Zuckerman, D.M. Middle-way flexible docking: Pose prediction using mixed-resolution Monte Carlo in estrogen receptor α. *PLoS ONE* **2019**, *14*, e0215694. [CrossRef] [PubMed]
53. Wolohan, P.; Reichert, D.E. CoMSIA and docking study of rhenium based estrogen receptor ligand analogs. *Steroids* **2007**, *72*, 247–260. [CrossRef] [PubMed]
54. Wierbowski, S.D.; Wingert, B.M.; Zheng, J.; Camacho, C.J. Cross-docking benchmark for automated pose and ranking prediction of ligand binding. *Protein Sci.* **2019**, *29*, 298–305. [CrossRef]
55. Shtaiwi, A.; Adnan, R.; Khairuddean, M.; Khan, S.U. Computational investigations of the binding mechanism of novel benzophenone imine inhibitors for the treatment of breast cancer. *RSC Adv.* **2019**, *9*, 35401–35416. [CrossRef]
56. Ling, J.; Oh, D.L.; Chia, A.Y.Y. Molecular Docking Studies of Glycyrrhizic Acid (GA), Glycyrrhetic Acid (GE) and Glabridin (GLA) with Estrogen Receptors (ERs). *Biosci. Biotechnol. Res. Asia* **2017**, *14*, 1–11. [CrossRef]
57. Pang, X.; Fu, W.; Wang, J.; Kang, D.; Xu, L.; Zhao, Y.; Liu, A.; Du, G.-H. Identification of Estrogen Receptor α Antagonists from Natural Products via In Vitro and In Silico Approaches. *Oxidative Med. Cell. Longev.* **2018**, *2018*, 6040149. [CrossRef]
58. Muchtaridi, M.; Dermawan, D.; Yusuf, M. Molecular Docking, 3D Structure-Based Pharmacophore Modeling, and ADME Prediction of Alpha Mangostin and Its Derivatives against Estrogen Receptor Alpha. *J. Young Pharm.* **2018**, *10*, 252–259. [CrossRef]
59. Yu, E.; Xu, Y.; Shi, Y.; Yu, Q.; Liu, J.; Xu, L. Discovery of novel natural compound inhibitors targeting estrogen receptor α by an integrated virtual screening strategy. *J. Mol. Model.* **2019**, *25*, 278. [CrossRef]
60. Abdelsamie, A.S.; Salah, M.; Siebenbürger, L.; Hamed, M.M.; Börger, C.; Van Koppen, C.J.; Frotscher, M.; Hartmann, R.W. Development of potential preclinical candidates with promising in vitro ADME profile for the inhibition of type 1 and type 2 17β-Hydroxysteroid dehydrogenases: Design, synthesis, and biological evaluation. *Eur. J. Med. Chem.* **2019**, *178*, 93–107. [CrossRef]
61. Hasan, T.N.; B, L.G.; A Masoodi, T.; Shafi, G.; Alshatwi, A.A.; Sivashanmugham, P. Affinity of estrogens for human progesterone receptor A and B monomers and risk of breast cancer: A comparative molecular modeling study. *Adv. Appl. Bioinform. Chem.* **2011**, *4*, 29–36. [CrossRef] [PubMed]
62. Fu, X.; Wang, P.; Zhu, B.T. Characterization of the Estradiol-Binding Site Structure of Human Protein Disulfide Isomerase (PDI). *PLoS ONE* **2011**, *6*, e27185. [CrossRef] [PubMed]
63. Da Silva, A.J.; Dos Santos, E.S. Aqueous solution interactions with sex hormone-binding globulin and estradiol: A theoretical investigation. *J. Boil. Phys.* **2018**, *44*, 539–556. [CrossRef] [PubMed]
64. Thomas, M.P.; Potter, B.V. The structural biology of oestrogen metabolism. *J. Steroid Biochem. Mol. Boil.* **2013**, *137*, 27–49. [CrossRef]
65. Maltais, R.; Fournier, D.; Poirier, D.; Maltais, R. Quantitative Structure-Activity Relationship (QSAR) Study with a Series of 17Î±-Derivatives of Estradiol: Model for the Development of Reversible Steroid Sulfatase Inhibitors. *QSAR Comb. Sci.* **2009**, *28*, 1284–1299. [CrossRef]
66. Granados, S.T.; Castillo, K.; Bravo-Moraga, F.; Sepúlveda, R.V.; Carrasquel-Ursulaez, W.; Rojas, M.; Carmona, E.; Lorenzo-Ceballos, Y.; González-Nilo, F.D.; González, C.; et al. The molecular nature of the 17β-Estradiol binding site in the voltage- and Ca2+-activated K+ (BK) channel β1 subunit. *Sci. Rep.* **2019**, *9*, 9965. [CrossRef]
67. McCullough, C.; Neumann, T.S.; Gone, J.R.; He, Z.; Herrild, C.; (Nee Lukesh) Wondergem, J.; Pandey, R.K.; Donaldson, W.A.; Sem, D.S. Probing the human estrogen receptor-α binding requirements for phenolic mono- and di-hydroxyl compounds: A combined synthesis, binding and docking study. *Bioorg. Med. Chem.* **2013**, *22*, 303–310. [CrossRef]
68. Kerdivel, G.; Le Guével, R.; Habauzit, D.; Brion, F.; Ait-Aissa, S.; Pakdel, F. Estrogenic Potency of Benzophenone UV Filters in Breast Cancer Cells: Proliferative and Transcriptional Activity Substantiated by Docking Analysis. *PLoS ONE* **2013**, *8*, e60567. [CrossRef]
69. Sauvée, C.; Schäfer, A.; Sundén, H.; Ma, J.-N.; Gustavsson, A.-L.; Burstein, E.S.; Olsson, R. The A-CD analogue of 16β,17α-estriol is a potent and highly selective estrogen receptor β agonist. *MedChemComm* **2013**, *4*, 1439. [CrossRef]
70. Kucinska, M.; Giron, M.D.; Piotrowska, H.; Lisiak, N.; Granig, W.H.; Lopez-Jaramillo, F.-J.; Salto, R.; Murias, M.; Erker, T. Novel Promising Estrogenic Receptor Modulators: Cytotoxic and Estrogenic Activity of Benzanilides and Dithiobenzanilides. *PLoS ONE* **2016**, *11*, e0145615. [CrossRef]

71. Gonzalez, T.L.; Rae, J.M.; Colacino, J.A.; Richardson, R.J. Homology models of mouse and rat estrogen receptor-α ligand-binding domain created by in silico mutagenesis of a human template: Molecular docking with 17β-estradiol, diethylstilbestrol, and paraben analogs. *Comput. Toxicol.* **2019**, *10*, 1–16. [CrossRef] [PubMed]
72. Yarger, J.G.; E Babine, R.; Bittner, M.; Shanle, E.; Xu, W.; Hershberger, P.; Nye, S.H. Structurally similar estradiol analogs uniquely alter the regulation of intracellular signaling pathways. *J. Mol. Endocrinol.* **2012**, *50*, 43–57. [CrossRef] [PubMed]
73. Maltais, R.; Trottier, A.; Barbeau, X.; Lagüe, P.; Perreault, M.; Thériault, J.-F.; Lin, S.X.; Poirier, D. Impact of structural modifications at positions 13, 16 and 17 of 16β-(m-carbamoylbenzyl)-estradiol on 17β-hydroxysteroid dehydrogenase type 1 inhibition and estrogenic activity. *J. Steroid Biochem. Mol. Boil.* **2016**, *161*, 24–35. [CrossRef] [PubMed]
74. E Starkey, N.J.; Li, Y.; Drenkhahn-Weinaug, S.K.; Liu, J.; Lubahn, D.B. 27-Hydroxycholesterol Is an Estrogen Receptor β-Selective Negative Allosteric Modifier of 17β-Estradiol Binding. *Endocrinology* **2018**, *159*, 1972–1981. [CrossRef] [PubMed]
75. Singh, K.; Munuganti, R.S.N.; Leblanc, E.; Lin, Y.L.; Leung, E.; Lallous, N.; Butler, M.S.; Cherkasov, A.; Rennie, P.S. In silico discovery and validation of potent small-molecule inhibitors targeting the activation function 2 site of human oestrogen receptor α. *Breast Cancer Res.* **2015**, *17*, 27. [CrossRef]
76. Durrant, J.D.; Carlson, K.E.; Martin, T.A.; Offutt, T.L.; Mayne, C.G.; Katzenellenbogen, J.A.; Amaro, R.E. Neural-Network Scoring Functions Identify Structurally Novel Estrogen-Receptor Ligands. *J. Chem. Inf. Model.* **2015**, *55*, 1953–1961. [CrossRef]
77. Ng, H.W.; Zhang, W.; Shu, M.; Luo, H.; Ge, W.; Perkins, R.; Tong, W.; Hong, H. Competitive molecular docking approach for predicting estrogen receptor subtype α agonists and antagonists. *BMC Bioinform.* **2014**, *15*, S4. [CrossRef]
78. Amr, A.E.-G.E.; Elsayed, E.; Al-Omar, M.; Eldin, H.O.B.; Nossier, E.S.; Abdalla, M.M. Design, Synthesis, Anticancer Evaluation and Molecular Modeling of Novel Estrogen Derivatives. *Molecules* **2019**, *24*, 416. [CrossRef]
79. Stjernschantz, E.; Reinen, J.; Meinl, W.; George, B.J.; Glatt, H.; Vermeulen, N.P.; Oostenbrink, C. Comparison of murine and human estrogen sulfotransferase inhibition in vitro and in silico—Implications for differences in activity, subunit dimerization and substrate inhibition. *Mol. Cell. Endocrinol.* **2010**, *317*, 127–140. [CrossRef]
80. Hanson, R.N.; McCaskill, E.; Tongcharoensirikul, P.; Dilis, R.; Labaree, D.; Hochberg, R.B. Synthesis and evaluation of 17α-(dimethylphenyl)vinyl estradiols as probes of the estrogen receptor-α ligand binding domain. *Steroids* **2012**, *77*, 471–476. [CrossRef]
81. Grande, F.; Rizzuti, B.; Occhiuzzi, M.A.; Ioele, G.; Casacchia, T.; Gelmini, F.; Guzzi, R.; Garofalo, A.; Statti, G. Identification by Molecular Docking ofHomoisoflavones from Leopoldia comosa as Ligands of Estrogen Receptors. *Molecules* **2018**, *23*, 894. [CrossRef] [PubMed]
82. Powers, C.N.; Setzer, W.N. A molecular docking study of phytochemical estrogen mimics from dietary herbal supplements. *Silico Pharmacol.* **2015**, *3*, 4. [CrossRef] [PubMed]
83. Yuan, P.; Liang, K.; Ma, B.; Zheng, N.; Nussinov, R.; Huang, J. Multiple-Targeting and Conformational Selection in the Estrogen Receptor: Computation and Experiment. *Chem. Boil. Drug Des.* **2011**, *78*, 137–149. [CrossRef] [PubMed]
84. Chen, Y.; Wang, J.; Hong, D.-Y.; Chen, L.; Zhang, Y.-Y.; Xu, Y.-N.; Pan, D.; Fu, L.-Y.; Tao, L.; Luo, H.; et al. Baicalein has protective effects on the 17β-estradiol-induced transformation of breast epithelial cells. *Oncotarget* **2017**, *8*, 10470–10484. [CrossRef] [PubMed]
85. Fokialakis, N.; Alexi, X.; Aligiannis, N.; Boulaka, A.; Meligova, A.K.; Lambrinidis, G.; Kalpoutzakis, E.; Pratsinis, H.; Cheilari, A.; Mitsiou, D.J.; et al. Biological evaluation of isoflavonoids from Genista halacsyi using estrogen-target cells: Activities of glucosides compared to aglycones. *PLoS ONE* **2019**, *14*, e0210247. [CrossRef]
86. Ayoub, N.M.; Siddique, A.B.; Ebrahim, H.Y.; Mohyeldin, M.M.; El Sayed, K.A. The olive oil phenolic (-)-oleocanthal modulates estrogen receptor expression in luminal breast cancer in vitro and in vivo and synergizes with tamoxifen treatment. *Eur. J. Pharmacol.* **2017**, *810*, 100–111. [CrossRef]
87. Yuseran, H.; Hartoyo, E.; Nurseta, T.; Kalim, H. Molecular docking of genistein on estrogen receptors, promoter region of BCLX, caspase-3, Ki-67, cyclin D1, and telomere activity. *J. Taibah Univ. Med. Sci.* **2018**, *14*, 79–87. [CrossRef]

88. Nanashima, N.; Horie, K.; Maeda, H. Phytoestrogenic Activity of Blackcurrant Anthocyanins Is Partially Mediated through Estrogen Receptor Beta. *Molecules* **2017**, *23*, 74. [CrossRef]
89. Alam, S.; Khan, F. Virtual screening, Docking, ADMET and System Pharmacology studies on Garcinia caged Xanthone derivatives for Anticancer activity. *Sci. Rep.* **2018**, *8*, 5524. [CrossRef]
90. Puranik, N.V.; Srivastava, P.; Bhatt, G.; Mary, D.J.S.J.; Limaye, A.M.; Sivaraman, J. Determination and analysis of agonist and antagonist potential of naturally occurring flavonoids for estrogen receptor (ERα) by various parameters and molecular modelling approach. *Sci. Rep.* **2019**, *9*, 7450. [CrossRef]
91. Wang, T.; Wang, Y.; Zhuang, X.; Luan, F.; Zhao, C.; Cordeiro, M.N.D.S. Interaction of Coumarin Phytoestrogens with ERα and ERβ: A Molecular Dynamics Simulation Study. *Molecules* **2020**, *25*, 1165. [CrossRef]
92. Wang, X.; Wang, G.-C.; Rong, J.; Wang, S.W.; Ng, T.B.; Zhang, Y.B.; Lee, K.F.; Zheng, L.; Wong, H.-K.; Yung, K.K.-L.; et al. Identification of Steroidogenic Components Derived From Gardenia jasminoides Ellis Potentially Useful for Treating Postmenopausal Syndrome. *Front. Pharmacol.* **2018**, *9*, 9. [CrossRef]
93. Jeong, J.; Kim, H.; Choi, J. In Silico Molecular Docking and In Vivo Validation with *Caenorhabditis elegans* to Discover Molecular Initiating Events in Adverse Outcome Pathway Framework: Case Study on Endocrine-Disrupting Chemicals with Estrogen and Androgen Receptors. *Int. J. Mol. Sci.* **2019**, *20*, 1209. [CrossRef]
94. Neves, B.J.; Braga, R.C.; Melo-Filho, C.C.; Moreira-Filho, J.T.; Muratov, E.N.; Andrade, C.H. QSAR-Based Virtual Screening: Advances and Applications in Drug Discovery. *Front. Pharmacol.* **2018**, *9*, 1275. [CrossRef]
95. Dror, O.; Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H.J. Novel Approach for Efficient Pharmacophore-Based Virtual Screening: Method and Applications. *J. Chem. Inf. Model.* **2009**, *49*, 2333–2343. [CrossRef]
96. Fiser, A. Template-based protein structure modeling. *Breast Cancer* **2010**, *673*, 73–94. [CrossRef]
97. Zhao, X.; Chen, M.; Huang, B.; Ji, H.; Yuan, M. Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA) Studies on α1A-Adrenergic Receptor Antagonists Based on Pharmacophore Molecular Alignment. *Int. J. Mol. Sci.* **2011**, *12*, 7022–7037. [CrossRef]
98. Sharma, R.; Dhingra, N.; Patil, S. CoMFA, CoMSIA, HQSAR and Molecular Docking Analysis of Ionone-based Chalcone Derivatives as Antiprostate Cancer Activity. *Indian J. Pharm. Sci.* **2016**, *78*, 54–64. [CrossRef]
99. Vuorinen, A.; Engeli, R.T.; Leugger, S.; Kreutz, C.R.; Schuster, D.; Odermatt, A.; Matuszczak, B. Phenylbenzenesulfonates and -sulfonamides as 17β-hydroxysteroid dehydrogenase type 2 inhibitors: Synthesis and SAR-analysis. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 2982–2985. [CrossRef]
100. Vuorinen, A.; Engeli, R.; Meyer, A.; Bachmann, F.; Griesser, U.J.; Schuster, D.; Odermatt, A. Ligand-Based Pharmacophore Modeling and Virtual Screening for the Discovery of Novel 17β-Hydroxysteroid Dehydrogenase 2 Inhibitors. *J. Med. Chem.* **2014**, *57*, 5995–6007. [CrossRef]
101. Chang, Y.-H.; Chen, J.-Y.; Hor, C.-Y.; Chuang, Y.-C.; Yang, C.-B.; Yang, C.-N. Computational Study of Estrogen Receptor-Alpha Antagonist with Three-Dimensional Quantitative Structure-Activity Relationship, Support Vector Regression, and Linear Regression Methods. *Int. J. Med. Chem.* **2013**, *2013*, 1–13. [CrossRef]
102. Sodero, A.C.R.; Romeiro, N.C.; Da Cunha, E.F.F.; Magalhães, U.D.O.; De Alencastro, R.B.; Rodrigues, C.R.; Cabral, L.M.; Castro, H.C.; Albuquerque, M.G. Application of 4D-QSAR Studies to a Series of Raloxifene Analogs and Design of Potential Selective Estrogen Receptor Modulators. *Molecules* **2012**, *17*, 7415–7439. [CrossRef]
103. Wang, P.; McInnes, C.; Zhu, B.T. Structural Characterization of the Binding Interactions of Various Endogenous Estrogen Metabolites with Human Estrogen Receptor α and β Subtypes: A Molecular Modeling Study. *PLoS ONE* **2013**, *8*, e74615. [CrossRef]
104. Bhhatarai, B.; Wilson, D.M.; Price, P.S.; Marty, S.; Parks, A.K.; Carney, E. Evaluation of OASIS QSAR Models Using ToxCast™ in Vitro Estrogen and Androgen Receptor Binding Data and Application in an Integrated Endocrine Screening Approach. *Environ. Health Perspect.* **2016**, *124*, 1453–1461. [CrossRef]
105. Bohari, M.; Srivastava, H.K.; Sastry, G.N. Analogue-based approaches in anti-cancer compound modelling: The relevance of QSAR models. *Org. Med. Chem. Lett.* **2011**, *1*, 3. [CrossRef]
106. Dems, M.A.E.; Laib, S.; Latelli, N.; Ouddai, N. A DFT-based Quantitative structure activity relationship Study of organometallic estradiol derivatives. *JCPS* **2017**, *10*, 483–487.
107. Zhang, T.; Wei, D.-Q.; Chou, K.-C. A pharmacophore model specific to active site of CYP1A2 with a novel molecular modeling explorer and CoMFA. *Med. Chem.* **2012**, *8*, 198–207.

108. Poirier, D.; Roy, J.; Cortes-Benitez, F.; Dutour, R. Targeting cytochrome P450 (CYP) 1B1 with steroid derivatives. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 5272–5276. [CrossRef]
109. Kar, S.; Roy, K.; Leszczynski, J. Impact of Pharmaceuticals on the Environment: Risk Assessment Using QSAR Modeling Approach. *Methods Mol. Biol.* **2018**, *1800*, 395–443. [CrossRef]
110. Colosi, L.M.; Huang, Q.; Weber, W.J. QSAR-assisted design of an environmental catalyst for enhanced estrogen remediation. *Chemosphere* **2010**, *81*, 897–903. [CrossRef]
111. Rokhina, E.V.; Suri, R. Application of density functional theory (DFT) to study the properties and degradation of natural estrogen hormones with chemical oxidizers. *Sci. Total Environ.* **2012**, *417*, 280–290. [CrossRef] [PubMed]
112. Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Boil.* **1976**, *103*, 227–249. [CrossRef]
113. Jorgensen, W. The Many Roles of Computation in Drug Discovery. *Science* **2004**, *303*, 1813–1818. [CrossRef] [PubMed]
114. Simonson, T.; Archontis, G.; Karplus, M. Free Energy Simulations Come of Age: Protein−Ligand Recognition. *Acc. Chem. Res.* **2002**, *35*, 430–437. [CrossRef]
115. Fratev, F.; Steinbrecher, T.; Jónsdóttir, S.O. Prediction of Accurate Binding Modes Using Combination of Classical and Accelerated Molecular Dynamics and Free-Energy Perturbation Calculations: An Application to Toxicity Studies. *ACS Omega* **2018**, *3*, 4357–4371. [CrossRef]
116. Schindler, C.; Rippmann, F.; Kuhn, D. Relative binding affinity prediction of farnesoid X receptor in the D3R Grand Challenge 2 using FEP+. *J. Comput. Mol. Des.* **2017**, *32*, 265–272. [CrossRef]
117. Olsson, M.A.; García-Sosa, A.T.; Ryde, U. Binding affinities of the farnesoid X receptor in the D3R Grand Challenge 2 estimated by free-energy perturbation and docking. *J. Comput. Mol. Des.* **2017**, *32*, 211–224. [CrossRef]
118. Åqvist, J.; Luzhkov, V.B.; Brandsdal, B.O. Ligand Binding Affinities from MD Simulations. *Acc. Chem. Res.* **2002**, *35*, 358–365. [CrossRef]
119. Van Lipzig, M.M.H.; Ter Laak, A.M.; Jongejan, A.; Vermeulen, N.P.; Wamelink, M.; Geerke, D.P.; Meerman, J.H.N. Prediction of Ligand Binding Affinity and Orientation of Xenoestrogens to the Estrogen Receptor by Molecular Dynamics Simulations and the Linear Interaction Energy Method. *J. Med. Chem.* **2004**, *47*, 1018–1030. [CrossRef]
120. Rifai, E.A.; Van Dijk, M.; Vermeulen, N.P.; Geerke, D.P. Binding free energy predictions of FXR agonists using LIE with reliability estimation: Application to the D3R Grand Challenge 2. *J. Comput. Mol. Des.* **2017**, *32*, 239–249. [CrossRef]
121. Simonson, T. Protein: Ligand recognition: Simple models for electrostatic effects. *Curr. Pharm. Des.* **2013**, *19*, 4241–4256. [CrossRef] [PubMed]
122. Gelpí, J.L.; Hospital, A.; Goñi, J.R.; Orozco, M. Molecular dynamics simulations: Advances and applications. *Adv. Appl. Bioinform. Chem.* **2015**, *8*, 37–47. [CrossRef] [PubMed]
123. Freindorf, M.; Furlani, T.R.; Kong, J.; Cody, V.; Davis, F.B.; Davis, P.J. Combined QM/MM Study of Thyroid and Steroid Hormone Analogue Interactions with αvβ3 Integrin. *J. Biomed. Biotechnol.* **2012**, *2012*, 1–12. [CrossRef] [PubMed]
124. Kalaiarasi, C.; Manjula, S.; Kumaradhas, P. Combined quantum mechanics/molecular mechanics (QM/MM) methods to understand the charge density distribution of estrogens in the active site of estrogen receptors. *RSC Adv.* **2019**, *9*, 40758–40771. [CrossRef]
125. Costa, A.H.L.; Clemente, W.S.; Bezerra, K.S.; Neto, J.X.L.; De Albuquerque, É.L.; Fulco, U.; Clemente, W.S., Jr. Computational biochemical investigation of the binding energy interactions between an estrogen receptor and its agonists. *New J. Chem.* **2018**, *42*, 19801–19810. [CrossRef]
126. Hilder, T.A.; Hodgkiss, J.M. Molecular Mechanism of Binding between 17β-Estradiol and DNA. *Comput. Struct. Biotechnol. J.* **2016**, *15*, 91–97. [CrossRef] [PubMed]
127. Eisold, A.; LaBudde, D. Detailed Analysis of 17β-Estradiol-Aptamer Interactions: A Molecular Dynamics Simulation Study. *Molecules* **2018**, *23*, 1690. [CrossRef] [PubMed]
128. Lakhin, A.V.; Tarantul, V.Z.; Gening, L.V. Aptamers: Problems, Solutions and Prospects. *Acta Naturae* **2013**, *5*, 34–43. [CrossRef] [PubMed]
129. Dutkiewicz, Z.; Mikstacka, R. Structure-Based Drug Design for Cytochrome P450 Family 1 Inhibitors. *Bioinorg. Chem. Appl.* **2018**, *2018*, 3924608. [CrossRef] [PubMed]

130. Moldogazieva, N.T.; Ostroverkhova, D.S.; Kuzmich, N.N.; Kadochnikov, V.V.; Terentiev, A.A.; Porozov, Y. Elucidating Binding Sites and Affinities of ERα Agonists and Antagonists to Human Alpha-Fetoprotein by In Silico Modeling and Point Mutagenesis. *Int. J. Mol. Sci.* **2020**, *21*, 893. [CrossRef] [PubMed]
131. Ma, K.; Saha, P.K.; Chan, L.; Moore, D.D. Farnesoid X receptor is essential for normal glucose homeostasis. *J. Clin. Investig.* **2006**, *116*, 1102–1109. [CrossRef] [PubMed]
132. Yang, F.; Huang, X.; Yi, T.; Yen, Y.; Moore, D.D.; Huang, W. Spontaneous Development of Liver Tumors in the Absence of the Bile Acid Receptor Farnesoid X Receptor. *Cancer Res.* **2007**, *67*, 863–867. [CrossRef] [PubMed]
133. Levin, E.R. Cellular Functions of the Plasma Membrane Estrogen Receptor. *Trends Endocrinol. Metab.* **1999**, *10*, 374–377. [CrossRef]
134. Soltysik, K.; Czekaj, P. Membrane estrogen receptors - is it an alternative way of estrogen action? *J. Physiol. Pharmacol. Off. J. Pol. Physiol. Soc.* **2013**, *64*, 129–142.
135. Guo, J.J.; Yang, D.-P.; Tian, X.; Vemuri, V.K.; Yin, D.; Li, C.; Duclos, R.I.; Shen, L.; Ma, X.; Janero, D.R.; et al. 17β-estradiol (E2) in membranes: Orientation and dynamic properties. *Biochim. Biophys. Acta (BBA) Biomembr.* **2016**, *1858*, 344–353. [CrossRef]
136. Vogel, A.; Scheidt, H.A.; Feller, S.E.; Metso, J.; Badeau, R.M.; Tikkanen, M.J.; Wähälä, K.; Jauhiainen, M.; Huster, D. The Orientation and Dynamics of Estradiol and Estradiol Oleate in Lipid Membranes and HDL Disc Models. *Biophys. J.* **2014**, *107*, 114–125. [CrossRef]
137. Goh, J.Y.; Goh, K.S.; Yip, Y.M.; Ng, C.K. High salinity enhances adsorption of 17α-ethinyl estradiol by polyethersulfone membrane: Isotherm modelling and molecular simulation. *engrXiv Preprints* **2019**. [CrossRef]
138. Lakshmanan, S.; Kanwal, A.; Liu, S.; Patlolla, A.; Iqbal, Z.; Mitra, S.; Thomas, G.A.; Fagan, J.A.; Farrow, R.C. Improved Electrophoretic Deposition of Vertical Single Wall Carbon Nanotubes with Nanoscopic Electrostatic Lenses. *Micromachines* **2020**, *11*, 324. [CrossRef]
139. Ulissi, Z.W.; Zhang, J.; Sresht, V.; Blankschtein, D.; Strano, M.S. 2D Equation-of-State Model for Corona Phase Molecular Recognition on Single-Walled Carbon Nanotube and Graphene Surfaces. *Langmuir* **2014**, *31*, 628–636. [CrossRef]
140. Sun, W.; Li, M.; Zhang, W.; Wei, J.; Chen, B.; Wang, C. Sediments inhibit adsorption of 17β-estradiol and 17α-ethinylestradiol to carbon nanotubes and graphene oxide. *Environ. Sci. Nano* **2017**, *4*, 1900–1910. [CrossRef]
141. Jiang, L.; Liu, Y.; Liu, S.; Zeng, G.; Hu, X.-J.; Hu, X.; Guo, Z.; Tan, X.; Wang, L.; Wu, Z. Adsorption of Estrogen Contaminants by Graphene Nanomaterials under Natural Organic Matter Preloading: Comparison to Carbon Nanotube, Biochar, and Activated Carbon. *Environ. Sci. Technol.* **2017**, *51*, 6352–6359. [CrossRef] [PubMed]
142. Boateng, L.K.; Heo, J.; Flora, J.R.V.; Park, Y.-G.; Yoon, Y. Molecular level simulation of the adsorption of bisphenol A and 17α-ethinyl estradiol onto carbon nanomaterials. *Sep. Purif. Technol.* **2013**, *116*, 471–478. [CrossRef]
143. Zaib, Q.; Khan, I.A.; Saleh, N.B.; Flora, J.R.V.; Park, Y.-G.; Yoon, Y. Removal of Bisphenol A and 17β-Estradiol by Single-Walled Carbon Nanotubes in Aqueous Solution: Adsorption and Molecular Modeling. *Water Air Soil Pollut.* **2012**, *223*, 3281–3293. [CrossRef]
144. Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864. [CrossRef]
145. Kohn, W.; Sham, L.J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138. [CrossRef]
146. Grimme, S. Density functional theory with London dispersion corrections. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 211–228. [CrossRef]
147. Clark, S.J.; Segall, M.D.; Pickard, C.J.; Hasnip, P.J.; Probert, M.I.J.; Refson, K.; Payne, M.C. First principles methods using CASTEP. *Z. Krist. Cryst. Mater.* **2005**, *220*, 567–570. [CrossRef]
148. Amusia, M.Y.; Msezane, A.Z.; Shaginyan, V.R. Density Functional Theory versus the Hartree–Fock Method: Comparative Assessment. *Phys. Scr.* **2003**, *68*, C133–C140. [CrossRef]
149. Mota, K.; Neto, J.X.L.; Costa, A.L.; Oliveira, J.; Bezerra, K.; De Albuquerque, É.L.; Caetano, E.W.S.; Freire, V.; Fulco, U. A quantum biochemistry model of the interaction between the estrogen receptor and the two antagonists used in breast cancer treatment. *Comput. Theor. Chem.* **2016**, *1089*, 21–27. [CrossRef]

150. Mazurek, A.H.; Szeleszczuk, Ł.; Pisklak, D.M. Periodic DFT Calculations—Review of Applications in the Pharmaceutical Sciences. *Pharmaceutics* **2020**, *12*, 415. [CrossRef]
151. Stevenson, E.L.; Lancaster, R.W.; Buanz, A.B.M.; Price, L.S.; Tocher, D.A.; Price, S.L. The solid state forms of the sex hormone 17-β-estradiol. *CrystEngComm* **2019**. [CrossRef]
152. Du, R.; Xu, J.; Zhang, L.; Ning, L.; Li, S. Ethinyl estradiol cocrystals assembled by chain structures: Improvement in stability and solubility. *New J. Chem.* **2019**, *43*, 16889–16897. [CrossRef]
153. Wang, J.-R.; Yang, Y.; Chen, X.; Mei, X. Solid-state characterization of 17β-estradiol co-crystals presenting improved dissolution and bioavailability. *CrystEngComm* **2016**, *18*, 3498–3505. [CrossRef]
154. Watanabe, H.; Okiyama, Y.; Nakano, T.; Tanaka, S. Incorporation of solvation effects into the fragment molecular orbital calculations with the Poisson–Boltzmann equation. *Chem. Phys. Lett.* **2010**, *500*, 116–119. [CrossRef]
155. Szeleszczuk, Ł.; Pisklak, D.M.; Zielińska-Pisklak, M.; Wawer, I. Effects of structural differences on the NMR chemical shifts in cinnamic acid derivatives: Comparison of GIAO and GIPAW calculations. *Chem. Phys. Lett.* **2016**, *653*, 35–41. [CrossRef]
156. Charpentier, T. The PAW/GIPAW approach for computing NMR parameters: A new dimension added to NMR study of solids. *Solid State Nucl. Magn. Reson.* **2011**, *40*, 1–20. [CrossRef] [PubMed]
157. Elyashberg, M.E. Identification and structure elucidation by NMR spectroscopy. *TrAC Trends Anal. Chem.* **2015**, *69*, 88–97. [CrossRef]
158. Szeleszczuk, Ł.; Pisklak, D.M.; Zielińska-Pisklak, M.; Jurczak, E. A new polymorph of 17-β-estradiol and the application of different analytical techniques (ssNMR, PXRD, DSC, and FTIR) for its study. *J. Mol. Struct.* **2019**, *1183*, 274–280. [CrossRef]
159. Singh, H.; Singh, S.; Srivastava, A.; Tandon, P.; Bharti, P.; Kumar, S.; Maurya, R. Conformational analysis and vibrational study of daidzein by using FT-IR and FT-Raman spectroscopies and DFT calculations. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2014**, *120*, 405–415. [CrossRef] [PubMed]
160. Machado, N.; De Carvalho, L.A.E.B.; Otero, J.C.; Marques, M.P.M. A conformational study of hydroxyflavones by vibrational spectroscopy coupled to DFT calculations. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2013**, *109*, 116–124. [CrossRef]
161. A Minaeva, V.; Minaev, B.F.; Hovorun, D.M. Vibrational spectra of the steroid hormones, estradiol and estriol, calculated by density functional theory. The role of low-frequency vibrations. *Ukr. Biokhimichnyi Zhurnal* **2009**, *80*, 82–95.
162. Scheidt, H.A.; Badeau, R.M.; Huster, D. Investigating the membrane orientation and transversal distribution of 17β-estradiol in lipid membranes by solid-state NMR. *Chem. Phys. Lipids* **2010**, *163*, 356–361. [CrossRef] [PubMed]
163. Vedad, J.; Mojica, E.-R.E.; Desamero, R.Z. Raman spectroscopic discrimination of estrogens. *Vib. Spectrosc.* **2018**, *96*, 93–100. [CrossRef] [PubMed]
164. Oren, I.; Fleishman, S.J.; Kessel, A.; Tal, N.B. Free Diffusion of Steroid Hormones Across Biomembranes: A Simplex Search with Implicit Solvent Model Calculations. *Biophys. J.* **2004**, *87*, 768–779. [CrossRef] [PubMed]
165. Ellena, J.; De Paula, K.; De Melo, C.C.; Da Silva, C.C.P.; Bezerra, B.P.; Venâncio, T.; Ayala, A.P. Temperature-Driven Isosymmetric Reversible Phase Transition of the Hormone Estradiol 17β Valerate. *Cryst. Growth Des.* **2014**, *14*, 5700–5709. [CrossRef]
166. Morishima, F.; Inokuchi, Y.; Ebata, T. Laser Spectroscopic Study of β-Estradiol and Its Monohydrated Clusters in a Supersonic Jet. *J. Phys. Chem. A* **2012**, *116*, 8201–8208. [CrossRef]
167. Borah, M.M.; Devi, T.G. The vibrational spectroscopic studies and molecular property analysis of Estradiol, Tamoxifen and their interaction by density functional theory. *J. Mol. Struct.* **2018**, *1163*, 205–220. [CrossRef]
168. Cherkasova, O.; Nazarov, M.; Mankova, A.; Fedulova, E.; Volodin, V.; Minaeva, V.A.; Minaev, B.F.; Baryshnikov, G.V. Terahertz time-domain spectroscopy of testosterone, estradiol and estriol. In Proceedings of the 2010 International Kharkov Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves, Kharkiv, Ukraine, 21–26 June 2010; pp. 1–3.
169. Hafizi, R.; Taheri, R.A.; Moghimi, H. Liquid phase extraction of nanosized biologically active estrogenic pollutants by using an efficient adsorbent. *J. Mol. Liq.* **2018**, *266*, 535–539. [CrossRef]
170. Donini, C.A.; Silva, M.K.L.; Simões, R.; Cesarino, I. Reduced graphene oxide modified with silver nanoparticles for the electrochemical detection of estriol. *J. Electroanal. Chem.* **2018**, *809*, 67–73. [CrossRef]

171. Klamt, A.; Schüürmann, G. COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans. 2* **1993**, *2*, 799–805. [CrossRef]
172. Mennucci, B.; Tomasi, J.; Cammi, R.; Cheeseman, J.R.; Frisch, M.J.; Devlin, F.J.; Gabriel, S.; Stephens, P.J. Polarizable Continuum Model (PCM) Calculations of Solvent Effects on Optical Rotations of Chiral Molecules. *J. Phys. Chem. A* **2002**, *106*, 6102–6113. [CrossRef]
173. Chinnasamy, K.; Kumaradhas, P. Intermolecular interactions and charge density distribution of endocrine-disrupting molecules (xenoestrogens) with ERα: QM/MM perspective. *Struct. Chem.* **2020**, *31*, 1013–1028. [CrossRef]
174. Ruiz, P.; Ingale, K.; Wheeler, J.S.; Mumtaz, M. 3D QSAR studies of hydroxylated polychlorinated biphenyls as potential xenoestrogens. *Chemosphere* **2016**, *144*, 2238–2246. [CrossRef] [PubMed]
175. Kim, M.; Li, L.Y.; Grace, J.R. Predictability of physicochemical properties of polychlorinated dibenzo-p-dioxins (PCDDs) based on single-molecular descriptor models. *Environ. Pollut.* **2016**, *213*, 99–111. [CrossRef]
176. Eguchi, A.; Hanazato, M.; Suzuki, N.; Matsuno, Y.; Todaka, E.; Mori, C. Maternal–fetal transfer rates of PCBs, OCPs, PBDEs, and dioxin-like compounds predicted through quantitative structure–activity relationship modeling. *Environ. Sci. Pollut. Res.* **2015**, *25*, 7212–7222. [CrossRef]
177. Delfosse, V.; Grimaldi, M.; Cavaillès, V.; Balaguer, P.; Bourguet, W. Structural and Functional Profiling of Environmental Ligands for Estrogen Receptors. *Environ. Health Perspect.* **2014**, *122*, 1306–1313. [CrossRef]
178. Nwachukwu, J.C.; Srinivasan, S.; Bruno, N.E.; Nowak, J.; Wright, N.J.; Minutolo, F.; Rangarajan, E.S.; Izard, T.; Yao, X.-Q.; Grant, B.J.; et al. Systems Structural Biology Analysis of Ligand Effects on ERα Predicts Cellular Response to Environmental Estrogens and Anti-hormone Therapies. *Cell Chem. Boil.* **2017**, *24*, 35–45. [CrossRef]
179. Cozzini, P.; Dellafiora, L. In silico approach to evaluate molecular interaction between mycotoxins and the estrogen receptors ligand binding domain: A case study on zearalenone and its metabolites. *Toxicol. Lett.* **2012**, *214*, 81–85. [CrossRef]
180. Zhang, J.; Wu, W.; Song, Y.; Hou, L.; Li, T.; Guan, T.; Zhang, T.; Wang, Y. Homogeneous assay for zearalenone analogues and their docking studies with apo-/holo-estrogen receptors. *Anal. Methods* **2019**, *11*, 192–199. [CrossRef]
181. Dellafiora, L.; Oswald, I.P.; Dorne, J.-L.; Galaverna, G.; Battilani, P.; Dall'Asta, C. An in silico structural approach to characterize human and rainbow trout estrogenicity of mycotoxins: Proof of concept study using zearalenone and alternariol. *Food Chem.* **2020**, *312*, 126088. [CrossRef]
182. Yang, J.; Hu, C.T.; Zhu, X.; Zhu, Q.; Ward, M.D.; Kahr, B. DDT Polymorphism and the Lethality of Crystal Forms. *Angew. Chem.* **2017**, *129*, 10299–10303. [CrossRef]
183. Zhang, H.; He, W.; Luo, X.; Lin, X.; Lu, X. Adsorption of 2,3,7,8-tetrochlorodibenzo-p-dioxins on intrinsic, defected, and Ti (N, Ag) doped graphene: A DFT study. *J. Mol. Model.* **2014**, *20*, 1–7. [CrossRef] [PubMed]
184. Liu, C.; Li, H.; Johnston, C.T.; Boyd, S.A.; Teppen, B.J. Relating Clay Structural Factors to Dioxin Adsorption by Smectites: Molecular Dynamics Simulations. *Soil Sci. Soc. Am. J.* **2012**, *76*, 110–120. [CrossRef]
185. Zhang, J.; Wu, W.; Wang, Y.; Xing, X.; Zhong, S.; Guan, T.; Zhang, T.; Hou, L.; Li, T. Estrogen receptor-based fluorescence polarization assay for bisphenol analogues and molecular modeling study of their complexation mechanism. *Anal. Chim. Acta* **2018**, *1032*, 107–113. [CrossRef] [PubMed]
186. Sengupta, S.; Obiorah, I.; Maximov, P.; Curpan, R.; Jordan, V.C. Molecular mechanism of action of bisphenol and bisphenol A mediated by oestrogen receptor alpha in growth and apoptosis of breast cancer cells. *Br. J. Pharmacol.* **2013**, *169*, 167–178. [CrossRef] [PubMed]
187. Liu, Y.; Qu, K.; Hai, Y.; Zhao, C. Bisphenol A (BPA) binding on full-length architectures of estrogen receptor. *J. Cell. Biochem.* **2018**, *119*, 6784–6794. [CrossRef] [PubMed]
188. Zhang, J.; Li, T.; Wang, T.; Yuan, C.; Zhong, S.; Guan, T.; Li, Z.; Wang, Y.; Yu, H.; Luo, Q.; et al. Estrogenicity of halogenated bisphenol A: In vitro and in silico investigations. *Arch. Toxicol.* **2017**, *92*, 1215–1223. [CrossRef] [PubMed]
189. Cao, H.; Wang, F.; Liang, Y.; Wang, H.; Zhang, A.; Song, M. Experimental and computational insights on the recognition mechanism between the estrogen receptor α with bisphenol compounds. *Arch. Toxicol.* **2017**, *91*, 3897–3912. [CrossRef]
190. Zhuang, S.; Zhang, C.; Liu, W. Atomic Insights into Distinct Hormonal Activities of Bisphenol A Analogues toward PPARγ and ERα Receptors. *Chem. Res. Toxicol.* **2014**, *27*, 1769–1779. [CrossRef]

191. Li, L.; Wang, Q.; Zhang, Y.; Niu, Y.; Yao, X.; Liu, H. The Molecular Mechanism of Bisphenol A (BPA) as an Endocrine Disruptor by Interacting with Nuclear Receptors: Insights from Molecular Dynamics (MD) Simulations. *PLoS ONE* **2015**, *10*, e0120330. [CrossRef]
192. Delfosse, V.; Grimaldi, M.; Pons, J.-L.; Boulahtouf, A.; Le Maire, A.; Cavaillès, V.; Labesse, G.; Bourguet, W.; Balaguer, P. Structural and mechanistic insights into bisphenols action provide guidelines for risk assessment and discovery of bisphenol A substitutes. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 14930–14935. [CrossRef] [PubMed]
193. Wei, D.; Li, J.; Chen, Z.; Liang, L.; Ma, J.; Wei, M.; Ai, Y.; Wang, X. Understanding bisphenol-A adsorption in magnetic modified covalent organic frameworks: Experiments coupled with DFT calculations. *J. Mol. Liq.* **2020**, *301*, 112431. [CrossRef]
194. Bao, S.; Wu, S.; Huang, L.; Xu, X.; Xu, R.; Li, Y.; Liang, Y.; Yang, M.; Yoon, D.K.; Lee, M.; et al. Supramolecular Nanopumps with Chiral Recognition for Moving Organic Pollutants from Water. *ACS Appl. Mater. Interfaces* **2019**, *11*, 31220–31226. [CrossRef] [PubMed]
195. Díaz, I.; Díez, E.; Camacho, J.; León, S.; Ovejero, G.; Cabanillas, S.L. Comparison between three predictive methods for the calculation of polymer solubility parameters. *Fluid Phase Equilibria* **2013**, *337*, 6–10. [CrossRef]
196. Wang, C.-Y.; Zhang, Y.-J.; Wang, W.-K.; Pei, D.-N.; Huang, G.-X.; Chen, J.-J.; Zhang, X.; Yu, H.-Q. Enhanced photocatalytic degradation of bisphenol A by Co-doped BiOCl nanosheets under visible light irradiation. *Appl. Catal. B Environ.* **2018**, *221*, 320–328. [CrossRef]
197. Motta, A.; La Mantia, F.P.; Ascione, L.; Mistretta, M. Theoretical study on the decomposition mechanism of bisphenol A polycarbonate induced by the combined effect of humidity and UV irradiation. *J. Mol. Graph. Model.* **2020**, *99*, 107622. [CrossRef]
198. Liu, Y.; Liu, Y.; Liu, Z.; Du, F.; Qin, G.; Li, G.; Hu, X.; Xu, Z.; Cai, Z. Supramolecularly imprinted polymeric solid phase microextraction coatings for synergetic recognition nitrophenols and bisphenol A. *J. Hazard. Mater.* **2019**, *368*, 358–364. [CrossRef]
199. Dvorakova, M.; Kejlová, K.; Rucki, M.; Jírová, D. Selected bisphenols and phthalates screened for estrogen and androgen disruption by in silico and in vitro methods. *Neuro Endocrinol. Lett.* **2018**, *39*, 409–416.
200. Zhu, Q.; Liu, L.; Zhou, X.; Ma, M. In silico study of molecular mechanisms of action: Estrogenic disruptors among phthalate esters. *Environ. Pollut.* **2019**, *255*, 113193. [CrossRef]
201. Josh, M.S.; Pradeep, S.; Adarsh, V.; Amma, K.V.; Devi, R.S.; Balachandran, S.; Sreejith, M.; Jaleel, U.A.; Benjamin, S. In silicoevidences for the binding of phthalates onto human estrogen receptor α, β subtypes and human estrogen-related receptor γ. *Mol. Simul.* **2013**, *40*, 408–417. [CrossRef]
202. Sheikh, I.A.; Turki, R.F.; Abuzenadah, A.M.; Damanhouri, G.A.; A Beg, M. Endocrine Disruption: Computational Perspectives on Human Sex Hormone-Binding Globulin and Phthalate Plasticizers. *PLoS ONE* **2016**, *11*, e0151444. [CrossRef] [PubMed]
203. Wang, S.; Wang, S.; Chen, M.; Xu, D.; Tang, L.; Wang, S. Elucidating Adsorption Mechanisms of Phthalate Esters upon Carbon Nanotubes/Graphene and Natural Organic Acid Competitive Effects in Water by DFT and MD Calculations. *Bull. Korean Chem. Soc.* **2015**, *36*, 1631–1636. [CrossRef]
204. Liu, Y.; Zhang, R.; Wang, X.; Sun, P.; Chen, W.; Shen, J.; Xue, G. Hydrogenation induced deviation of temperature and concentration dependences of polymer-solvent interactions in poly(vinyl chloride) and a new eco-friendly plasticizer. *Eur. Phys. J. Plus* **2015**, *130*, 11. [CrossRef]
205. Jacob, R.B.; Andersen, T.; McDougal, O.M. Accessible High-Throughput Virtual Screening Molecular Docking Software for Students and Educators. *PLoS Comput. Boil.* **2012**, *8*, e1002499. [CrossRef]
206. Chaput, L.; Mouawad, L. Efficient conformational sampling and weak scoring in docking programs? Strategy of the wisdom of crowds. *J. Chem.* **2017**, *9*, 37. [CrossRef]
207. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **2015**, *10*, 449–461. [CrossRef]
208. Software Website. GOLD—Protein Ligand Docking Software. Available online: www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/ (accessed on 25 August 2020).
209. Clark, M.; Cramer, R.D.; Van Opdenbosch, N. Validation of the general purpose tripos 5.2 force field. *J. Comput. Chem.* **1989**, *10*, 982–1012. [CrossRef]
210. Software Website. Available online: https://www.charmm.org/ (accessed on 25 August 2020).

211. Brooks, B.R.; Brooks, C.L.; Mackerell, A.D., Jr.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545–1614. [CrossRef]
212. Software Website. Available online: http://www.swissdock.ch/ (accessed on 25 August 2020).
213. Grosdidier, A.; Zoete, V.; Michielin, O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.* **2011**, *39*, W270–W277. [CrossRef]
214. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; Beer, T.A.P.D.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. [CrossRef]
215. Dominguez, C.; Boelens, R.; Bonvin, A.M.J.J. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **2003**, *125*, 1731–1737. [CrossRef] [PubMed]
216. Jian, Y.; He, Y.; Yang, J.; Han, W.; Zhai, X.; Zhao, Y.; Li, Y. Molecular Modeling Study for the Design of Novel Peroxisome Proliferator-Activated Receptor Gamma Agonists Using 3D-QSAR and Molecular Docking. *Int. J. Mol. Sci.* **2018**, *19*, 630. [CrossRef]
217. Software Website. Available online: https://www.schrodinger.com/maestro (accessed on 25 August 2020).
218. McAliley, J.H.; Bruce, D.A. Development of Force Field Parameters for Molecular Simulation of Polylactide. *J. Chem. Theory Comput.* **2011**, *7*, 3756–3767. [CrossRef] [PubMed]
219. Halgren, T.A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Compt. Chem.* **1996**, *17*, 490–519. [CrossRef]
220. Software Website. Available online: https://gaussian.com/ (accessed on 25 August 2020).
221. Forli, S.; Huey, R.; Pique, M.E.; Sanner, M.F.; Goodsell, D.S.; Olson, A.J. Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* **2016**, *11*, 905–919. [CrossRef]
222. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2009**, *31*, 455–461. [CrossRef]
223. Software Website. Available online: http://autodock.scripps.edu/ (accessed on 25 August 2020).
224. Lipparini, F.; Mennucci, B. Perspective: Polarizable continuum models for quantum-mechanical descriptions. *J. Chem. Phys.* **2016**, *144*, 160901. [CrossRef]
225. Li, D.-D.; Meng, X.-F.; Wang, Q.; Yu, P.; Zhao, L.-G.; Zhang, Z.-P.; Wang, Z.; Xiao, W. Consensus scoring model for the molecular docking study of mTOR kinase inhibitor. *J. Mol. Graph. Model.* **2018**, *79*, 81–87. [CrossRef]
226. Software Website. Available online: https://ambermd.org/ (accessed on 25 August 2020).
227. Case, D.A.; Iii, T.E.C.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688. [CrossRef]
228. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C.L. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins Struct. Funct. Bioinform.* **2006**, *65*, 712–725. [CrossRef]
229. Available online: http://www.cosmo-model.org/ (accessed on 25 August 2020).
230. Grubmüller, H.; Heller, H.; Windemuth, A.; Schulten, K. Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-range Interactions. *Mol. Simul.* **1991**, *6*, 121–142. [CrossRef]
231. Softwae Website. Available online: www.gromacs.org (accessed on 25 August 2020).
232. Softwae Website. Available online: https://www.vasp.at/ (accessed on 25 August 2020).
233. Softwae Website. Available online: https://www.ks.uiuc.edu/Research/namd/ (accessed on 25 August 2020).
234. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [CrossRef] [PubMed]
235. Karamertzanis, P.G.; Pantelides, C.C. Ab initio crystal structure prediction. I. Rigid molecules. *J. Comput. Chem.* **2005**, *26*, 304–324. [CrossRef] [PubMed]
236. Karamertzanis, P.G.; Pantelides, C.C. Ab initio crystal structure prediction. II. Flexible molecules. *Mol. Phys.* **2007**, *105*, 273–291. [CrossRef]
237. Vasileiadis, M.; Pantelides, C.C.; Adjiman, C.S. Prediction of the crystal structures of axitinib, a polymorphic pharmaceutical molecule. *Chem. Eng. Sci.* **2015**, *121*, 60–76. [CrossRef]
238. Delley, B. DFT studies: From molecules and molecular environments to surfaces and solids. *Comput. Mater. Sci.* **2000**, *17*, 122–126. [CrossRef]

239. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868. [CrossRef]
240. Softwae Website. Available online: http://www.castep.org/ (accessed on 25 August 2020).
241. Becke, A.D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652. [CrossRef]
242. Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954–9960. [CrossRef]

molecules

MDPI

Review

Current Status of Quantum Chemical Studies of Cyclodextrin Host–Guest Complexes

Anna Helena Mazurek [1] and Łukasz Szeleszczuk [2,*]

[1] Department of Physical Chemistry, Chair of Physical Pharmacy and Bioanalysis, Faculty of Pharmacy, Doctoral School, Medical University of Warsaw, Banacha 1 Str., 02-093 Warsaw, Poland; anna.mazurek@wum.edu.pl

[2] Department of Physical Chemistry, Chair of Physical Pharmacy and Bioanalysis, Faculty of Pharmacy, Medical University of Warsaw, Banacha 1 Str., 02-093 Warsaw, Poland

* Correspondence: lukasz.szeleszczuk@wum.edu.pl; Tel.: +48-501-255-121

Abstract: This article aims to review the application of various quantum chemical methods (semi-empirical, density functional theory (DFT), second order Møller–Plesset perturbation theory (MP2)) in the studies of cyclodextrin host–guest complexes. The details of applied approaches such as functionals, basis sets, dispersion corrections or solvent treatment methods are analyzed, pointing to the best possible options for such theoretical studies. Apart from reviewing the ways that the computations are usually performed, the reasons for such studies are presented and discussed. The successful applications of theoretical calculations are not limited to the determination of stable conformations but also include the prediction of thermodynamic properties as well as UV–Vis, IR, and NMR spectra. It has been shown that quantum chemical calculations, when applied to the studies of CD complexes, can provide results unobtainable by any other methods, both experimental and computational.

Keywords: cyclodextrin; host–guest complexes; DFT; QC; quantum chemistry; density functional theory; CD complexes

Citation: Mazurek, A.H.; Szeleszczuk, Ł. Current Status of Quantum Chemical Studies of Cyclodextrin Host–Guest Complexes. *Molecules* **2022**, *27*, 3874. https://doi.org/10.3390/molecules27123874

Academic Editors: Rosa Iacovino, Marina Isidori and Margherita Lavorgna

Received: 30 May 2022
Accepted: 13 June 2022
Published: 16 June 2022

1. Introduction

Due to their unique structural, physical, and chemical properties, cyclodextrins (CDs) and their derivatives have been of great interest for more than a century [1]. The biodegradability, biocompatibility, and versatility of CDs and CDs-based materials extend their applications to new areas every year; however, the main property that makes CDs so popular is their ability to form host–guest complexes with a variety of compounds [2].

CDs are commonly used in pharmaceutical formulations as they increase the solubility of poorly soluble drugs and protect substances against external factors, such as light, humidity, and heat [3]. CDs can mask unpleasant smells or flavors of drugs, which is especially important in formulations dedicated to children [4]. More than 100 original drugs are currently being manufactured with CDs as excipients [5–7].

Interactions between CDs (host) and guest molecules may yield a stable complex with a high equilibrium constant; for example, β-CD forms highly stable inclusion complexes with adamantyl derivatives with a binding constant of ~10^4–10^5 M^{-1} [8,9]. It is not surprising then that the number of newly obtained cyclodextrin host–guest complexes is constantly increasing. However, only a small amount of those complexes is being reported with their crystal structures. This is caused by the fact that most of those complexes are either amorphous or polycrystalline, and even for the crystalline complexes it is usually very hard to obtain a crystal of a size suitable for single crystal X-ray measurements [10,11]. This is probably one of the reasons why a lot of experimental works describing the structure and properties of CDs complexes are supported by theoretical calculations.

To fully understand the behavior and physicochemical properties of a complex, knowledge of its structure is essential. However, it is not just the desire to reveal how the complex looks that makes the application of molecular modeling methods so popular in the studies of CDs complexes. By choosing an appropriate computational approach, it is possible to determine (or explain) the molar stoichiometry of the complex, the differences observed in the spectra (UV–VIS, IR, NMR) of host–guest physical mixtures and their complexes, and also to predict the stability of such structures under various conditions such as different solvents, temperature or pressure.

CDs host–guest complexes are surely very flexible structures, which is the common cause of their polycrystallinity. This is why a lot of the molecular modelling studies devoted to those complexes utilize molecular dynamics simulation at the molecular mechanics level. Those works have been recently reviewed by us [12]. However, the types of intermolecular forces that stabilize such complexes such as hydrogen bonding, van der Waals, and hydrophobic and dipole–dipole interactions, usually cannot be modeled with the required accuracy using the molecular mechanics methods. This is why the number of works in which the calculations of CD complexes at the higher level of theory, namely, quantum chemical (QC), has constantly increased since 2005 (Figure 1). Now, after 20 years of studies, the number of such articles is large enough to draw some general conclusions and trends as well as the advantages and disadvantages of such an approach. Therefore, the aim of this review was to gather and analyze the works in which the CD host–guest complexes have been modeled using QC methods.

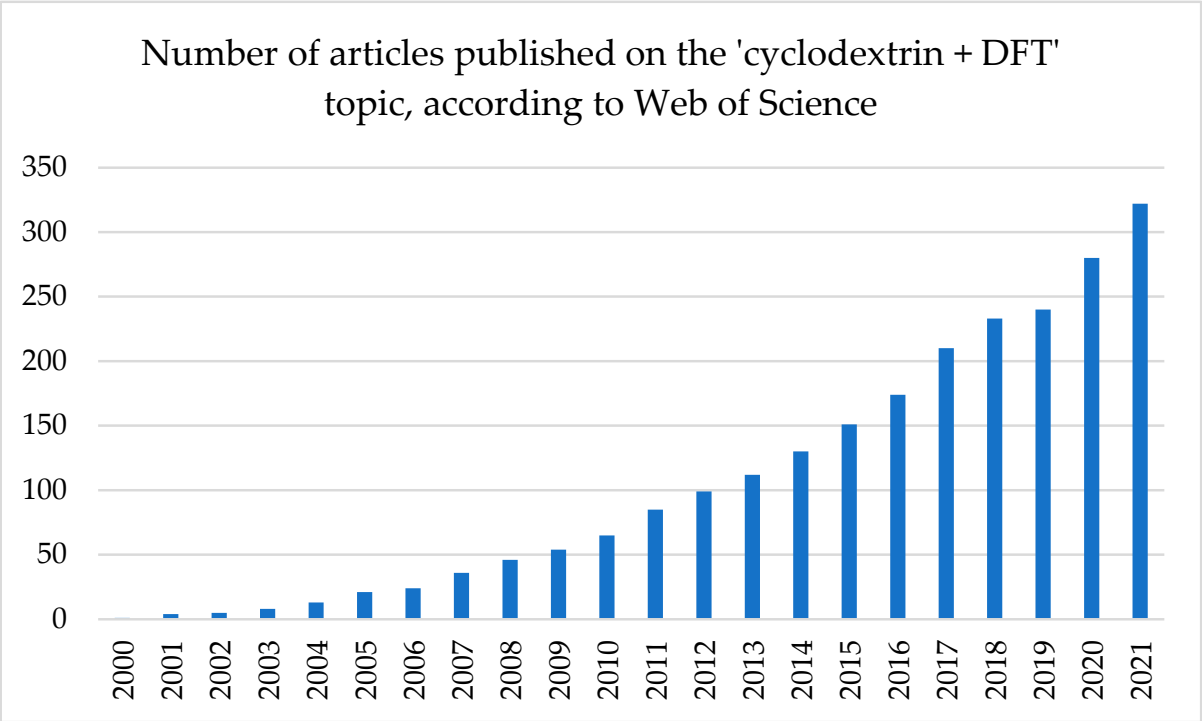


Figure 1. Number of the search results for the 'cyclodextrin AND DFT' phrase in the Web of Science. Each column shows the number of articles in the given year and all years before. For example, the column entitled '2010' depicts the number of articles published in the period 2000–2010, including 2010.

There were at least a few reasons behind writing this review. First, it was worth analyzing whether some conclusions can be made on the choice of the most accurate method, including the applied DFT functional, basis set, dispersion correction or solvent model. Those aspects are discussed at the very beginning. Then, it was interesting to check the main reasons behind the QC method chosen by the authors of the reviewed works. Was

it solely to predict or suggest the possible structure of the complex, or were the calculations used for something more such as the simulation of the spectra or explanation of the reaction mechanism to support the experimental findings? Further, we wanted to find out what type of CDs and what possible guests were studied in those computational works. Therefore, an informative table presenting the most essential information such as the composition of the complex, applied functional, basis set, and solvent treatment has been prepared to serve as an informative and easy-to-follow guide for future studies. Finally, the chosen examples are described in a more detailed way, suggesting the possible solutions and future indications.

As the authors of this review have been using the QC methods to model the structure and properties of CD complexes and found this approach very useful, it was our hope and desire to convince other researchers to try such solutions in their works.

2. Applied Calculational Methods and Parameters

2.1. Choice of QC Method

The computational methods that were used in the reviewed works nicely correspond with the general increase in the computational power available to researchers worldwide. The earliest (before 2005) QC works studying CD complexes were done using the least demanding semi-empirical methods such as AM1, PM3 and later PM6 or PM7. Subsequently, those methods have been gradually superseded by density functional theory (DFT) calculations, while recently a few works have been published in which the Møller–Plesset perturbation theory (MP) was applied. It is worth noticing that CD complexes are not small objects, in terms of QC calculations, especially when the γ-CD or substituted CDs are the hosts with the large ligand as a guest. Application of QC can also be a problem when studying complexes with a host:guest ratio higher than 1:1. This is why even in the 2020s in some works, semiempirical methods have been applied; however, the ratio of DFT to the semi-empirical ones is constantly increasing (Figure 2).

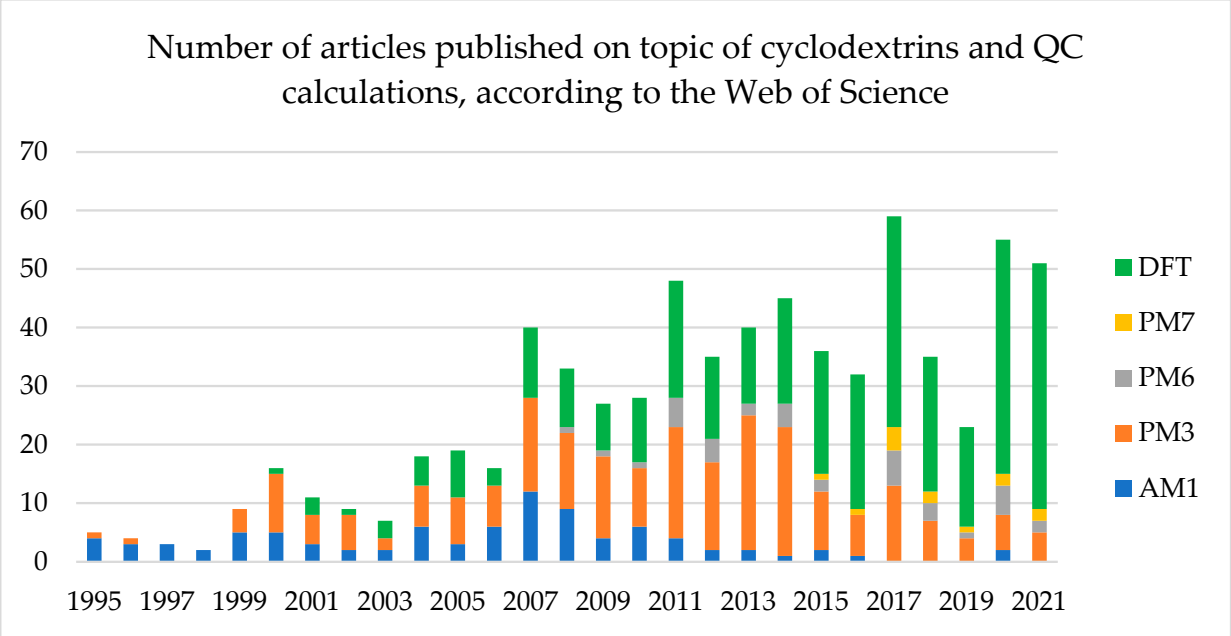


Figure 2. Changes in the number of articles published on the topic of cyclodextrin and either different semi-empirical or DFT methods over the years.

2.2. General Remarks

- Calculations of CD other than in the form of typical complexes

As stated in the title of this manuscript, this review focuses on the application of QC methods in the analysis of CD complexes. However, computational studies are also

performed for the systems where CDs play different roles, i.e., as nanocarriers, corrosion inhibitors or building blocks for even more complex structures, often with some additives such as gold particles. Those studies have not been listed in Table 1. The selected newest articles on the topic are [13–18].

It should be noticed that to increase the accuracy of DFT calculations in solid state, periodic boundary conditions are often applied, using crystal unit cells as simulation boxes. During the computations, only the properties of the original unit cell need to be calculated and then propagated in the chosen dimensions. However, to perform such computations, the crystal structure of the studied object is mandatory. More information on such calculations can be found in a recent review [19] with an example of such calculations for CDs presented in [20].

- Software

In almost all of the reviewed works presenting the results of the DFT calculations, Gaussian software was used. There are only a few cases when ORCA [21–23], VASP [24] or ADF [25] were applied instead. Only for the solid state calculations or those involving adhesion, nanocarriers, etc., is DMol3 also commonly used.

- ONIOM

In several works that included CD where DFT was applied, the ONIOM method was presented (see Table 1). ONIOM stands for Our own N-layered Integrated molecular Orbital and Molecular mechanics and is a hybrid method which combines QC (either ab inito or semi-empirical) and molecular mechanics methods in order to reduce the computational cost [26]. On the basis of the ONIOM results, other properties such as thermodynamic ones are calculated. This approach was for years popular in the computation of the CD systems, as CDs are relatively big structures, and for a long time it was not possible to calculate both the CD and a guest using DFT. Therefore, CD was considered an outer layer and there a lower level of theory was applied, and the guest molecule was computed using a higher level of theory. However, with the general increase in computational power available, this approach is currently rarely used in the studies of CD complexes due to its lower accuracy when compared with pure QC calculations.

- Molecular dynamics

An MD simulation is a well-established technique used for the study of various molecules complexes and mixtures in any state of matter and at almost any temperature and pressure condition. It can be used to determine structural, energetic, and thermodynamic properties as well as a means to scan the potential energy surface of a studied system.

For MD simulations of large molecular complexes, such as ligand–protein, molecular mechanics (MM) methods are commonly used. On the contrary, when MD simulations are performed on relatively small molecules, it is usually at the quantum mechanics (QM) level of theory, which significantly increases the accuracy of calculations, but also their computational costs. In terms of the sizes of the modeled objects, CD complexes are somewhere in between. While geometry optimization calculations on the static structures of CD complexes are, nowadays, performed mostly at the QM level, usually by the means of DFT, the MD simulations are still being performed at the MM level [12].

2.3. Semi-Empirical Methods

Semi-empirical methods are based on the Hartree–Fock equation but simplified by the application of the empirical corrections [27]. Semi-empirical calculations are much faster than their ab initio counterparts, mostly due to the use of the zero differential overlap approximation. Their results, however, can be very wrong if the molecule being computed is not similar enough to the molecules in the database used to parameterize the method. Here, we will concentrate on the most popular semi-empirical methods, that is AM1, PM3, PM6, and PM7.

According to Figure 2, the most popular semi-empirical method applied for the cyclodextrin complexes is PM3. However, it may be argued whether PM3 delivers better results than the newer generations, PM6 and PM7. The reason for wider application of PM3 instead of PM6 and PM7 may be the fact that some researchers got used to PM3 and some software does not support the newer parametrizations (PM6 and PM7). Since 2019, the DFT methods have strongly overtaken the role of a leading quantum chemistry-based calculation method in the cyclodextrin systems. Naturally, DFT is more precise than any semi-empirical approach, which leaves no space for further investigation, and among the semi-empirical methods, seems to be the most reliable in the complexes in question. Below, the application of the semi-empirical methods in the cyclodextrin-including complexes is described based on the examples from 2015–2022. However, it should also be noticed that the PM3, PM6, and PM7 approaches often co-exist with the DFT ones. In those cases, the systems are first optimized with the semiempirical method to obtain the initial structure for the DFT calculations.

The oldest of all here presented semi-empirical methods is the AM1 approach. There are just a few examples of the cyclodextrin complexes calculated using this approach. In some of them, AM1 was applied to perform the geometrical optimization of a whole guest–CD complex [28]; in others, both substrates were calculated by AM1 whereas the complex underwent the DFT treatment [29], and in others, AM1 was applied solely for the CD.

The next generation of semi-empirical methods is PM3. There are studies showing that PM3 predicts the presence and energy of hydrogen bonds better than AM1 [30]. Moreover, it has been reported that after α- and β-CD optimization with AM1 and PM3, the former resulted in badly distorted geometries, whereas the latter reproduced the crystalline structure rather well [31].

Semi-empirical methods are repetitively reported to deliver good insight into the complex formation process as well as reliable order of the configuration stabilities [32–35]. These methods help to determine global or local minima. However, it is stressed that to obtain reliable complexation energy values, DFT calculations should be performed [36]. Often, the PM3 approach is undertaken along with the ONIOM DFT/PM3 approach [34,37–44].

The most often and standard PM3 application in the cyclodextrin complexes is to move the guest along the selected axis going through the CD cavity. The guest is stopped every 1 Å, usually between −8 Å or −10 Å and respectively +8 Å or +10 Å. Additionally, the guest is rotated from 0°to 360° usually every 20° or 45°. At each such stopping point, the complexation energy is measured with PM3 [37,44–47].

In the way of their application to the cyclodextrin systems, the PM6 and PM7 approaches follow the same pattern as PM3. Namely, they are used to gain insight into the complex structure [48] and thermodynamic properties [48–52], which allows to determine the most stable complex [52–55], and, as in case of ofloxacin enantiomers, rank the eluted substances in the order in which they will be eluted [49]. Similarly, as in the case of PM3, FT-IR spectra can be simulated [50,56]. Further, PM6 and PM7 are often combined with the DFT methods in form of the ONIOM approach [42,52,57,58].

In contrast to PM6, in PM7, the description of dispersion interactions and hydrogen bonding has been improved, and consequently the errors associated with modelling large molecules and complexes have been reduced [49]. The description of properties such as heat of formation or height of the reaction barrier has been improved [59]. In turn, it has been reported that PM6, compared with PM3, can yield better agreement with the experimental values [59]. In another study, when compared to the experiment, PM3 provided wrong and at the same time opposite results to those obtained by PM7 [60]. To depict another example, in a study where β-CD, dimethyl-β-CD, and hydroxypropyl-β-CD were analyzed by both PM6 and PM7, in all cases PM7 delivered complexation energies of significantly lower values [58].

A separate topic is ADMP, the Atom Centered Density Matrix Propagation Molecular Dynamics approach, which can be performed with semi-empirical, Hartree–Fock or DFT

methods. It provides equivalent functionality to Born–Oppenheimer molecular dynamics at a considerably reduced computational cost. The ADMP method has a number of attractive features. Systems can be simulated by accurately treating all electrons or by using pseudopotentials. Through the use of a tensorial fictitious mass and smaller values of the mass, reasonably large time steps can be employed, and lighter atoms such as hydrogens need not be replaced with heavier isotopes. A wide variety of exchange-correlation functionals can be utilized, including hybrid density functionals [61]. In the last decade, only two cases have been published: β-CD-olsalazine with PM3-ADMP [62] and β-CD-propranolol with PM6-ADMP, ONIOM(DFT/PM3)-ADMP, and DFT-ADMP approaches [63]. ADMP results confirm the importance of the non-bonded interactions in the complex stabilization.

To sum up this part of the review, a relatively new and complex study should be cited. In the article entitled 'Prediction of correct intermolecular interactions in host–guest systems involving cyclodextrins' [63] published in 2020, the following approaches were tested: AM1, PM3, PM6, and DFT with the most standard B3LYP/6-31G(d,p), as well as PM6 and DFT with and without dispersion correction. The study involved 15 α-CD and 28 β-CD inclusion complexes in terms of both geometrical parameters and complexation energy values. The results showed that the most accurate was the B3LYP/6-31G(d,p)-D3 approach, followed by PM6-D3. Nevertheless, taking into account the high computational requirements of the DFT methods, the authors suggest that PM6-D3 is the most accurate and cost-effective approach. However, it must be stated that as the availability of the computational power is developing quickly, the DFT approach might shortly be, or already is, the best option for the computational analysis of the structure and energy of CD systems.

2.4. Density Functional Tight Binding (DFTB)

An approach that can be positioned between semi-empirical methods and DFT is the Density Functional Tight Binding Self Consistent Charge method (DFTB-SCC, referred here to as DFTB), which is often described as DFT approximation [64]. According to the Web of Science, just a few articles referring to DFTB and CD complexes have been published. However, those works show a relatively wide spectrum of possible DFTB applications. In the oldest works [64,65], DFTB was applied to verify the experimental NMR results and deliver some additional structural information. In the first case [64], the method's application confirmed which conformation of spironolactone was preferred in the complex. In the second case [65], DFTB confirmed that in the analyzed peptide, tyrosine was a favored residue to access the CD's cavity. In this second study, DFTB was applied within the QM/MM approach. In both of the described cases, the authors claimed good agreement of the obtained computation data with the experimental data.

A more complex situation is described in [66], where the topic is self-inclusion (in the own cavity) of the CD's substituents. Here, DFTB was compared with the DFT approach. DFTB is said to be method of choice as it predicts the stability order of the analyzed complexes properly, delivers the energy data that are close to the DFT data, and is faster than DFT. Both options including dispersion correction and the absence of this correction have been tested in DFTB and in DFT. The results plainly show that in both cases, application of the correction is necessary. In the case of DFTB, the dispersion-corrected version of calculations result in lower complexation energies and the order is maintained. Nevertheless, the authors emphasize that the application of an empirical dispersion correction may significantly overestimate dispersion interactions, and therefore a comparison with a rigorous DFT method should be done.

Some drawbacks of DFTB have been pointed out in the work that targeted the largest CD for which the crystallographic structure is known [67]. The heavy atoms' RMSD between the optimized structure and the crystallographic one were between 0.89 Å and 1.35 Å for DFT, depending on the parameters applied, with the best results for the B3LYP functional and 0.95 Å for DFTB. However, even if the overall RMSD looks good, large

discrepancies in the angles sizes when compared with the experiment have been reported in the case of the DFTB approach as opposed to all DFT methods.

In turn, the article from 2018 utilizes the DFTB approach to analyze the tautomerization process during encapsulation of genistein in CD [68]. The results are clear and deliver important information on the complexation, namely, 'DFTB-based MD simulations reveal that spontaneous keto-enol tautomerization occurs even within a hundred picoseconds, which suggests that the encapsulated genistein is complexed in the ordinary enol form of the drug molecule'.

2.5. Density Functional Theory (DFT)

- Functionals

While performing the DFT calculations, the main two parameters that must be decided on are type of functional and a basis set. Functionals mathematically define the electronic energy, which when added to the kinetic and electrostatic energy of the system, sums up to the total system's energy [68,69]. According to the literature (see Table 1) for the systems that included CD, the hybrid B3LYP [70], semi-empirical GGA (generalized gradient approximation): wB97XD, B97D3 [71] or meta-GGA kinetic energy density incorporating Minnesota [72] functionals have been applied so far. Among the last category, M06-2X, M05-2X, and M06-L are used, with M06-2X being the most common in the analysis of the non-covalent interactions, whereas M05-2X includes 0% Hartree–Fock (HF) exchange, and M06-2X has 54% HF exchange [73]. In one of the studies including CDs, it was concluded that M06-L delivered poor results [74]. This was expected as the analyzed complex was β-CD-alprazolam, whereas M06L has been designed for calculations of the systems including transition metals, inorganic or organometallics [21]. wB97X and 97D3 are comprised of 22% Hartree–Fock exchange in the short range and 100% Hartree–Fock in the long range [75].

- Dispersion correction

Noncovalent forces, such as hydrogen bonding and van der Waals interactions, are crucial for the formation, stability, and function of most CD complexes. At present, ubiquitous van der Waals interactions can only be accounted for properly by high-level quantum-chemical wavefunctions or by the Quantum Monte Carlo method. In contrast, the correct long-range interaction tail is absent from all popular local-density or gradient corrected exchange-correlation functionals of DFT, as well as from the Hartree–Fock (HF) approximation. A long-range electron correlation effect, known as the London part of the dispersion energy term, is not included in the Kohn–Sham DFT equation [76]. For years this was an issue affecting the accuracy of the DFT calculations. Nowadays, several dispersion correction methods are available. Nevertheless, their inclusion not always improves the calculation effect, hence this should be tested separately for each system in question. The most widely used dispersion corrections are TS (Tkatchenko-Scheffler) [77], GD (Grimme Dispersion, written also as D) [78], and MBD (Many-Body Dispersion) [79]. However, to perform calculations on the systems that included CD, almost solely the semi-empirical Grimme dispersion correction was applied (see the Table 1). It occurs in the D2, D3, D4, and D3(BJ) versions, where BJ indicates Becke Johnson damping. This last one is rarely used in CD-complex calculations. It is claimed that 'the damping function in DFT-D methods has only a minor impact on the quality of the results' [80] and even the comparison between D3 and D3 (BJ) published by Stefan Grimme clearly indicates that 'the differences between the two methods are much smaller than the overall dispersion effect' [80]. D3 includes less empirical input than D2, can be called a newer D2-version, and is the dispersion correction that is currently the most widely used.

No dispersion correction is applied to the Minnesota functionals that are parametrized for dispersion. The same applies to wB97X (written also as ωB97X), which is the dispersion-incorporating version of B97X. B3LYP is sometimes used as B3LYP-CAM (Cambridge extension) [81], which includes the long-range correction; however, this is not common,

and there is no strong evidence that this type of dispersion correction gives better results than application of the Grimme correction.

Sometimes it seems reasonable to perform the same calculations with two chosen functionals as has been done for the β-CD-2,2′-bipyridine complex [82]. The authors' conclusion is that wB97XD showed reliability in elucidating weak interactions, whereas B3LYP allowed one to achieve a good time–precision compromise.

Another interesting example is comparison of three types of functionals performed for the β-CD-procaine HCl system [83]. According to the authors, B3LYP showed the highest efficiency and quality of results, wB97XD was specifically used to analyze the long-range interactions, and M06-2X was applied to predict the presence of hydrogen bonds. On the other hand, there are works such as [84] (β-CD-8-Anilinonaphthalene-1-sulfonate) where it is claimed that among tested functionals, B3LYP, wB97XD and M06-2X, the overall best results were delivered by wB97XD.

To take one more example, for the β-CD-benzocaine system [53], where B3LYP, CAM-B3LYP, M05-2X, and M06-2X were tested, M06-2X was claimed to deliver excellent results when used to obtain the NMR spectra. In another study, for UV–Vis spectrum simulation, the B3LYP-D3-based results showed the best agreement with the experimental data. The tested functionals were BLYP-D3, B3LYP-D3, and M06-2X-D3 [22]. In the case of β-CD-5-fluorouracil, inclusion of the dispersion correction changed the interaction energy by 15% and by 20% in water and ethanol solvent, respectively [85].

This only allows us to draw three important conclusions. Firstly, a couple of functionals, preferably representing all three groups (hybrid, GGA, Minnesota) should be tested for each system. Secondly, the choice of a functional depends on the goal of the study: geometry optimization, thermodynamic parameters, NMR spectra, etc. Thirdly, this literature review sends a clear message that the three most used and effective functionals are B3LYP-(D3), wB97XD, and M06-2X.

- Basis set

It should not be forgotten that in the study preparation, the functional and dispersion correction and the basis set choice play a significant role. The basis sets applied for the CD complexes so far are the Pople, correlation-consistent (cc-pVDZ), and Karlsruhe (def-TZVP, def2-SVP) basis sets (see Table 1). However, the Pople ones are definitely the most common and the variety among them is large, for instance, 6-31G(d), 6-31G(d,p), 6-31+G(d,p), 6-311++G(d,p), 6-31G**, where '1′ means basis set enlargement, '+' means an additional diffuse function, and '*' means a polarization function. The choice among different Pople basis sets depends partly on the available computational possibilities and partly on the type of studied objects. Inclusion of diffuse functions is needed to properly calculate the long-range interactions, such as hydrogen bonds. In turn, by extending the size of a basis set, the addition of polarization functions is meaningless. This is why, in the works presented in Table 1, the most presented basis sets are 6-31G(d,p) and 6-311G(d,p).

In Table 1, it is noticeable that the '6-31+G* for H, N, O and 4-31G for C' basis set combination is present. However, first of all, several of these works have been published with one affiliation, in other words, there is one laboratory that uses such an approach, and secondly, such a combination of basis sets for one system has been used mainly in the past when insufficient computation possibilities were at hand. In order to perform computation using the limited available tools, it is common practice to perform the geometry optimization in a lower basis set and later, for instance, for single point calculations, a larger basis set is used. An analogical approach applies to the more and less computationally demanding functionals.

2.6. Solvent

The last parameter to decide on refers to the environment of the system. Calculations can be performed in gas or in solvent. To simulate a solvent in DFT, implicit solvent models are typically used. For the CD systems, the most popular is the family of the Polarizable Continuum Models (PCM) [86]. The other possibility is the Solvation Model Based on

Density (SMD). IEFPCM [87] is a reformulation of the dielectric PCM, and the C-PCM is a conductor-like PCM, closer to the COSMO model. SMD defines the free energy of solvation via two components: the one is electrostatic contribution arising from the self-consistent reaction field, the other comes from the short-range interactions between the solute and solvent molecules [88].

Both PCM and SMD treat the solvent as a continuum because using Quantum Mechanics, it would not be possible to calculate a system with a large number of explicit solvent molecules. However, this issue can be partly approached as has been done in the study [64]. There, for general calculations, IEFPCM has been applied, but additionally a separate set of calculations has been done on the limited number of water molecules placed inside the CD cavity. Such an approach is useful if there is the probability that the CD–guest interaction is influenced significantly by the solvent's presence. For example, when it is assumed that the hydrogen bonds between the guest and host molecules are water mediated. Sometimes, inclusion of the solvent effect decreases the complexation energy significantly, as in the case of the dexamethasone and SMD model [89]. However, it must be pointed out that inclusion of a solvent in the calculated system not always results in better (closer to the experimental data) complexation energies.

Taking into account all what has been written above, in order to make a fully justified selection of the parameters for the DFT calculations, a cross-study including different but most commonly used functionals, dispersion corrections (presence or absence), basis sets, and the environment (gas or solvent, type of solvation) should be performed. To the authors' best knowledge, so far no such study has been undertaken on any system that includes CD. The already published benchmark studies are quite uncommon and usually focus on modification of one of the parameters, i.e., functionals, dispersion correction or the solvation method [67,90–92].

Only afterwards, with the parameters chosen carefully for the analyzed system (e.g., CDs+steroidal hormones, CDs+flavonoids etc.), should further calculations be performed.

2.7. Møller–Plesset Perturbation Theory 2 (MP2)

Only a few (five in the period from 2014–2021) articles in which the MP2 method has been applied for CD analysis have been published. The reason is the fact that this technique is computationally more demanding than DFT and since CD complexes are relatively large, as the objects for QC studies, this method is currently not affordable for most of the computational researchers.

The most recent work in this topic was published in 2017 and concerns β-CDs with one large substituent that can either be located in or outside of the CD's cavity [68]. The geometry optimization has been performed with the B3LYP-D3 functional but single point calculations already with both B3LYP-D3 and MP2 using various basis sets, for MP2: 6-31G* and 6-311G*.

In another study (β-CD-sertraline) [93], MP2/6-31G(d,p) was applied for the single point calculations, although even the authors of the work state that such a small basis set does not allow one to obtain results with the required accuracy.

3. Preparation of Structures, Post-Processing Methods, and Some Examples

3.1. Preparation of the CD Complexes for the QC Calculations

To obtain the structure of the CD complex that can be used for DFT calculations, both the structure of the chosen CD as well as the structure of the guest molecule must be prepared beforehand. The method of in silico complex preparation is also important, as it may have a major influence on the results.

Since the crystal structures of all of the native and also some of the modified CDs can be found in the CCDC [94], they are usually used as the starting points for calculations. The structure of the guest is either simply drawn using one of the multiple available software packages or taken from the CCDC, assuming that its crystal structure has been deposited previously.

Very important, for the accuracy of the results, is the method of preparation of the complex from its components. This must be done unless a crystal structure of the complex has been obtained or deposited previously in the CCDC, which is unfortunately quite uncommon. Usually, one of the two approaches is used. In the first one, molecular docking is applied, treating CD as a macromolecule and the guest as a ligand. In the reviewed works, the authors usually use the popular Auto Dock [95] software for that purpose. However, some other programs are also used such as Schrodinger Maestro [96] or BIOVIA Discovery Studio [97]. Surprisingly, in the reviewed studies, not much attention was being paid to the description of this part, which was justified by the fact that the initial (docked) structures would be optimized at the higher theory level. This may, however, lead to some inaccurate or even wrong results as the energetically lowest conformation obtained from the docking part may not necessarily be close to the global DFT minimum. This is nicely reflected in Figure 3, where the energetically lowest orientation obtained from molecular docking is substantially different from the experimental one, even after optimization using DFT. However, when the other pose from molecular docking was optimized using QC, much better agreement between the experimental and theoretical results was obtained. Therefore, in some studies, the authors decided to optimize not one but a few different complexes from molecular docking. While this approach is reasonable as it increases the likelihood of finding the deep minimum, it should be noticed that the time of calculations increases linearly with the number of initial structures.

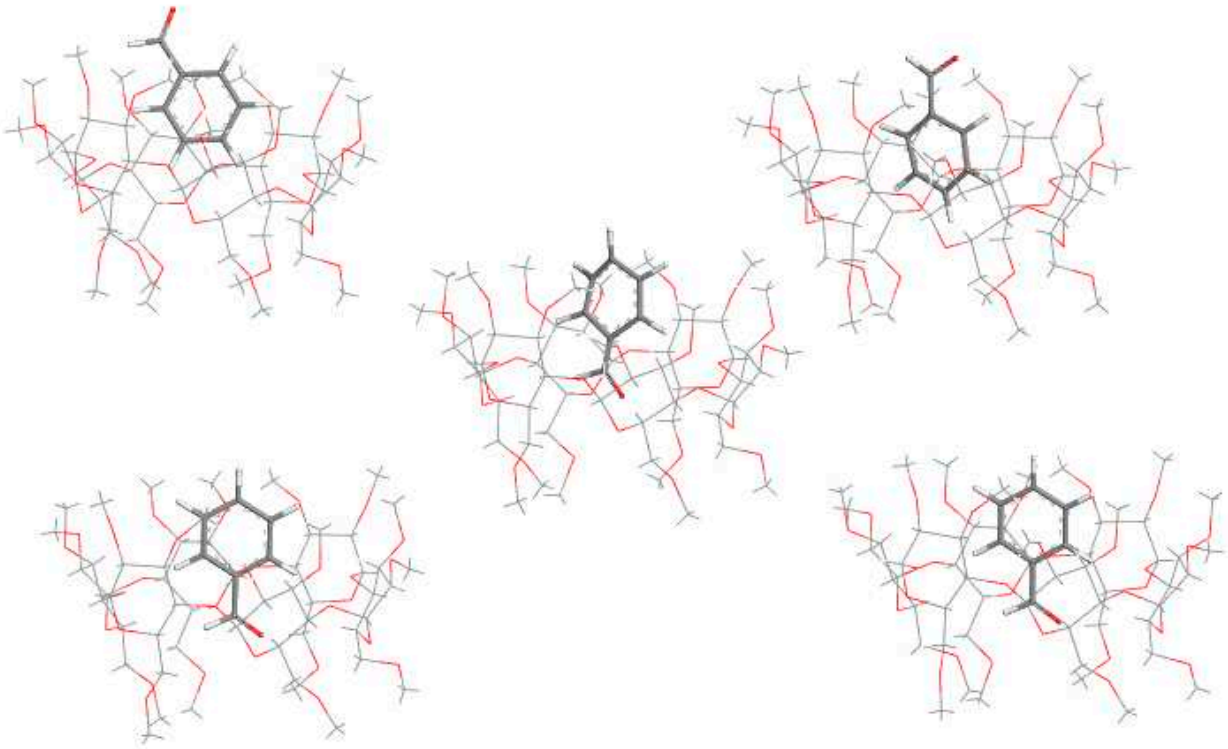


Figure 3. Comparison between the structures of the αCD complex with benzaldehyde. Top left: the best pose from molecular docking; top right: the best pose from molecular docking after optimization using DFT; bottom left: one of the poses obtained from molecular docking; bottom right: "bottom left" structure after optimization using DFT; middle one: experimental structure (CSDC ref. code: BOHWUQ). It should be noted that while the top left structure has energy lower than the bottom left by 3.4 kcal/mol, the top right structure has energy higher than the bottom right by 4.2 kcal/mol. Source: author's archive.

Instead of molecular docking based on the molecular mechanics calculations, in some cases, the authors decided to manually dock the guest into the CD. In order to find the best pose within the cavity, the guest molecule is put in different positions along the selected

axis, so that the guest has different levels of immersion into the CD's cavity. For example, in [98], the guest was moved along the *Z*-axis from +7.5 Å to −7.5 Å with an interval of 0.3 Å. Additionally, in this particular study, the guest was rotated around the *Z*-axis by 3° from 0° to 360°. In each step, the generated systems underwent geometry optimization calculations. This type of systematic search seems to be the most accurate approach, especially for the ligands with limited conformational space, with the only drawback being the increase in the calculation time.

The other important factor that is usually neglected in the QC studies of CD complexes is the conformational flexibility of the guest molecule. The optimal conformation of the guest found *in vacuo* is not necessarily the one that it takes in the complex. To increase the likelihood of finding the deep energetical minimum, the conformational space search of the guest molecule should be performed.

Another aspect that must be taken into consideration is the host–guest molar ratio of the complex. When there are some experimental indications for a specific value, the assumption can be tested using QC calculations. Otherwise, it seems reasonable to prepare the complexes of various stoichiometry and confirm their stability via geometry optimization.

3.2. Description of QC Results

After the host, guest, and complex are optimized at the chosen QC level, the interaction and stabilization energies are obtained. Stabilization energy is defined as the difference between the energy of the fully optimized geometry complex and complex components: CD and guest (Equation (1)) [99]. Interaction energy is defined as the corresponding single point energy.

$$E_{stb} = E_{cplx\_opt} - (E_{CD\_opt} - E_{guest\_opt}) \quad (1)$$

Sometimes the solvation energy is taken into account as well. It is calculated as the difference between the complex energy in water and in gas. Thermodynamic parameters (TD) are often calculated, as they give more insight into the stability of the analyzed systems.

The collected data allow us to draw conclusions, which forces determination of the complex creation: dispersion [100], van der Waals [63] interactions or hydrogen bonds [61]. TD results allow us to determine whether the complexation process is enthalpy driven [51,84,100], which is said to relate to the number and strength of the intermolecular interactions within the system, or entropy driven, which is quite rare for those complexes. Inclusion of the temperature effects allows us to observe if and how the temperature affects the complex stoichiometry, as in the case of β-CD-pentoxifilline [101].

Additionally, often IR or UV–Vis spectra or NMR chemical shifts are calculated (see Table 1) and in the majority of cases, the results are claimed to have very good agreement with the experimental data.

Another common practice is the application of the QTAIM method (Quantum Theory of Atoms In Molecules) for the DFT-optimized complexes in order to analyze weak interactions and therefore obtain a better understanding of the complex's structure at the molecular level. Several articles about CD complexes including this approach have been published (see Table 1).

3.3. Analyzed CD Complexes

Among the large variety of CDs, in the DFT studies, mainly only β-CD has been applied so far, and a variety of guests in CD complexes has been analyzed, as presented in Table 1. The guests are mainly drugs among which the antidepressants seem to be especially targeted (Table 1A.1) as well as plant derivatives with (potential) medical use (Table 1B). A separate group consists of substances that could be defined as functionalized food (Table 1C). In those cases, CDs serve to protect or even increase the antioxidative capacity of the substances in question, for example (−)-gallocatechin, (−)-catechin gallate and (−)-gallocatechin present in tea [102–104], or to reduce the bitter taste of coffee [105,106]. Moreover, CDs can be used as chiral selectors, and this has its reflection in the DFT articles (Table 1D).

Some of the analyzed systems in Table 1 have been already described in the previous paragraphs as examples regarding the applied parameters and computation. Other selected examples showing a particular usefulness and applicability of the DFT methods as well as the obtained results are described below.

The examples that directly show a vast area of DFT applicability for CD systems analysis are studies of the β-CD complexes, with 8-anilinophthalene-sulfonate [84], benzyl isothiocyanthe [23], metheonine [98] or vanillina [107]. In these works, the DFT approach was used to perform geometry optimization, obtain interaction and stabilization energies, calculate thermodynamic properties, use QTAIM and NBO analysis approaches, and simulate NMR and absorption spectra. The obtained data allow the screening of possible conformations, to define the interactions (van der Waals, hydrogen bonds, etc.) determining the CD–guest interaction, rank the complexes according to their stability, complement experimental spectra, and support the signal assignment.

Further, such chemical information happens to be a crucial part of new theses. For instance, DFT calculations revealed that the nicotine forms have considerably stronger binding with β-CD rather than with Mβ-CD in the same orientation with lower complexation energy. This explains why after 21 days the remaining nicotine increased from 65.56% in pure nicotine to 89.32% and 76.22% in β-CD-nicotine and Mβ-CD-nicotine complexes, respectively [34].

Another example is imipramine and desipramine β-CD complexes [108]. DFT calculations revealed an alternative inclusion scenario: via a guest's side chain and not via the aromatic moiety. Thus, the controversy in the experiments has been explained because such bimodal complexation increased the therapeutic effect of the substances.

For another antidepressant, paroxetine, the DFT calculations helped to confirm the existence of a new CD inclusion polymorph: a new 2:1 stoichiometry complex has been described [109]. It is characterized by a stronger presence of the dispersion interactions and is more energetically favorable than the 1:1 complex, which improves the drug's bioavailability.

Again, when it comes to structural information, the DFT approach showed that in 1:2 Cu-flavonoid and 1:3 Fe-flavonoid β-CD complexes, in morin and quercetin, the 3-OH site, and in primuletin, the 5-OH site, were utilized as preferable chelation sites [25]. These data are helpful for scientists trying to obtain an effective CD-flavonoid antidiabetic formulation.

Table 1. Selected articles published in the years 2015–2022 on the application of DFT methods for systems that included CD. The functional and basis set information concerns CD complexes, not guests. Abbreviations used in table: DM-CD (2,6-dimethylo-CD), TM-CD (trimetylo-CD), per-M-CD (permethylated-CD), geo. opt. (geometry optimization), SP (single point calculations), NMR (1H NMR spectra simulation), NBO (Natural Bond Orbitals), BJ (Becke Johnson damping function),TD (thermodynamics calculations), n.i.p. (no information provided). In the case where the DFT application in the published research occurs only as an ONIOM component, the article has not been included in the table. The ONIOM approach along with the examples has been described in Section 2.2.

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
A (potential) drugs							
1	β	(s)-2-Isopropyl-1-(o-nitrophenyl) Sulfonyl) Aziridine	B3LYP, WB97X-D, B97D3	6-31G(d)	gas, water		[110]
2	β	boron-based aromatic systems	BLYP-D3(BJ)	def2-SVP	vacuum, CPCM	geo. opt., natural bond orbital calculations (NBO), complexation energy	[100]

Table 1. *Cont.*

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
3	α, β, γ	alprazolam	B3LYP, M06L	def-TZVP	vacuum	geo. opt. in gas, NMR spectra	[21]
4	β	lenalidomide	B3LYP, M06-2X	6-31G(d,p)	PCM		[111]
5	β	dexamethasone	BLYP-D4	def2-TZVP	gas, water	geo. opt., complexation energy	[89]
6	β	2,2′-Bipyridine	B3LYP, wB97XD	6-31G(d)	PCM (eight solvents)	geo. opt., UV–Vis spectrum, HOMO-LUMO	[82]
7	β	2,2′-Dipyridylamine	B3LYP	6-311++G(d,p)	PCM		[112]
8		vardenafil hydrochloride	B3LYP	6-311G(2d,2p)	vacuum	geo. opt., FT-IR	[113]
9	amino-CD	doxorubicin	B3LYP	6-31G	vacuum	geo. opt., complexation energy, HOMO-LUMO, dipole moment, chemical potential, electrophilicity	[114]
10	β	5-fluorouracil	B3LYP-D3	6-31+G(d,p)	vacuum, PCM	geo. opt., complexation energy, harmonic frequency calculations	[85]
11	HP-β	2-methyl mercapto phenothiazine	B97-D3, BP86-D3	6-31G(d,p)	gas, CPCM	geo. opt., vibrational spectra, NBO, QTAIM, HOMO-LUMO	[115]
12	β	vemurafenib	ωB97XD	6-31+G(d)	vacuum, PCM	Geo. opt., vibrational spectra, MD, NBO, TD, HOMO-LUMO	[116]
13	β	procaine hydrochloride	B3LYP, M06-2X, WB97XD	6-31G(d,p)	gas, PCM	Geo. opt., NBO	[83]
14	β, SBE-β	fluorometholone, cholesterol	M06-2X	6-31G**	PCM	Geo opt., interaction energy	[117]
15	α, β, γ	chlordecone	M06-2X-D3	6-31G(d,p)	SMD	Geo. opt., QTAIM	[118]
16	β, methyl-β	nicotine	M06-2X	6-31G(d,p)	n.i.p.	Geo., opt., complexation enrgy	[34]
17	β	8-Anilinonaphthalene-1-sulfonate	B3LYP, M06-2X, WB97X-D	6-31G(d)	gas, water	Geo. opt., interaction energy, NMR, TD, NBO	[84]
18	β	benzocaine	B3LYP, CAM-B3LYP, M05-2X, M06-2X	6-31G(d,p)	PCM	Geo. opt., QTAIM, NBO, NMR, HOMO-LUMO, TD	[53]
19	β	aryl pentazole	M06-2X	6-31+G(d,p)	PCM	Geo. opt.	[119]
20	β	2,4D, dicamba pesticides	PBE1PBE (PBE0), B97-D, M06-2X	6-31G(d,p)	gas, SMD	Geo. opt.	[120]

Table 1. *Cont.*

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
21	Monochlorotriazinyl-β	permethrin, cyppermethrin	BLYP (geo. opt.); BLYP-D3, B3LYP-D3, M06-2X-D3 (UV–Vis)	def2-SV(P) (geo. opt.); TZVP (UV–Vis)	COSMO	Geo. opt.	[22]
22	β	dopamine	B3LYP, MPW1PW91, M05-2X, M06-2X, ωB97X-D	3-21G*	CPCM	Geo. opt., complexation energy, QTAIM, NBO	[61]
23	α	benzoate derivatives	M06L (geo. opt.); M06-2X//M06-L (SP)	6-31+G(d,p)	gas	Geo. opt.	[121]
24	α, β, γ	cholic, deoxycholic acid	B97-D, M06-2X, B3LYP	6-31G(d)	PCM	Geo. opt., interaction energy	[122]
25	α	benzoate derivatives	M06-2X//M06-L, M06-2X//BLYP, BLYP, M06-2X	6-31+G(d,p)	gas	Geo. opt., interaction energy	[123]
26	γ	cetirizine	B3LYP	def-TZVP	n.i.p.	Geo.opt., interaction energy, HOMO-LUMO, DOS, NMR	[124]
27	succinyl-β	uranium	M06-2X	6-31G(d,p)	SMD	Geo. opt.	[125]
28	β-CD, DM -β	thymidine-carbonate	B3LYP-GD2	6-31G(d,p)	PCM	Geo. opt., complexation energy, TD, HOMO-LUMO, NMR	[126]
29	β	glycyl-L-phenylalanine	B3LYP	3-21G(d)	PCM	Geo. opt., interaction energy, HOMO-LUMO	[127]
30	β	sodium salicylate	B3LYP	6-31G(d)	gas, PCM	Geo. opt., solvation energy, relative stabilization energy, complexation energy, change of volume	[128]
31	β	benzyl isothiocyanthe	B97-D3	def2-SVP	vacuum	Geo. opt., complexation energy, HOMO-LUMO, NBO, NMR	[23]
32	α	iodine solution	CAM-B3LYP	6-31*G	PCM	Geo. opt., absorption spectra, HOMO-LUMO	[129]
33	β	meta-aminophenol	M06-2X	6-31G(d,p)	IEFPCM	Geo. opt., complexation energy, HOMO-LUMO, TD, NBO	[130]

Table 1. *Cont.*

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
35	β	L-glutamine	B97-D3	6-31G(d)	n.i.p.	Geo. opt., complexation energy, TD, NBO, QTAIM	[131]
36	β	R and S ibuprofen	M062X	6-31G(d,p) (geo. opt.); 6-311++G(d,p) (SP)	gas, SMD	Geo. opt., solvation energy	[132]
37	α, β	thioureides	B97-D3	6–31G(d,p)		Geo. opt., interaction energy	[133]
38	β	mepivacaine	B97-D3	6-31G(d,p)	gas, SMD	Geo. opt., interaction energy, TD	[134]
39	β	L-metheonine	WB97-D3	6-31G(d)	PCM	geo. opt., interaction energy, QTAIM, TD, NMR	[98]
40	β	prazosin, losartan	B3LYP	6–311+G(d,p)	gas	Geo. opt.	[135]
41	β	olsalazine	B3LYP, WB97-D3, CAM-B3LYP (UV-vis)	6-31+G(d)	PCM	Geo. opt., ADMP	[62]
42	β	aspirin	B3LYP-D3	cc-pVDZ	gas	Geo. opt., qTAIM, NBO	[136]
43	β	quinine	B3PW91	6-311++G(d,p)	PCM	Geo. opt.	[137]
44	β	erlotinib	B3LYP	6-31+G*	n.i.p.	Geo. opt., harmonic frequencies, HOMO-LUMO	[138]
45	γ	rocuronium, vecuronium	B3LYP	6–31+G(d,p)	n.i.p.	Geo. opt., NBO, HOMO-LUMO	[139]
46	α, β, γ	cathinone	M05-2X	6-31G(d)	gas, CPCM (water, chloroform, methanol)	Geo. opt., QTAIM, NBO, IR spectra, TD	[140]
47	α	CO2	B3LYP	G-31G*	PCM	NMR	[141]
48	β	flutafemic acid	B3LYP, M05-2X	6-31G(d)	vacuum, water	Geo. opt., complexation energy, TD, NMR	[142]
49	2-HP-β	Cu (II) and Fe (III) complexes of quercetin, morin, primuletin	B3LYP	6-311++G**	n.i.p.	Geo. opt., complexation energy, HOMO-LUMO	[25]
50	β	6-thioguanine, 6-mercaptopurine	B3LYP	6-31+g(d,p)	IEFPCM (DMSO)	Geo. opt., interaction energy, TD	[37]
51	β	N-(2-chloroethyl),N -nitroso,N′,N′-dicyclohexylsulfamid	B3LYP	6-31G(d)	PCM (DMSO)	Geo. opt., NBO, QTAIM	[143]
52	β	benzaldehyde	B97-D	6-31G(d,p) (geo. opt.); 6-311++G(2d,p) (SP)	gas, SMD	Geo. opt., interaction energy, TD	[144]

Table 1. *Cont.*

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
53	α	chitibiose	M06-2X	6-311++G**	n.i.p.	Geo. opt., NBO, QTAIM	[145]
54	α	hydrated and nonhydrated IIA/IIB group metal cations	M06-2X	6-31G(d,p)	gas, PCM	Geo. opt., interaction energy, TD	[146]
55	β	nabumetone	WB97X-D, B97-D, B3LYP, M05-2X, M06-2X	6-31G(d)	IEFPCM	Geo. opt., NBO, QTAIM	[40]
56	β	propranolol	B3LYP, ωB97XB (ONIOM)	6-31+G(d)	gas, IEFPCM, explicit solvent effect: explicit water molecules inside of the complex	Geo. opt., interaction energy, ADMP, TD	[59]
57	functionalized CDs	8-hydroxyquinoline ligands	B3LYP	6-31G**	n.i.p.	Geo. opt.	[147]
58	β	pentoxifilline	M06-2X	6-31g(d,p)	gas	Geo. opt., NBO, HOMO-LUMO	[101]
59	β	p-nitropenthyl acetate	B3LYP	6-31G(d,p)	n.i.p.	Geo. opt., interaction energy, NBO, HOMO-LUMO	[148]
60	β	norfloxacin	B97D (geo. opt.), B3LYP (SP, NMR)	6-31G(d,p)	IEFPCM	Geo. opt., interaction and stabilization energy, NMR, TD	[149]
			A1. Antidepressants				
61	β	paroxetine	B3LYP (geo. opt.); B97D (SP)	6-31+G* for H, N, O and 4-31G for C	vacuum	Geo. opt., interaction energy, TD	[109]
62	2,6-DM-β	mianserin	B3LYP-GD2 (geo. opt.); M05-GD3, M06-GD3, M062X-GD3, ωB97XD, mPW1PW91, M11 (SP)	6-31G(d,p)	PCM, vacuum	Geo. opt., interaction energy, NMR	[150]
63	β	sertraline HCl, fluoxetine HCl	B3LYP	6-31+G* for H, N, O and 4-31G for C	gas	Geo. opt., interaction energy	[151]
64	β	protriptyline, maprotiline	B3LYP	6-31+G* for H, N, O and 4-31G for C	vacuum	Geo. opt., interaction and stabilization energy	[152]
65	β	clomipramine, doxepin	B3LYP	31+G(d) for H, N, O, Cl, and 4-31G for C	gas	Geo. opt., interaction energy	[153]

Table 1. *Cont.*

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
66	β	desipramine, imipramine	B3LYP	6-31þG(d) for H, N, O and 4- 31G for C	gas, implicit solvent (water)	Geo. opt., interaction and stabilization energy	[108]
67	β	amitryptyline, nortryptiline	n.i.p.	6-31+G* for H, N, O, Cl and 4-31G for C	vacuum, SMD	Geo. opt., interaction and stabilization energy	[154]
			B. Plant derivatives				
68	HP-β	thymoquinone	B3LYP-D2, B3LYP-D3	6-31G(d,p)	PCM	Geo. opt., NBO, QTAIM, HOMO-LUMO, NMR	[155]
69	α	β-carotene	B3LYP	cc-pVDZ	vacuum	Geo. opt., interaction energy, Raman spectra	[156]
70	γ	3-hydroxyflavone	PBE0	def2-SV	PCM	Geo. opt., HOMO-LUMO, IT spectra	[157]
71	β	vanillina	B3LYP, ωB97xD, M06- 2X	6-311G(d,p)	vacuum, CPCM	Geo. opt., interaction energy, NMR, HOMO-LUMO, NBO, QTAIM, UV–Vis	[107]
72	β	alfa-terpineol	B3LYP (for UV–Vis), B3LYP/CAM, M062X, WB97-D3	6-311G(d,p)	vacuum, CPCM	Geo. opt., complexation energy, NBO, QTAIM, TD, UV-vis	[158]
73	TM-β, β	naringenin	B3LYP, M06-2X, wB97X-D	6-31G(d)	vacuum	Geo. opt., interaction energy, NBO, QTAIM, NMR, HOMO-LUMO	[159]
74	2,6-DMβ, 2HP-β, 2,6-DH-β, β	eucalyptol	M06-2X	6-31G(d,p)		Geo. opt., interaction energy	[160]
75	β	fisetin	M06-2X	6-31G(d,p)	gas, PCM	Geo. opt., interaction energy	[161]
76	β	gallic acid	B97-D3	6-31G*, for GIAO: 6-311++g**	gas, solvent	Geo. opt., HOMO-LUMO, NBO, NMR	[162]
77	β	gabapentin	B3LYP-D3	6- 31G(d)	vacuum, PCM	Geo. opt., interaction energy, NBO, HOMO-LUMO	[163]
78	B, γ	tropane alkaloids	B3LYP	6-31+G(d,p)	PCM	Geo. opt., interaction energy, NMR	[164]
79	β	coumarins	EDF2	6-311G(d,p)	PCM	Geo. opt.	[165]
80	2-HP-β	quercetin	B3LYP	6-31G*		Geo. opt.	[166]
81	β	carvacrol, thymol	B3LYP	6-31G, 6-31+G(d)	SMD	Geo. opt., interaction energy, NBO, HOMO-LUMO	[167]
82	β	thymol	B3LYP, PBEPBE, CAM-B3LYP	6-31G(d,p)	PCM	Geo. opt., interaction energy, UV–Vis	[168]
83	β	carvacrol	B3LYP, M05-2X	6-31G(d)	PCM	Geo. opt., HOMO-LUMO, NBO	[42]

Table 1. *Cont.*

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
C. Functionalized food							
84	β	(−)-gallocatechin, (−)-catechin gallate, (−)-gallocatechin gallate	B3LYP	6-31+G* for H, O and 4-31G for C	gas	Geo. opt., interaction and stabilization energy	[102,103]
85	β	(−)-epigallocatechin, (−)-epigallocatechin gallate	B3PW91	cc-pVDZ	gas	Geo. opt., interaction energy	[104]
86	β	catechol derivatives: protocatechuic aldehyde, protocatechuic acid	B3LYP	6-31+G* for H, O and 4-31G for C	gas (geo. opt.), implicit solvent (TD)	Geo. opt., interaction energy, TD	[169]
87	β	oleuropein, hydroxytyrosol, tyrosol	n.i.p.	6-31+G* for H, O and 4-31G for C	gas	Geo. opt., interaction energy, TD	[170]
88	β	chlorogenic, caffeic, quinic acids	B3LYP	6-31+G* for H, O and 4-31G for C	gas	Geo. opt., interaction energy, TD	[105,106]
D. CD as a chiral selector							
89	β	D- and L-penicillamine	B3LYP-D3 (geo. opt.); M062X-D3, xB97X-D, B3LYP-D3 (interaction energy)	6-31G(d, p) (geo. opt.); G-311+G(d,p) (interaction energy)	water	Geo. opt., interaction energy	[171]
90	metal-ion coupled β	D- and L-penicillamine	DFT, M062X	6-31G(d,p)	vacuum	Geo. opt.	[172]
91	β	R- and S-propranolol	B3LYP	6-311+G(d,p)	vacuum	Geo. opt., vibrational spectra	[173]
92	per-M β	D- and L-isoleucine	B3LYP (geo. opt.), wB97X-D (IR)	6-31G*, 6-311G**	gas	Geo. opt., interaction energy, IR spectra, TD	[174]
93	per-M β	D- and L-alanine	B3LYP, wB97X-D, M06-2X	6-31G**, 6-311G**		Geo. opt., IR spectra	[175]
94	2,3,6-TM-β	cis-(2S,4R) and -(2R,4S) ketoconazole	B3LYP	6-311G(d,p)	gas (geo. opt.), PCM (SP)	Geo. opt., interaction energy	[36]
95	2-HP-β	abacavir enentiomers	PBE	6-31G*	PCM	Geo. opt., interaction energy	[176]

4. Conclusions

The number of studies concerning CDs complexes in which the theoretical calculations at the QC level have been used has constantly increased since the beginning of the 21st century. Solely for the DFT-based works in this topic, the number of published articles has already exceeded 300. While this number is still relatively low, when compared to the amount of reported molecular dynamics simulations at the molecular mechanics level [12], the reviewed works reveal that the application of QC calculations in the studies of CD complexes can be essential, providing the results unobtainable by any other method, both experimental and computational.

Initially, for those kind of studies, less computationally demanding semi-empirical methods have been applied (mostly PM3, PM6, and PM7). However, since 2015, each year there have been more papers in which DFT has been chosen instead of the semi-empirical approach. Nevertheless, even in some works published after 2020, the authors have found that the PM6 or PM7, when used with appropriate dispersion correction, can provide results with similar accuracy to those obtained using DFT.

Regarding calculations of geometries and interaction energies with DFT methods, in most of the reviewed works, the inclusion of dispersion correction was found to be crucial to obtain accurate energies, irrespective of the basis set and functional used. As for the functionals, there is no surprise that B3LYP, which is the most commonly applied one in the field of organic molecule calculations, is also the one that is most extensively used for the studies of CD complexes. However, in some works where the authors have used wB97X or M06-2X instead, different results have been obtained both in terms of predicted geometry as well as stability ranking.

When preparing the complexes, some authors prefer to manually dock the molecules, systematically moving the guest towards the CD cavity and rotating it; however, in most of the works, the molecular mechanics docking procedure has been applied, usually employing the popular and freeware Auto Dock. It should be noticed that the best pose from docking was not always the one with the lowest DFT energy; therefore, at least a few different poses should be optimized in order to achieve credible results.

Though in most of the reviewed works the authors have limited their calculations solely to geometry optimization of one or a few conformations, in some of the articles, the complex properties have been computed. Successful application of DFT methods include prediction of UV–Vis, IR, and NMR spectra as well as HOMO-LUMO and NBO calculations.

The major problem with the DFT calculations seems to be the solvent treatment. In the vast majority of the reviewed works, the authors decided to apply an implicit solvation model, usually PCM or SMD. However, other studies have shown that the role of water in the complex formation can be crucial as the water-mediated hydrogen bonds between the host and guest have been observed many times. Further, the role of water release from the cyclodextrin cavity during the complexation significantly affects the thermodynamics of such a process, which can be modeled accurately only using explicit water models.

Finally, it should be emphasized that even when carefully choosing the appropriate DFT method (applied functional, basis set, solvation scheme, dispersion correction, credible initial conformation), the obtained results can still be different from the corresponding experimental ones. This is due to the high flexibility and dynamics of most of CD complexes. Therefore, it seems reasonable to explore the application of ab initio molecular dynamics simulations. While, at this moment, it may be computationally not affordable for many researchers, combining the benefits of molecular dynamics simulations with the accuracy of DFT calculations seems to be the solution to obtain even more accurate results.

Author Contributions: Conceptualization, A.H.M. and Ł.S. writing—original draft preparation, A.H.M., Ł.S.; writing—review and editing, A.H.M., Ł.S.; supervision, Ł.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Aiassa, V.; Garnero, C.; Longhi, M.R.; Zoppi, A. Cyclodextrin Multicomponent Complexes: Pharmaceutical Applications. *Pharmaceutics* **2021**, *13*, 1099. [CrossRef] [PubMed]
2. Poulson, B.G.; Alsulami, Q.A.; Sharfalddin, A.; El Agammy, E.F.; Mouffouk, F.; Emwas, A.-H.; Jaremko, L.; Jaremko, M. Cyclodextrins: Structural, Chemical, and Physical Properties, and Applications. *Polysaccharides* **2022**, *3*, 1. [CrossRef]
3. Hădărugă, N.G.; Bandur, G.N.; David, I.; Hădărugă, D.I. A review on thermal analyses of cyclodextrins and cyclodextrin complexes. *Environ. Chem. Lett.* **2019**, *17*, 349–373. [CrossRef]
4. Szejtli, J.; Szente, L. Elimination of bitter, disgusting tastes of drugs and foods by cyclodextrins. *Eur. J. Pharm. Biopharm.* **2005**, *61*, 115–125. [CrossRef]
5. European Medicines Agency. Available online: https://www.ema.europa.eu/en (accessed on 29 May 2022).
6. U.S. Food & Drug Administration (FDA). Available online: http://www.fda.gov/cder/guidance/index.htm (accessed on 29 May 2022).
7. Pharmaceutical and Medical Devices Agencty. Available online: https://www.pmda.go.jp/english/index.html (accessed on 29 May 2022).
8. Tran, D.N.; Colesnic, D.; de Beaumais, S.A.; Pembouong, G.; Portier, F.; Queijo, A.; Tato, J.V.; Zhang, Y.; Ménand, M.; Bouteiller, L.; et al. Cyclodextrin-adamantane conjugates, self-inclusion and aggregation versus supramolecular polymer formation. *Org. Chem. Front.* **2014**, *1*, 703–706. [CrossRef]
9. Carrazana, J.; Jover, A.; Meijide, F.; Soto, A.V.H.; Tato, J.V. Complexation of Adamantyl Compounds by β-Cyclodextrin and Monoaminoderivatives. *J. Phys. Chem. B* **2005**, *109*, 9719–9726. [CrossRef]
10. Jug, M.; Mura, P.A. Grinding as Solvent-Free Green Chemistry Approach for Cyclodextrin Inclusion Complex Preparation in the Solid State. *Pharmaceutics* **2018**, *10*, 189. [CrossRef]
11. Wdowiak, K.; Rosiak, N.; Tykarska, E.; Żarowski, M.; Płazińska, A.; Płaziński, W.; Cielecka-Piontek, J. Amorphous Inclusion Complexes: Molecular Interactions of Hesperidin and Hesperetin with HP-B-CD and Their Biological Effects. *Int. J. Mol. Sci.* **2022**, *23*, 4000. [CrossRef]
12. Mazurek, A.H.; Szeleszczuk, Ł.; Gubica, T. Application of Molecular Dynamics Simulations in the Analysis of Cyclodextrin Complexes. *Int. J. Mol. Sci.* **2021**, *22*, 9422. [CrossRef]
13. Dehghani, A.; Bahlakeh, G.; Ramezanzadeh, B.; Mofidabadi, A.H.J.; Mostafatabar, A.H. Benzimidazole loaded β-cyclodextrin as a novel anti-corrosion system; Coupled experimental/computational assessments. *J. Colloid Interface Sci.* **2021**, *603*, 716–727. [CrossRef]
14. Roschi, E.; Gellini, C.; Ricci, M.; Sanchez-Cortes, S.; Focardi, C.; Neri, B.; Otero, J.C.; López-Tocón, I.; Smulevich, G.; Becucci, M. Surface-Enhanced Raman Spectroscopy for Bisphenols Detection: Toward a Better Understanding of the Analyte–Nanosystem Interactions. *Nanomaterials* **2021**, *11*, 881. [CrossRef] [PubMed]
15. Paulino, P.H.S.; Silva, C.F.; De Almeida, B.D.; Guimarães, L.; Nascimento, C.S. A theoretical study of poly(p-phenylenes) and their cyclodextrin-based insulated molecular wires. *Comput. Theor. Chem.* **2021**, *1197*, 113157. [CrossRef]
16. Dehghani, A.; Bahlakeh, G.; Ramezanzadeh, B. Synthesis of a non-hazardous/smart anti-corrosion nano-carrier based on beta-cyclodextrin-zinc acetylacetonate inclusion complex decorated graphene oxide (β-CD-ZnA-MGO). *J. Hazard. Mater.* **2020**, *398*, 122962. [CrossRef] [PubMed]
17. Tomer, A.; Kusema, B.T.; Paul, J.-F.; Przybylski, C.; Monflier, E.; Pera-Titus, M.; Ponchel, A. Cyclodextrin-assisted low-metal Ni-Pd/Al2O3 bimetallic catalysts for the direct amination of aliphatic alcohols. *J. Catal.* **2018**, *368*, 172–189. [CrossRef]
18. Quan, X.; Yi, S.; Wang, X. Theoretical study of an anti-Markovnikov addition reaction catalyzed by β-cyclodextrin. *J. Mol. Model.* **2018**, *24*, 77. [CrossRef]
19. Mazurek, A.H.; Szeleszczuk, Ł.; Pisklak, D.M. Periodic DFT Calculations—Review of Applications in the Pharmaceutical Sciences. *Pharmaceutics* **2020**, *12*, 415. [CrossRef]
20. Van De Streek, J.; Neumann, A.M. Validation of molecular crystal structures from powder diffraction data with dispersion-corrected density functional theory (DFT-D). *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2014**, *B70*, 1020–1032. [CrossRef]
21. Imtiaz, S.; Ali, S.M. Atom accurate structure determination of alprazolam/cyclodextrin inclusion complexes by ROESY and computational approaches. *J. Indian Chem. Soc.* **2022**, *99*, 100299. [CrossRef]
22. Lu, H.; Wang, Y.; Xie, X.; Chen, F.; Li, W. Molecular dynamics simulation and TDDFT study of the structures and UV–vis absorption spectra of MCT-β-CD and its inclusion complexes. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2015**, *149*, 564–570. [CrossRef]
23. Bouhadiba, A.; Rahali, S.; Belhocine, Y.; Allal, H.; Nouar, L.; Rahim, M. Structural and energetic investigation on the host/guest inclusion process of benzyl isothiocyanate into β-cyclodextrin using dispersion-corrected DFT calculations. *Carbohydr. Res.* **2020**, *491*, 107980. [CrossRef]

24. Wang, J.; Zhang, R.; Lu, Z.; Ai, Y. Experimental and theoretical studies of spherical β-cyclodextrin modified titanium dioxide composites for uranium removal. *Ecol. Eng.* **2020**, *149*, 105835. [CrossRef]
25. Jabeen, E.; Janjua, N.K.; Ahmed, S.; Murtaza, I.; Ali, T.; Masood, N.; Rizvi, A.S.; Murtaza, G. DFT predictions, synthesis, stoichiometric structures and anti-diabetic activity of Cu (II) and Fe (III) complexes of quercetin, morin, and primuletin. *J. Mol. Struct.* **2017**, *1150*, 459–468. [CrossRef]
26. Yuan, S.; Shi, W.; Li, B.; Wang, J.; Jiao, H.; Li, Y.-W. Theoretical ONIOM2 Study on Pyridine Adsorption in the Channels and Intersection of ZSM-5. *J. Phys. Chem. A* **2005**, *109*, 2594–2601. [CrossRef]
27. Christensen, A.; Kubař, T.; Cui, Q.; Elstner, M. Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications. *Chem. Rev.* **2016**, *116*, 5301–5337. [CrossRef]
28. Ferrero, R.; Pantaleone, S.; Delle Piane, M.; Caldera, F.; Corno, M.; Trotta, F.; Brunella, V. On the Interactions of Melatonin/β-Cyclodextrin Inclusion Complex: A Novel Approach Combining Efficient Semiempirical Extended Tight-Binding (xTB) Results with Ab Initio Methods. *Molecules* **2021**, *26*, 5881. [CrossRef] [PubMed]
29. Vidal, R.B.P.; Ibañez, G.A.; Escandar, G.M. Spectrofluorimetric study of phenolic endocrine disruptors in cyclodextrin media. *RSC Adv.* **2015**, *5*, 20914–20923. [CrossRef]
30. Li, X.-S.; Liu, L.; Mu, T.-W.; Guo, Q.-X. A Systematic Quantum Chemistry Study on Cyclodextrins. *Mon. Chem.* **2000**, *131*, 849–855. [CrossRef]
31. Fatiha, M.; Khatmi, D.E.; Largate, L. Theoretical approach in the study of the inclusion processes of sulconazole with β-cyclodextrin. *J. Mol. Liq.* **2010**, *154*, 1–5. [CrossRef]
32. Setiadji, S.; Sundari, C.D.D.; Ramdhani, M.A.; Umam, A.B.K.; Ivansyah, A.L. Theoretical Investigation of Inclusion Complex between Omeprazole Enantiomers and Carboxymethyl-β-Cyclodextrin. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *288*, 012138. [CrossRef]
33. Saroj, M.K.; Payal, R.; Jain, S.; Sharma, N.; Rastogi, R.C. Investigation of indole chalcones encapsulation in β-cyclodextrin: Determination of stoichiometry, binding constants and thermodynamic parameters. *J. Incl. Phenom. Macrocycl. Chem.* **2018**, *90*, 306–320. [CrossRef]
34. Nurhidayah, E.S.; Martoprawiro, M.; Zulfikar, M.A. PM3 and ONIOM2 modelling of inclusion complex of ibuprofen enantiomers with dimethyl-β-Cyclodextrin. *J. Chem. Tech. Metall.* **2019**, *54*, 673–678.
35. Srihakulung, O.; Triamchaisri, N.; Toochinda, P.; Lawtrakul, L. Theoretical study on ferrocenyl hydrazones inclusion complexes with β-cyclodextrin and its three methylated derivatives. *J. Incl. Phenom. Macrocycl. Chem.* **2020**, *98*, 79–91. [CrossRef]
36. Arsad, S.R.; Maarof, H.; Ibrahim, W.A.W.; Aboul-Enein, H.Y. Theoretical and Molecular Docking Study of Ketoconazole on Heptakis(2,3,6-tri-*O* -methyl)-β-cyclodextrin as Chiral Selector. *Chirality* **2016**, *28*, 209–214. [CrossRef] [PubMed]
37. Bensouilah, N.; Fisli, H.; Bensouilah, H.; Zaater, G.; Abdaoui, M.; Boutemeur-Kheddis, B. Host-guest complex of N-(2-chloroethyl), N-nitroso,N′,N′-dicyclohexylsulfamid with β-cyclodextrin: Fluorescence, QTAIM analysis and structure-chemical reactivity. *J. Mol. Struct.* **2017**, *1146*, 179–190. [CrossRef]
38. Ivansyah, A.L.; Martoprawiro, M. Buchari Computational modeling of inclusion complex of r/s-omeprazole with β-cyclodextrin using oniom2 method. *J. Phys. Conf. Ser.* **2017**, *812*, 012070. [CrossRef]
39. Yang, Z.; Huang, L.; Yao, X.; Ji, H. Host-guest complexes of estragole with β-cyclodextrin: An experimental and theoretical investigation. *Flavour Fragr. J.* **2017**, *32*, 102–111. [CrossRef]
40. Bensouilah, N.; Boutemeur-Kheddis, B.; Meddour, I.; Abdaoui, M. Host-guest complex of nabumetone: β-cyclodextrin: Quantum chemical study and QTAIM analysis. *J. Incl. Phenom. Macrocycl. Chem.* **2017**, *87*, 191–206. [CrossRef]
41. Reis, V.S.; Santos, E.S.; Bonsolhos, D.N.F.; Guimarães, L.; De Almeida, W.G.; Nascimento, C.S. Theoretical study on the formation process of Cross-Linked β-Cyclodextrin molecular tubes. *Chem. Phys. Lett.* **2017**, *677*, 13–18. [CrossRef]
42. Abdelmalek, L.; Fatiha, M.; Leila, N.; Mouna, C.; Nora, M.; Djameleddine, K. Computational study of inclusion complex formation between carvacrol and β-cyclodextrin in vacuum and in water: Charge transfer, electronic transitions and NBO analysis. *J. Mol. Liq.* **2016**, *224*, 62–71. [CrossRef]
43. Al Azzam, K.M.; Muhammad, E. Host-guest Inclusion Complexes between Mitiglinide and the Naturally Occurring Cyclodextrins α, β, and γ: A Theoretical Approach. *Adv. Pharm. Bull.* **2015**, *5*, 289–291. [CrossRef]
44. Yang, Z.; Yao, X.; Xiao, Z.; Chen, H.; Ji, H. Preparation and release behaviour of the inclusion complexes of phenylethanol withβ-cyclodextrin. *Flavour Fragr. J.* **2016**, *31*, 206–216. [CrossRef]
45. Nazarov, V.B.; Avakyan, V.G.; Bagrii, E.I.; Vershinnikova, T.G.; Alfimov, M.V. Long-lived phosphorescence of arenes in complexes with cyclodextrins 2. Room-temperature phosphorescence of ternary complexes of naphthalene and phenanthrene with β-cyclodextrin and adamantane derivatives in the presence of oxygen. *Russ. Chem. Bull.* **2005**, *54*, 2752–2756. [CrossRef]
46. Bouzit, H.; Stiti, M.; Abdaoui, M. Spectroscopic and molecular modelling investigations of supramolecular complex of β-cyclodextrin with N-[(4-sulfonamidophenyl)ethyl]-5-(1,2-dithiolan-3-yl)pentanamide. *J. Incl. Phenom. Macrocycl. Chem.* **2016**, *86*, 121–134. [CrossRef]
47. Laspidou, C.S.; Archimandritis, A.S.; Papadimitriou, T.; Kormas, K.A.; Yannakopoulou, K.; Lazarou, Y.G. Theoretical investigation of microcystin-LR, microcystin-RR and nodularin-R complexation with α-, β-, and γ-cyclodextrin as a starting point for the targeted design of efficient cyanotoxin traps. *Sustain. Chem. Pharm.* **2016**, *3*, 25–32. [CrossRef]

48. Suliman, F.O.; Elbashir, A.A.; Schmitz, O.J. Study on the separation of ofloxacin enantiomers by hydroxyl-propyl-β-cyclodextrin as a chiral selector in capillary electrophoresis: A computational approach. *J. Incl. Phenom. Macrocycl. Chem.* **2015**, *83*, 119–129. [CrossRef]
49. Mizera, M.; Lewandowska, K.; Miklaszewski, A.; Cielecka-Piontek, J. Machine Learning Approach for Determining the Formation of β-Lactam Antibiotic Complexes with Cyclodextrins Using Multispectral Analysis. *Molecules* **2019**, *24*, 743. [CrossRef] [PubMed]
50. Djilani, I.; Madi, F.; Nouar, L.; Haiahem, S.; Rahim, M.; Khatmi, D.E.; Bouhadiba, A. Theoretical investigation to characterize the inclusion complex of α-lipoic acid and β-cyclodextrin. *Comptes Rendus Chim.* **2015**, *18*, 170–177. [CrossRef]
51. Bouhadiba, A.; Belhocine, Y.; Rahim, M.; Djilani, I.; Nouar, L.; Khatmi, D.E. Host-guest interaction between tyrosine and β-cyclodextrin: Molecular modeling and nuclear studies. *J. Mol. Liq.* **2017**, *233*, 358–363. [CrossRef]
52. Sambrook, M.R.; Vincent, J.C.; Ede, J.A.; Gass, I.A.; Cragg, P.I. Experimental and computational study of the inclusion complexes of β-cyclodextrin with the chemical warfare agent soman (GD) and commonly used simulants. *RSC Adv.* **2017**, *7*, 38069–38076. [CrossRef]
53. Yahia, H.A.; Yahia, O.A.; Khatmi, D.; Belghiche, R.; Bouzitouna, A. Quantum chemical investigations on hydrogen bonding interactions established in the inclusion complex β-cyclodextrin/benzocaine through the DFT, AIM and NBO approaches. *J. Incl. Phenom. Macrocycl. Chem.* **2017**, *89*, 353–365. [CrossRef]
54. Ceborska, M.; Kędra-Królik, K.; Kowalska, A.A.; Koźbiał, M. Comparative study of molecular recognition of folic acid subunits with cyclodextrins. *Carbohydr. Polym.* **2018**, *184*, 47–56. [CrossRef] [PubMed]
55. Appell, M.; Evans, K.O.; Jackson, M.A.; Compton, D.L. Determination of ochratoxin A in grape juice and wine using nanosponge solid phase extraction clean-up and liquid chromatography with fluorescence detection. *J. Liq. Chromatogr. Relat. Technol.* **2018**, *41*, 949–954. [CrossRef]
56. López-Méndez, L.J.; Rojas-Aguirre, Y.; Vázquez-Lima, H.; Cassani, J.; Enríquez, R.G.; Rojo-Domínguez, A.; Guadarrama, P. On the conformational search of a βCD dendritic derivative: NMR and theoretical calculations working together reveal a donut-like amphiphilic structure. *J. Mol. Struct.* **2020**, *1204*, 127535. [CrossRef]
57. Sahra, K.; Dinar, K.; Seridi, A.; Kadri, M. Investigation on the inclusion of diclofenac with β-cyclodextrin: A molecular modeling approach. *Struct. Chem.* **2015**, *26*, 61–69. [CrossRef]
58. Srihakulung, O.; Maezono, R.; Toochinda, P.; Kongprawechnon, W.; Intarapanich, A.; Lawtrakul, A.L. Host-Guest Interactions of Plumbagin with β-Cyclodextrin, Dimethyl-β-Cyclodextrin and Hydroxypropyl-β-Cyclodextrin: Semi-Empirical Quantum Mechanical PM6 and PM7 Methods. *Sci. Pharm.* **2018**, *86*, 20. [CrossRef]
59. Bani-Yaseen, A.W. Computational molecular perspectives on the interaction of propranolol with β-cyclodextrin in solution: Towards the drug-receptor mechanism of interaction. *J. Mol. Liq.* **2017**, *227*, 280–290. [CrossRef]
60. Prabhu, A.A.M.; Fatiha, M.; Leila, N.; Raj, T.A.; Navarro-González, I.; Periago, M.J.; Yáñez-Gascón, M.J.; Pérez-Sánchez, H. Investigation of 3D Contour Map and Intermolecular Interaction of Dopamine with β-Cyclodextrin and 2-Hydroxypropyl-β-cyclodextrin. *J. Solut. Chem.* **2018**, *47*, 409–429. [CrossRef]
61. Iyengar, S.S.; Schlegel, B.H.; Voth, G.A. Atom-Centered Density Matrix Propagation (ADMP): Generalizations Using Bohmian Mechanics. *J. Phys. Chem. A* **2003**, *107*, 7269–7277. [CrossRef]
62. Al-Jaber, A.S.; Bani-Yaseen, A.D. On the encapsulation of Olsalazine by β-cyclodextrin: A DFT-based computational and spectroscopic investigations. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2019**, *214*, 531–536. [CrossRef]
63. Silva, D.A.; Xavier, M.J.; Dutra, J.D.L.; Gimenez, I.F.; Freire, R.O.; da Costa, N.B. Prediction of correct intermolecular interactions in host-guest systems involving cyclodextrins. *J. Mol. Struct.* **2020**, *1205*, 127517. [CrossRef]
64. Lula, I.; Gomes, M.F.; Piló-Veloso, D.; Noronha, A.L.O.; Duarte, H.A.; Santos, R.A.S.; Sinisterra, R.D. Spironolactone and its Complexes with β-cyclodextrin: Modern NMR Characterization and Structural DFTB-SCC Calculations. *J. Incl. Phenom. Macrocycl. Chem.* **2006**, *56*, 293–302. [CrossRef]
65. Lula, I.; Denadai, L.; Resende, J.M.; de Sousa, F.B.; de Lima, G.F.; Pilo-Veloso, D.; Heine, T.; Duarte, H.A.; Santos, R.A.; Sinisterra, R.D. Study of angiotensin-(1–7) vasoactive peptide and its β-cyclodextrin inclusion complexes: Complete sequence-specific NMR assignments and structural studies. *Peptides* **2007**, *28*, 2199–2210. [CrossRef] [PubMed]
66. Lukin, O.; Dolgonos, G.; Leszczynski, J. A comprehensive test of computational approaches for evaluation of cyclodextrin complexes. Self-inclusion in monosubstituted β-cyclodextrins—A case study. *Tetrahedron* **2017**, *73*, 5302–5306. [CrossRef]
67. Schnupf, U.; Momany, F.A. DFT Energy Optimization of a Large Carbohydrate: Cyclomaltohexaicosaose (CA-26). *J. Phys. Chem. B* **2011**, *116*, 6618–6627. [CrossRef] [PubMed]
68. Hanpaibool, C.; Chakcharoensap, T.; Hijikata, Y.; Irle, S.; Wolschann, P.; Kungwan, N.; Pongsawasdi, P.; Ounjai, P.; Rungrotmongkol, T. Theoretical analysis of orientations and tautomerization of genistein in β-cyclodextrin. *J. Mol. Liq.* **2018**, *265*, 16–23. [CrossRef]
69. Yunta, M. Using Molecular Modelling to Study Interactions between Molecules with Biological Activity. *IntechOpen* **2012**. [CrossRef]
70. Lu, L. Can B3LYP be improved by optimization of the proportions of exchange and correlation functionals? *Int. J. Quantum Chem.* **2015**, *115*, 502–509. [CrossRef]
71. Ivanov, P. Performance of some DFT functionals with dispersion on modeling of the translational isomers of a solvent-switchable [2]rotaxane. *J. Mol. Struct.* **2016**, *1107*, 31–38. [CrossRef]

72. Mardirossian, N.; Head-Gordon, M. How Accurate Are the Minnesota Density Functionals for Noncovalent Interactions, Isomerization Energies, Thermochemistry, and Barrier Heights Involving Molecules Composed of Main-Group Elements? *J. Chem. Theory Comput.* **2016**, *12*, 4303–4325. [CrossRef]
73. Valero, R.; Costa, R.; Moreira, I.D.P.R.; Truhlar, D.; Illas, F. Performance of the M06 family of exchange-correlation functionals for predicting magnetic coupling in organic and inorganic molecules. *J. Chem. Phys.* **2008**, *128*, 114103. [CrossRef]
74. Wang, Y.; Jin, X.; Yu, H.S.; Truhlar, D.G.; He, X. Revised M06-L functional for improved accuracy on chemical reaction barrier heights, noncovalent interactions, and solid-state physics. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 8487–8492. [CrossRef] [PubMed]
75. Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924. [CrossRef] [PubMed]
76. Kohn, W.; Sham, L.J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133–A1138. [CrossRef]
77. Bučko, T.; Lebègue, S.; Hafner, J.; Ángyán, J.G. Tkatchenko-Scheffler van der Waals correction method with and without self-consistent screening applied to solids. *Phys. Rev. B* **2013**, *87*, 064110. [CrossRef]
78. Grimme, S.; Steinmetz, M. Effects of London dispersion correction in density functional theory on the structures of organic molecules in the gas phase. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16031–16042. [CrossRef] [PubMed]
79. Xu, P.; Alkan, M.; Gordon, M.S. Many-Body Dispersion. *Chem. Rev.* **2020**, *120*, 12343–12356. [CrossRef]
80. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465. [CrossRef]
81. Yanai, T.; Tew, D.P.; Handy, N.C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57. [CrossRef]
82. Salma, A.; Fatiha, M.; Leila, N. Effect of solvent on absorption and emission spectra of 2,2′-Bipyridine and its inclusion complexinto β-cyclodextrin: DFT and TD-DFT study. *Comput. Theor. Chem.* **2021**, *1206*, 113481. [CrossRef]
83. Azayez, M.; Fergoug, T.; Meddah-Araibi, N.; Zelmat, C.; Bouhadda, Y. Theoretical Investigation of the Complexation Reaction of Procaine-hydrochloride by β-cyclodextrin. *Phys. Chem. Res.* **2020**, *8*, 155–165. [CrossRef]
84. Safia, H.; Ismahan, L.; Abdelkrim, G.; Mouna, C.; Leila, N.; Fatiha, M. Density functional theories study of the interactions between host β-Cyclodextrin and guest 8-Anilinonaphthalene-1-sulfonate: Molecular structure, HOMO, LUMO, NBO, QTAIM and NMR analyses. *J. Mol. Liq.* **2019**, *280*, 218–229. [CrossRef]
85. Buczek, A.; Staś, A.; Hebenstreit, C.; Maller, C.; Broda, M.A.; Kupka, T.; Kelterer, A. Interaction of 5-fluorouracil with β-cyclodextrin: A density functional theory study with dispersion correction. *Int. J. Quantum Chem.* **2021**, *121*, e26487. [CrossRef]
86. Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* **2003**, *24*, 669–681. [CrossRef] [PubMed]
87. Tomasi, J.; Mennucci, B.; Cancès, E. The IEF version of the PCM solvation method: An overview of a new method addressed to study molecular solutes at the QM ab initio level. *J. Mol. Struct.* **1999**, *464*, 211–226. [CrossRef]
88. Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396. [CrossRef]
89. Belhocine, Y.; Rahali, S.; Allal, H.; Assaba, I.M.; Ghoniem, M.G.; Ali, F.A.M. A Dispersion Corrected DFT Investigation of the Inclusion Complexation of Dexamethasone with β-Cyclodextrin and Molecular Docking Study of Its Potential Activity against COVID-19. *Molecules* **2021**, *26*, 7622. [CrossRef]
90. Oqmhula, K.; Hongo, K.; Maezono, R.; Ichibha, T. Ab Initio Evaluation of Complexation Energies for Cyclodextrin-Drug Inclusion Complexes. *ACS Omega* **2020**, *5*, 19371–19376. [CrossRef]
91. Sambrook, M.R.; Gass, I.A.; Cragg, P.J. Spectroscopic and inclusion properties of G-series chemical warfare agents and their simulants: A DFT study. *Supramol. Chem.* **2017**, *30*, 206–217. [CrossRef]
92. Sure, R.; Grimme, S. Comprehensive Benchmark of Association (Free) Energies of Realistic Host–Guest Complexes. *J. Chem. Theory Comput.* **2015**, *11*, 3785–3801. [CrossRef]
93. Lopes, J.F.; Nascimento, C.S.; Anconi, C.P.A.; Santos, H.F.; Almeida, W.B. Inclusion complex thermodynamics: The β-cyclodextrin and sertraline complex example. *J. Mol. Graph. Model.* **2015**, *62*, 11–17. [CrossRef]
94. CCDC, Cambridge Crystallographic Data Centre. Available online: https://www.ccdc.cam.ac.uk/ (accessed on 29 May 2022).
95. Autodock. Available online: https://autodock.scripps.edu/ (accessed on 29 May 2022).
96. Available online: https://www.schrodinger.com/products/maestro (accessed on 29 May 2022).
97. Biovia. Available online: https://www.3ds.com/ (accessed on 29 May 2022).
98. Nora, M.; Ismahan, L.; Abdelkrim, G.; Mouna, C.; Leila, N.; Fatiha, M.; Nada, B.; Brahim, H. Interactions in inclusion complex of β-cyclodextrin/l-Metheonine: DFT computational studies. *J. Incl. Phenom. Macrocycl. Chem.* **2020**, *96*, 43–54. [CrossRef]
99. Pan, A.; Kar, T.; Rakshit, A.K.; Moulik, S.P. Enthalpy–Entropy Compensation (EEC) Effect: Decisive Role of Free Energy. *J. Phys. Chem. B* **2016**, *120*, 10531–10539. [CrossRef] [PubMed]
100. Rahali, S.; Belhocine, Y.; Allal, H.; Bouhadiba, A.; Assaba, I.M.; Seydou, M. A DFT Investigation of The Host-Guest Interactions Between Boron-Based Aromatic Systems and β-Cyclodextrin. *Res. Sq.* **2022**, 1–2. [CrossRef]

101. Morais, C.A.S.; Silva, B.L.; Denadai, A.M.L.; Lopes, J.F.; De Sousa, F.B. Structural and thermodynamic investigation of pentoxifylline-cyclodextrin inclusion complex. *Chem. Phys. Lett.* **2017**, *682*, 43–48. [CrossRef]
102. Aree, T.; Jongrungruangchok, S. β-Cyclodextrin encapsulation elevates antioxidant capacity of tea: A closing chapter on non-epicatechins, atomistic insights from X-ray analysis, DFT calculation and DPPH assay. *Carbohydr. Polym.* **2018**, *194*, 24–33. [CrossRef] [PubMed]
103. Aree, T.; Jongrungruangchok, S. Enhancement of antioxidant activity of green tea epicatechins in β-cyclodextrin cavity: Single-crystal X-ray analysis, DFT calculation and DPPH assay. *Carbohydr. Polym.* **2016**, *20*, 1139–1151. [CrossRef] [PubMed]
104. Ikeda, H.; Ohata, T.; Yukawa, M.; Tsutsumi, H.; Fujisawa, M.; Aki, H. Calculation study on complex formation of catechins with β-cyclodextrin using density function theory. *J. Incl. Phenom. Macrocycl. Chem.* **2021**, *100*, 99–107. [CrossRef]
105. Aree, T. Understanding structures and thermodynamics of β-cyclodextrin encapsulation of chlorogenic, caffeic and quinic acids: Implications for enriching antioxidant capacity and masking bitterness in coffee. *Food Chem.* **2019**, *293*, 550–560. [CrossRef]
106. Aree, T. Inclusion complex of β-cyclodextrin with coffee chlorogenic acid: New insights from a combined crystallographic and theoretical study. *Acta Crystallogr. Sect. C Struct. Chem.* **2019**, *75*, 15–21. [CrossRef]
107. Meryem, G.; Rabah, K.; Fatiha, M.; Leila, N.; Aziz, B.A.; Imane, D.; Rachid, M. Computational investigation of vanillin@βéta-cyclodextrin inclusion complex: Electronic and intermolecular analysis. *J. Mol. Liq.* **2021**, *321*, 114839. [CrossRef]
108. Aree, T. β-Cyclodextrin Inclusion Complexation With Tricyclic Antidepressants Desipramine and Imipramine: A Structural Chemistry Perspective. *J. Pharm. Sci.* **2020**, *109*, 3086–3094. [CrossRef] [PubMed]
109. Aree, T. Inclusion Scenarios and Conformational Flexibility of the SSRI Paroxetine as Perceived from Polymorphism of β-Cyclodextrin–Paroxetine Complex. *Pharmaceuticals* **2022**, *15*, 98. [CrossRef] [PubMed]
110. Keniche, A.; Slimani, M.Z.; Miranda, J.I.; Aizpurua, J.M.; Mulengi, J.K. NMR Investigation of the complexation of (S)-2-isopropyl-1-(o-nitrophenyl)sulfonyl)aziridine with β-cyclodextrin. *Mediterr. J. Chem.* **2013**, *2*, 620–631. [CrossRef]
111. Harati, H.; Morsali, A.; Bozorgmehr, M.R.; Beyramabadi, S.A. β-cyclodextrin-lenalidomide anticancer drug delivery nanosystem: A quantum chemical approach. *J. Mol. Liq.* **2021**, *344*, 117762. [CrossRef]
112. Rohman, M.A.; Phanrang, P.T.; Chamlagai, D.; Mitra, S. Deciphering Spectroscopic and Structural Insights into the Photophysical Behavior of 2,2′-Dipyridylamine: An Efficient Environment Sensitive Fluorescence Probe. *J. Phys. Chem. A* **2021**, *125*, 6964–6975. [CrossRef]
113. Wiergowska, G.; Ludowicz, D.; Wdowiak, K.; Miklaszewski, A.; Lewandowska, K.; Cielecka-Piontek, J. Combinations of Freeze-Dried Amorphous Vardenafil Hydrochloride with Saccharides as a Way to Enhance Dissolution Rate and Permeability. *Pharmaceuticals* **2021**, *14*, 453. [CrossRef]
114. Akhondi, M.; Jamalizadeh, E.; Mohebbi, A. MD and DFT calculations on the structural variations of amino-cyclodextrin as a pH-sensitive carrier for smart carriage and release of Doxorubicin. *J. Mol. Struct.* **2021**, *1230*, 129855. [CrossRef]
115. Mezari, Y.; Nouar, L.; Madi, F.; Guendouzi, A.; Djellala, I.; Lafifi, I.; Merdes, R.; Bouhadiba, A.; Houari, B. Theoretical investigation of inclusion complex of 2-methyl mercapto phenothiazine with hydroxy propyl β-cyclodextrin by DFT approaches. *Bulg. Chem. Comm.* **2021**, *53*, 196–210. [CrossRef]
116. Bani-Yaseen, A.D. The supramolecular host-guest complexation of Vemurafenib with β-cyclodextrin and cucurbit[7]uril as drug photoprotecting systems: A DFT/TD-DFT study. *Comput. Theor. Chem.* **2020**, *1191*, 113026. [CrossRef]
117. Jafari, G.; Raissi, H.; Hashemzadeh, H. Molecular insight into the interaction of fluorometholone and cholesterol molecules with β-cyclodextrin and sulfobutylether-β-cyclodextrin. *Comput. Theor. Chem.* **2022**, *1208*, 113554. [CrossRef]
118. Gamboa-Carballo, J.J.; Ferino-Pérez, A.; Rana, V.K.; Levalois-Grützmacher, J.; Gaspard, S.; Montero-Cabrera, L.A.; Haza, U.J.J. Theoretical Evaluation of the Molecular Inclusion Process between Chlordecone and Cyclodextrins: A New Method for Mitigating the Basis Set Superposition Error in the Case of an Implicit Solvation Model. *J. Chem. Inf. Model.* **2020**, *60*, 2115–2125. [CrossRef] [PubMed]
119. Yang, Y.-Z.; Liu, X.-F.; Zhang, R.-B.; Pang, S.-P. Joint experimental and theoretical studies of the surprising stability of the aryl pentazole upon noncovalent binding to β-cyclodextrin. *Phys. Chem. Chem. Phys.* **2017**, *19*, 31236–31244. [CrossRef] [PubMed]
120. Pereira, R.A.; Borges, W.M.D.S.; Peraro, C.R.; Anconi, C.P.A. Theoretical inclusion of deprotonated 2,4-D and dicamba pesticides in ß-cyclodextrin. *J. Incl. Phenom. Macrocycl. Chem.* **2016**, *86*, 343–349. [CrossRef]
121. Li, Z.; Couzijn, E.P.A.; Zhang, X. Intrinsic Properties of α-Cyclodextrin Complexes with Benzoate Derivatives in the Gas Phase: An Experimental and Theoretical Study. *J. Phys. Chem. B* **2012**, *116*, 943–950. [CrossRef] [PubMed]
122. Yao, L.; Mori, Y.; Takano, K. Theoretical Study on Intermolecular Interactions in Complexes of Cyclodextrins with Bile Acids: DFT and Ab Initio Fragment Molecular Orbital Calculations. *Bull. Chem. Soc. Jpn.* **2014**, *87*, 258–266. [CrossRef]
123. Li, Z.; Couzijn, E.P.A.; Zhang, X. A quantitative study of intrinsic non-covalent interactions within complexes of α-cyclodextrin and benzoate derivatives. *Chem. Commun.* **2012**, *48*, 9864–9866. [CrossRef]
124. Muzaffar, S.; Imtiaz, S.; Ali, S.M. Demonstrating accuracy of the proposed protocol for structure elucidation of cyclodextrin inclusion complexes by validation using DFT studies. *J. Mol. Struct.* **2020**, *1217*, 128419. [CrossRef]
125. Li, N.; Yang, L.; Ji, X.; Ren, J.; Gao, B.; Deng, W.-Q.; Wang, Z. Bioinspired succinyl-β-cyclodextrin membranes for enhanced uranium extraction and reclamation. *Environ. Sci. Nano* **2020**, *7*, 3124–3135. [CrossRef]
126. Ignaczak, A.; Orszański, Ł.; Adamiak, M.; Olejniczak, A.B. Comparative DFT study of inclusion complexes of thymidine-carborane conjugate with β-cyclodextrin and heptakis(2,6-O-dimethyl)-β-cyclodextrin in water. *J. Mol. Liq.* **2020**, *315*, 113767. [CrossRef]

127. Ramos, M.L.; Dias, D.C.; Justino, L.L.G.; Verissimo, L.M.P.; Valente, A.J.M.; Esteso, M.A.; Ribeiro, A.C.F.; Leaist, D.G.; Pina, J.; Cabral, A.; et al. Interactions between glycyl-L-phenylalanine and β-cyclodextrin from diffusion, spectroscopic and computational studies. *J. Mol. Liq.* **2020**, *315*, 113704. [CrossRef]
128. Deosarkar, S.D.; Sawale, R.T.; Pinjari, R.V.; Kalyankar, T.M. Interactions of sodium salicylate and β-cyclodextrin in water: A volumetric, ultraacoustic and optical study. *J. Mol. Liq.* **2020**, *130*, 113151. [CrossRef]
129. Okuda, M.; Hiramatsu, T.; Yasuda, M.; Ishigaki, M.; Ozaki, Y.; Hayashi, M.; Tominaga, K.; Chatani, E. Theoretical Modeling of Electronic Structures of Polyiodide Species Included in α-Cyclodextrin. *J. Phys. Chem. B* **2020**, *124*, 4089–4096. [CrossRef] [PubMed]
130. Majhi, K.; Bandyopadhyay, P.; Khatun, R.; Sinha, S. Prediction of the most preferable rotamer of meta-aminophenol in β-cyclodextrin cavity in aqueous medium by using spectroscopic and DFT computational studies. *J. Incl. Phenom. Macrocycl. Chem.* **2020**, *97*, 77–86. [CrossRef]
131. Ismahan, L.; Leila, N.; Fatiha, M.; Abdelkrim, G.; Mouna, C.; Nada, B.; Brahim, H. Computational study of inclusion complex of l-Glutamine/beta-Cycldextrin: Electronic and intermolecular interactions investigations. *J. Mol. Struct.* **2020**, *1206*, 127740. [CrossRef]
132. Pereva, S.; Nikolova, V.; Sarafska, T.; Angelova, S.; Spassov, T.; Dudev, T. Inclusion complexes of ibuprofen and β-cyclodextrin: Supramolecular structure and stability. *J. Mol. Struct.* **2020**, *1205*, 127575. [CrossRef]
133. Stoicescu, C.S.; Neacşu, A.D.; Bădiceanu, C.D.; Munteanu, G. Inclusion complexes of some thiourea derivatives in cyclodextrins. *J. Incl. Phenom. Macrocycl. Chem.* **2019**, *96*, 275–283. [CrossRef]
134. Paulino, P.H.S.; Sousaab, S.M.R.; Da Silva, H.C.; De Almeida, W.B.; Ferrari, J.L.; Guimarãesa, L.; Nascimento, C.S., Jr. A theoretical investigation on the encapsulation process of mepivacaine into β-cyclodextrin. *Chem. Phys. Lett.* **2020**, *740*, 137060. [CrossRef]
135. Wu, J.; Ma, H.; Bu, X.; Zhu, L.; Hao, B.; Zhao, B.; Tian, Y. SERS determination of the antihypertensive drugs prazosin and losartan by using silver nanoparticles coated with β-cyclodextrin. *Mikrochim. Acta* **2019**, *186*, 801. [CrossRef]
136. Bezzina, B.; Djemil, R.; Bensouilah, N. Quantitative and qualitative analyses of intermolecular interactions in neutral/deprotonated aspirin@β-CD inclusion complexes: QTAIM and NBO analyses. *Theor. Chim. Acta* **2019**, *138*, 43. [CrossRef]
137. Wójcik, J.; Ejchart, A.; Nowakowski, M. Shape adaptation of quinine in cyclodextrin cavities: NMR studies. *Phys. Chem. Chem. Phys.* **2019**, *21*, 6925–6934. [CrossRef]
138. Bakirhan, N.K.; Tok, T.T.; Ozkan, S.A. The redox mechanism investigation of non-small cell lung cancer drug: Erlotinib via theoretical and experimental techniques and its host–guest detection by β-Cyclodextrin nanoparticles modified glassy carbon electrode. *Sens. Actuators B Chem.* **2019**, *278*, 172–180. [CrossRef]
139. Li, L.; Zhou, Y.; Wang, Z.; Wu, C.; Li, Z.; Sun, C.; Sun, T. Theoretical studies on the mechanism of sugammadex for the reversal of aminosteroid-induced neuromuscular blockade. *J. Mol. Liq.* **2018**, *265*, 450–456. [CrossRef]
140. Khavani, M.; Kalantarinezhad, R.; Izadyar, M. A joint QM/MD study on α-, β- and γ-cyclodextrins in selective complexation with cathinone. *Supramol. Chem.* **2017**, *30*, 687–696. [CrossRef]
141. Eftaiha, A.F.; Qaroush, A.K.; Alsoubani, F.; Pehl, T.M.; Troll, C.; Rieger, B.; Al-Maythalony, B.A.; Assaf, K.I. A green sorbent for CO_2 capture: α-cyclodextrin-based carbonate in DMSO solution. *RSC Adv.* **2018**, *8*, 37757–37764. [CrossRef]
142. Belhocine, Y.; Bouhadiba, A.; Rahim, M.; Nouar, L.; Djilani, I.; Khatmi, D.I. Inclusion Complex Formation of β-Cyclodextrin with the Nonsteroidal Anti-inflammatory Drug Flufenamic Acid: Computational Study. *Macroheterocycles* **2018**, *11*, 203–209. [CrossRef]
143. Sierpe, R.; Noyong, M.; Simon, U.; Aguayo, D.; Huerta, J.; Kogan, M.J.; Yutronic, N. Construction of 6-thioguanine and 6-mercaptopurine carriers based on βcyclodextrins and gold nanoparticles. *Carbohydr. Polym.* **2017**, *177*, 22–31. [CrossRef] [PubMed]
144. Anconi, C.P.A.; Santos, T.M.R.; Souza, A.C.; Borges, W.M.S.; Sales, A.L.R. Host–guest intermolecular hydrogen bonds and stability in aqueous media: The benzaldehyde/β-CD case study. *J. Incl. Phenom. Macrocycl. Chem.* **2017**, *89*, 137–142. [CrossRef]
145. Cao, B.; Du, J.; Cao, Z.; Sun, X.; Sun, H.; Fu, H. DFT study on the dissolution mechanisms of α-cyclodextrin and chitobiose in ionic liquid. *Carbohydr. Polym.* **2017**, *169*, 227–235. [CrossRef]
146. Angelova, S.; Nikolova, V.; Molla, N.; Dudev, T. Factors Governing the Host–Guest Interactions between IIA/IIB Group Metal Cations and α-Cyclodextrin: A DFT/CDM Study. *Inorg. Chem.* **2017**, *56*, 1981–1987. [CrossRef]
147. Oliveri, V.; Pietropaolo, A.; Sgarlata, C.; Vecchio, G. Zinc Complexes of Cyclodextrin-bearing 8-Hydroxyquinoline Ligands: A Comparative Study. *Chem. Asian J.* **2016**, *12*, 110–115. [CrossRef]
148. Cheng, Y.; Wang, X.; Chang, D.; Li, W. DFT study on the effects of catalysis by β-cyclodextrin in the reaction of p-nitrophenyl acetate. *J. Mol. Model.* **2017**, *23*, 21. [CrossRef] [PubMed]
149. Maia, P.P.; de Sousa, S.M.R.; De Almeida, W.B.; Guimaraes, L.; Nascimento, C.S. Computational investigation on the host–guest inclusion process of norfloxacin into β-cyclodextrin. *J. Mol. Model.* **2016**, *22*, 220. [CrossRef] [PubMed]
150. Ignaczak, A.; Orszański, Ł. In Search of the Most Stable Molecular Configuration of Heptakis(2,6-O-dimethyl)-β-cyclodextrin and Its Complex with Mianserin: A Comparison of the B3LYP-GD2 and M062X-GD3 Results. *J. Phys. Chem. B* **2021**, *125*, 13077–13087. [CrossRef] [PubMed]
151. Aree, T. Advancing insights on β-cyclodextrin inclusion complexes with SSRIs through lens of X-ray diffraction and DFT calculation. *Int. J. Pharm.* **2021**, *609*, 121113. [CrossRef] [PubMed]

152. Aree, T. Distinctive Supramolecular Features of β-Cyclodextrin Inclusion Complexes with Antidepressants Protriptyline and Maprotiline: A Comprehensive Structural Investigation. *Pharmaceuticals* **2021**, *14*, 812. [CrossRef]
153. Aree, T. Supramolecular Complexes of β-Cyclodextrin with Clomipramine and Doxepin: Effect of the Ring Substituent and Component of Drugs on Their Inclusion Topologies and Structural Flexibilities. *Pharmaceuticals* **2020**, *13*, 278. [CrossRef]
154. Aree, T. β-Cyclodextrin encapsulation of nortriptyline HCl and amitriptyline HCl: Molecular insights from single-crystal X-ray diffraction and DFT calculation. *Int. J. Pharm.* **2020**, *575*, 118899. [CrossRef]
155. Rayene, K.; Imane, D.; Abdelaziz, B.; Leila, N.; Fatiha, M.; Abdelkrim, G.; Bouzid, G.; Ismahan, L.; Brahim, H.; Rabah, O. Molecular modeling study of structures, Hirschfield surface, NBO, AIM, RDG, IGM and 1HNMR of thymoquinone/hydroxypropyl-β-cyclodextrin inclusion complex from QM calculations. *J. Mol. Struct.* **2022**, *1249*, 131565. [CrossRef]
156. Macernis, M.; Bockuviene, A.; Gruskiene, R.; Krivorotova, T.; Sereikaite, J. Raman study for β-ring positioning in β-Carotene complexes with Cyclodextrins and Chitooligosaccharides. *J. Mol. Struct.* **2021**, *1226*, 129362. [CrossRef]
157. Kerdpol, K.; Daengngern, R.; Sattayanon, C.; Namuangruk, S.; Rungrotmongkol, T.; Wolschann, P.; Kungwan, N.; Hannongbua, S. Effect of Water Microsolvation on the Excited-State Proton Transfer of 3-Hydroxyflavone Enclosed in γ-Cyclodextrin. *Molecules* **2021**, *26*, 843. [CrossRef]
158. Bouchemela, H.; Madi, F.; Nouar, L. DFT investigation of host–guest interactions between α-Terpineol and β-cyclodextrin. *J. Incl. Phenom. Macrocycl. Chem.* **2019**, *95*, 247–258. [CrossRef]
159. Seridi, L.; Boufelfel, A. Naringenin encapsulation in β-CD and in heptakis(2,6-di-O-methyl)-β-CD:NMR, NBO and QTAIM analysis. *J. Incl. Phenom. Macrocycl. Chem.* **2018**, *90*, 287–304. [CrossRef]
160. Nutho, B.; Nunthaboot, N.; Wolschann, P.; Kungwan, N.; Rungrotmongkol, T. Metadynamics supports molecular dynamics simulation-based binding affinities of eucalyptol and beta-cyclodextrin inclusion complexes. *RSC Adv.* **2017**, *7*, 50899–50911. [CrossRef]
161. Nutho, B.; Khuntawee, W.; Rungnim, C.; Pongsawasdi, P.; Wolschann, P.; Karpfen, A.; Kungwan, N.; Rungrotmongkol, T. Binding mode and free energy prediction of fisetin/β-cyclodextrin inclusion complexes. *Beilstein J. Org. Chem.* **2014**, *10*, 2789–2799. [CrossRef] [PubMed]
162. Guendouzi, O.; Guendouzi, A.; Ouici, H.B.; Brahim, H.; Boumediene, M.; Elkeurti, M. A quantum chemical study of encapsulation and stabilization of gallic acid in β-cyclodextrin as a drug delivery system. *Can. J. Chem.* **2020**, *98*, 204–214. [CrossRef]
163. Yang, L.; Li, D.; Guo, B.; Wei, D. Theoretical Study on the Inclusion Interaction of β-Cyclodextrin with Gabapentin and Its Stability. *J. Struct. Chem.* **2019**, *60*, 564–574. [CrossRef]
164. Asztemborska, M.; Ceborska, M.; Pietrzak, M. Complexation of tropane alkaloids by cyclodextrins. *Carbohydr. Polym.* **2020**, *209*, 74–81. [CrossRef]
165. Abdel-Mottaleb, M.S.A.; Hamed, E.; Saif, M.; Hafez, H.S. Binding, and thermodynamics of β-cyclodextrin inclusion complexes with some coumarin laser dyes and coumarin-based enzyme substrates: A simulation study. *J. Incl. Phenom. Macrocycl. Chem.* **2018**, *92*, 319–327. [CrossRef]
166. Diamantis, D.A.; Ramesova, S.; Chatzigiannis, C.M.; Degano, I.; Gerogianni, P.S.; Karadima, K.E.; Perikleous, S.; Rekkas, D.; Gerothanassis, I.P.; Galaris, D.; et al. Exploring the oxidation and iron binding profile of a cyclodextrin encapsulated quercetin complex unveiled a controlled complex dissociation through a chemical stimulus. *Biochim. Biophys. Acta (BBA) Gen. Subj.* **2018**, *1862*, 1913–1924. [CrossRef]
167. Guendouzi, A.; Mekelleche, S.M.; Brahim, H.; Litim, K. Quantitative conformational stability host-guest complex of Carvacrol and Thymol with β-cyclodextrin: A theoretical investigation. *J. Incl. Phenom. Macrocycl. Chem.* **2017**, *89*, 143–155. [CrossRef]
168. Abdelaali, M.; Fatiha, M.; Leila, N.; Nora, M.; Mouna, C.; Sakina, H.; Eddine, K.D. Computational approach in the study of the inclusion processes of Thymol with β-cyclodextrin. *J. Mol. Liq.* **2017**, *242*, 714–721. [CrossRef]
169. Aree, T. β-Cyclodextrin Inclusion Complexes with Catechol-Containing Antioxidants Protocatechuic Aldehyde and Protocatechuic Acid—An Atomistic Perspective on Structural and Thermodynamic Stabilities. *Molecules* **2021**, *26*, 3574. [CrossRef] [PubMed]
170. Aree, T.; Jongrungruangchok, S. Structure–antioxidant activity relationship of β-cyclodextrin inclusion complexes with olive tyrosol, hydroxytyrosol and oleuropein: Deep insights from X-ray analysis, DFT calculation and DPPH assay. *Carbohydr. Polym.* **2018**, *199*, 661–669. [CrossRef] [PubMed]
171. Huang, F.; Zhuang, S.; Liu, W.; Lin, L.; Sun, L. Computational investigation on the chiral differentiation of D- and L-penicillamine by β-cyclodextrin. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2020**, *248*, 119277. [CrossRef]
172. Yang, S.; Wu, F.; Yu, F.; Gu, L.; Wang, H.; Liu, Y.; Chu, Y.; Wang, F.; Fang, X.; Ding, C.-F. Distinction of chiral penicillamine using metal-ion coupled cyclodextrin complex as chiral selector by trapped ion mobility-mass spectrometry and a structure investigation of the complexes. *Anal. Chim. Acta* **2021**, *1184*, 339017. [CrossRef]
173. Știufiuc, G.F.; Toma, V.; Onaciu, A.; Chiș, V.; Lucaciu, C.M.; Știufiuc, R.I. Proving Nanoscale Chiral Interactions of Cyclodextrins and Propranolol Enantiomers by Means of SERS Measurements Performed on a Solid Plasmonic Substrate. *Pharmaceutics* **2021**, *13*, 1594. [CrossRef]

174. Lee, S.-S.; Lee, J.-U.; Oh, J.H.; Park, S.; Hong, Y.; Min, B.K.; Lee, H.H.L.; Kim, H.I.; Kong, X.; Lee, S.; et al. Chiral differentiation of d- and l-isoleucine using permethylated β-cyclodextrin: Infrared multiple photon dissociation spectroscopy, ion-mobility mass spectrometry, and DFT calculations. *Phys. Chem. Chem. Phys.* **2018**, *20*, 30428–30436. [CrossRef]
175. Lee, S.; Park, S.; Hong, Y.; Lee, J.U.; Kim, J.H.; Yoon, D.; Kong, X.; Lee, S.; Bin Oh, H. Chiral differentiation of d- and l-alanine by permethylated β-cyclodextrin: IRMPD spectroscopy and DFT methods. *Phys. Chem. Chem. Phys.* **2017**, *19*, 14729–14737. [CrossRef]
176. Reyes-Reyes, M.L.; Roa-Morales, G.; Melgar-Fernández, R.; Reyes-Pérez, H.; Gómez-Oliván, L.M.; Gonzalez-Rivas, N.; Bautista-Renedo, J.; Balderas-Hernández, P. Chiral recognition of abacavir enantiomers by (2-hydroxy)propyl-β-cyclodextrin: UHPLC, NMR and DFT studies. *J. Incl. Phenom. Macrocycl. Chem.* **2015**, *82*, 373–382. [CrossRef]

Journal of Molecular Structure xxx (xxxx) 138710

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: www.elsevier.com/locate/molstruc

17-β-Estradiol—β-Cyclodextrin complex as an aqueous solution: Structural and physicochemical characterization supported by MM and QM calculations

Anna Helena Mazurek [a, b], Łukasz Szeleszczuk [a, *], Kostas Bethanis [c], Elias Christoforides [c], Marta Katarzyna Dudek [d], Ewelina Wielgus [d], Dariusz Maciej Pisklak [a]

[a] Department of Organic and Physical Chemistry, Faculty of Pharmacy, Medical University of Warsaw, Banacha 1 Str., 02-093 Warsaw, Poland
[b] Doctoral School, Medical University of Warsaw, Żwirki i Wigury 81 Str., 02-093 Warsaw, Poland
[c] Laboratory of Physics, Department of Biotechnology, Agricultural University of Athens, 11855 Athens, Greece
[d] Structural Studies Department, Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, Sienkiewicza 112 Str., 90-363 Łódź, Poland

ARTICLE INFO

Keywords:
Estradiol
Cyclodextrin
DFT calculations
Complex stability constant

ABSTRACT

17-β-estradiol (EST) is an Active Pharmaceutical Ingredient characterized by a low water solubility. Complexation with β-cyclodextrin (βCD) enhances its bioavailability, hence such complex is an interesting research object from pharmaceutical point of view. However, basic facts like description of complex's structure and definition of its molar ratio, were debatable already for decades. This work for the first time justifies the EST:βCD molar ratio as 1:2 using the HRMS (high-resolution mass spectrometry) and phase solubility studies. The latter are used to define complex stability constant, as well. The structure and stability is analyzed using a variety of computational approaches: Quantum Mechanics (QM) based methods (DFT, semiempirical approaches) and MD/MMGBSA approach. In case of the QM, for the first time in the computational analysis of cyclodextrin complexes, a thorough benchmarking test is presented. Different computational parameters (solvent model, presence/absence of dispersion correction etc.) are used. Obtained results are compared with the experimental data.

1. Introduction

Cyclodextrins (CDs) are oligosaccharides of a donut-like structure, which enables them to form inclusion complexes with non-polar substances. This characteristic is used by the pharmaceutical industry but at the same time it often poses a non-trivial questions about the structure and stability of the created complexes.

More precisely, CDs are cyclic structures composed of glucose subunits, joined by α−1,4 glycosidic bonds [1]. Because of the cyclic character of CDs, various chemical compounds can enter CD's void and this way inclusion complexes are created. Depending on the size of a chemical guest, different types of CDs are preferred. However, the most common one is a medium size beta-CD (βCD) which consists of 7 glucose units [1,2] (Fig. 1).

Due to the presence of hydroxyl groups, the external fragments of CDs are polar. When a non-polar substance enters the molecular hole of a CD, the formed host–guest complex is polar and more water soluble than a separate non-complexed guest molecule [2,3]. Therefore, knowing also that CDs are non-toxic for a human organism [4], CDs are commonly used in the pharmaceutical industry in order to increase the solubility of a complexed Active Pharmaceutical Ingredient (API) or protect it from external factors like light, humidity or heat [5-7]. An important API group characterized by a poor solubility in water are hormonal steroids. Encapsulation in CDs enhance their solubility in water and as a result also their bioavailability.

An example of the steroid hormones used as a medication is estradiol (EST) (Fig. 1). EST belongs to the estrogens group and is the most potent estrogen naturally produced by a human body [8]. Therefore, the first complexes between EST and a well-known βCD have been obtained already decades ago [9]. However, the structural analysis of such complexes happens not to be as straightforward as one would assume.

In 1997 in the article entitled ''Fluorometric Determination of Association Constants of Three Estrogens with Cyclodextrins'' the EST-βCD complex molar ratio has been defined as 1:1 [10]. This information has been used as a reference for instance in the following articles [11,12] by

* Corresponding author.
E-mail address: lukasz.szeleszczuk@wum.edu.pl (Ł. Szeleszczuk).

https://doi.org/10.1016/j.molstruc.2024.138710
Received 19 February 2024; Received in revised form 3 May 2024; Accepted 21 May 2024
0022-2860/

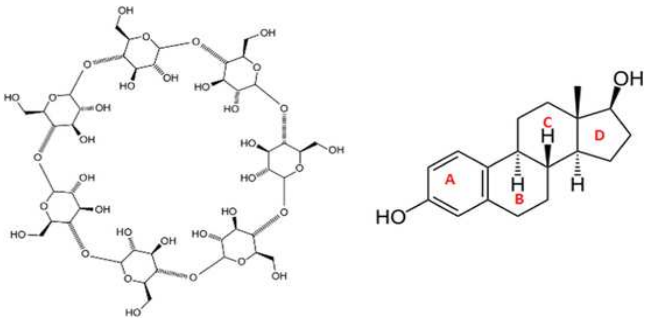


Fig. 1. Structures of β-CD and EST.

other scientists. Only two decades afterwards, this pre-defined EST-βCD complex molar ratio and its stability constant have been revisited [13,14]. However, again only either fluorescence or UV spectroscopy have been used for that purpose. Despite quite a few studies in this area, the 1:1 molar ratio of the complex in question has been accepted as a status quo and has never been questioned or verified by application of any other experimental methods such as high-resolution mass spectrometry (HRMS).

Recently, we have analyzed the EST-βCD complex in a solid state. For the first time, we have obtained and analyzed the structure of this complex's crystal structure. At the same time, it was one of the first ever obtained crystal structures of a steroid hormone complexed with any CD. The results have been published as [15] and the crystal structure has been deposited in the Cambridge Crystallographic Data Centre (CCDC). The results of the structural analysis in a solid state clearly indicate the 1:2 EST-βCD molar ratio, however they have also revealed significant structural disorder of this complex. Having obtained such results for a solid state, we were curious what is the molar ratio in the aqueous solution. Knowing the solid state structure of the complex, its 1:2 molar ratio after dissolution in water seems to be a scientifically sound hypothesis. Nevertheless, as we have gathered data on the previously performed experimental analyses repeatedly pointing out the 1:1 molar ratio, we decided against performing fluorescence or UV measurements and decided to concentrate on the molecular modelling techniques and HRMS spectrometry.

Complexes of various compounds with CDs have been analyzed using different molecular modelling approaches. At the beginning, computational methods used to predict the CD complexes structures and properties were the same as ones used for the analysis of much larger systems like receptors with ligands. Those were mainly Molecular Mechanics (MM) based methods and among them the most popular was molecular docking, sometimes followed by atomistic molecular dynamics (MD) simulations, at the same level of theory. However, in the recent years, with the increase of the computational power, such medium size systems like CD complexes started to be commonly analyzed using techniques based on the Quantum Mechanics (QM).

The conclusions of our recent review article [16] clearly show that the QM calculations era in the analysis of the CD complexes has already started more than a decade ago. The results show that the most commonly used are semi-empirical and Density Functional Theory (DFT) approaches. Among the former, PM6 and PM7 seem to be the most widely spread, also suggesting that they deliver the most appreciated results. When it comes to the DFT, a wider variety of computational options is used. The most frequent is application of B3LYP or M062X functionals both of the coming with or without dispersion correction. In terms of a solvent representation both possible approaches are practiced: either no solvent model is used or an implicit solvent model is applied.

As opposed to QM calculations, the Molecular Mechanics (MM) approach allows analysis of much larger systems but at the same time it delivers results of a significantly lower accuracy. However, for years CD complexes were too big for conducting QM calculations and Molecular Dynamics (MD) was the only option to obtain any information about the complexes' inner structure. Therefore, in the literature there are numerous examples of MD's application in the analysis of CD inclusion complexes. Even now this method is still in use. Therefore, we wanted to apply this well-described type of simulations in our study and compare them with the QM approaches.

Having gathered all the data on the previously conducted experimental analyses of the EST-βCD complex and having in mind our results concerning the solid state, we decided to try to determine the structure and the molar ratio of this complex. The aim of this study was to perform a benchmark studies on this topic using semi-empirical and DFT computational approaches. Additionally, we have managed to experimentally reveal the true complex's molecular ratio using a technique which has been never used in this particular case.

2. Materials and methods

2.1. Sample preparation

The EST-βCD complex was obtained by a method which is commonly used to obtain the CD inclusion complexes, a slow-cooling crystallization technique. This approach has already been used by us in a previous study describing the SCXRD analysis [15]. 60 mg of βCD was mixed in a flask with 1 mL distilled water and put into 70 °C water for 20 s to obtain a clear solution. Then the contents of the flask were poured into a beaker. In accordance with the molar mass of βCD and EST, the respective amount of EST was added to the beaker to maintain the 1:1 molar ratio. The beaker was put on a magnetic stirrer and left at room temperature for 15–20 min until a clear solution was obtained. Afterwards, the contents of the beaker were poured into a glass tube. The beaker was poured along with 0.5–1.0 mL water, which was also added to the glass tube. The glass tube was held in 70 °C water for 20 s to obtain a clear solution. Later, the tube was closed and put into a 70 °C water bath. A slow, gradual cooling process was performed over 10 days, reaching a temperature of 24 °C on the 10th day. At the end, a rotary evaporator was used.

2.2. HRMS

HRMS measurements were performed using Synapt G2-Si mass spectrometer (Waters) equipped with an ESI source and quadrupole-time-of-flight mass analyser. The mass spectrometer was operated in the positive ion detection mode. The optimized source parameters were: capillary voltage 3.0 kV, cone voltage 50 V, source temperature 110 °C, desolvation gas (nitrogen) flow rate 650 L/h with the temperature 450 °C, nebulizer gas pressure 6.5 bar. All samples were dissolved in water–methanol solution (1:1) and infused through a standard electrospray ion source into the instrument. The scan range was *m/z* 500–4000 and the acquisition method run time was 2 min. Mass calibration was performed using a cesium iodide solution. To ensure accurate mass measurements, data were collected in centroid mode and mass was corrected during acquisition using leucine enkephalin solution as an external reference (Lock-SprayTM) which generated reference ion at *m/z* 556.2771 Da ($[M + H]+$) in positive ESI mode. The results of the measurements were processed using the MassLynx 4.1 software (Waters) incorporated with the instrument.

2.3. Phase solubility studies

The UV–visible (UV–Vis) spectrophotometer (BioBase BK-S380, China) was utilized to assess the properties of EST and its inclusion complex. The EST showed a visible absorption peak at 280 nm. To generate a calibration curve, five standard solutions of EST in methanol

were measured, each replicated three times. The concentrations of the standards used were 0.1, 0.25, 0.50, 0.75, and 1.00 mM.

Phase solubility studies were carried out following the procedure outlined by Higuchi & Connors (1965) [17]. An excess amount of EST (50 mg) was added to 10 mL of deionized water containing various concentrations ranging from 0.20 to 20.00 mM for β-CD. The mixtures were then subjected to agitation using an orbital shaker (PHOENIX Instrument Laboratory Shaker RS-OS 5; Berlin, Germany) at 25 °C for 48 h to reach equilibrium. Afterwards, the collected samples were filtered through a 0.45 μm filter and assayed using a UV–Vis spectrophotometer at 280 nm.

2.4. QM calculations

All of the QM calculations were performed using the Gaussian 16 software [18]. All electron DFT computations were done employing the 6–311G(d,p) basis set and B3LYP or M062X functional, while the semi-empirical calculations have been done using PM6 and PM7 approaches. Those four methods were used either with or without Grimme's dispersion force corrections (D3). All of the calculations were performed either in vacuo or using one of the implicit solvation models: Polarizable Continuum Model (PCM) [19] or SMD (Solvation Model Density) [20], each time choosing water as the solvent with dielectric constant 78.540. For the details of computational models, please see Table 1 and Table 2. At the review stage, additional calculations have been performed at the ωB97X-D/6–31G(d,p)-PCM-Water level [21].

Vibrational frequencies were calculated to estimate thermodynamic parameters, including Zero Point Vibrational Energy (ZPVE) and Gibbs free energy (ΔG) at 298.15 K and 101.325 kPa.

According to our recent review, different types of QM approaches are commonly used when modeling CDs inclusion complexes. In this work we decided to use those methods which are the most commonly encountered in the recent literature. The goal was to compare the geometrically optimized 1:2 and 1:1 systems from a quantitative and qualitative perspective. For this reason, we have analyzed the structural aspects as well as values of the energy ΔE and Gibbs free energy ΔG defined as

$$\Delta E = E_{com} - (E_{EST} + nE_{CD}) \quad (1)$$

where E_{com} is energy of the EST-βCD complex, E_{EST} is the energy of EST, E_{CD} is the energy of βCD and n is the number of βCD molecules forming the complex, in this case $n = 1$ or $n = 2$

$$\Delta G = G_{com} - (G_{EST} + nG) \quad (2)$$

where G_{com} is free enthalpy of the EST-βCD complex, G_{EST} is the free enthalpy of EST, G_{CD} is the free enthalpy of βCD and n is the number of βCD molecules forming the complex, in this case $n = 1$ or $n = 2$

2.5. Molecular dynamics (MD) simulations

In addition to the crystallographically determined structure of the EST/β-CD inclusion complex, which is characterized by a host: guest ratio of 2:1 and a head-to-head inclusion mode, two additional models of the same complex assuming host: guest ratio of 1:1 but two different inclusion modes (down and up) were investigated through MD simulations. The initial coordinates for these latter two complexes were generated using molecular docking with Autodock Vina [22], where the simulation boxes were defined around the coordinates of the CD centers with size 40 Å in each direction and a grid spacing of 0.375 Å. The Lamarckian genetic algorithm in AutoDockTools was utilized for this purpose, enabling effective management of numerous degrees of freedom. The 3D structure of estradiol was retrieved from the PubChem database (Compound CID: 5757) [23], while the coordinates of β-CD correspond to its crystal structure in complex with EST [15]. The docking runs were set to 10 and the produced models with the most favorable binding energies for each one of the two inclusion modes were chosen for subsequent analysis.

The AMBER 12 software package [24] was used for the simulation of the EST/β-CD inclusion complexes in a aqueous environment. Three simulations were performed. In the first case, the stating 3D model was provided by the crystallographically determined atomic coordinates of a β-CD dimer[15], which includes one EST guest molecule (site A) inside the formed dimeric cavity. Thus, the host: guest stoichiometry of the entire system in the simulation was 2: 1. In the other two cases, monomers of EST/β-CD inclusion complexes with different inclusion modes from docking analysis were used as the starting models.

The geometry of EST was optimised following the AM1BCC methodology with the program Antechamber [25]. xLeaP, the GUI version of AMBER's LeaP program, was utilized for system preparation. The GLYCAM-06j [26] force field, which is suitable for β-CD atoms' treatment and the generalized AMBER (GAFF) (for the guest molecule) were applied for the simulation. Additionally, the TIP3P water model [27] was used to solvate the CD dimer in a periodic, octahedral box forming a 12 Å thick water shell around the structure.

Minimization and MD calculations that resulted in a single trajectory of the hydrated inclusion complex system were performed with Sander. The particle mesh Ewald summation approach [28] was followed in order to handle the long-range electrostatic interactions with a 10 Å cut-off limit for the direct space sum. Hydrogen bonds were handled using the SHAKE algorithm [29]. The simulation protocol was as following: Energy minimization for hydrogens and waters using 1000 steps of steepest descent (SD) followed by 500 steps of conjugated gradient (CG) methods, while the rest non-hydrogen atoms were fixed with positional restraints of 50 kcal mol 1 Å 2. Heating equilibration up to 300 K of the water in the canonical (NVT) ensemble for 50 ps using positional restraints and the Berendsen thermostat algorithm with coupling constants of 0.5 ps to control temperature and pressure. Energy minimization of all system atoms with weak positional restraints (10 kcal mol 1 Å 2), gradual temperature increase from 5 to 300 K with 10 kcal mol-1 Å-2 restraints on the atoms of the system followed by

Table 1
DFT computational approaches.

1	2	3	4	5	6	7	8	9	10	11	12	13
B3LYP In vacuo	M062X In vacuo	B3LYP-D3 In vacuo	M062X-D3 In vacuo	B3LYP PCM	M062 PCM	B3LYP-D3 PCM	M062X-D3 PCM	B3LYP SMD	M062X SMD	B3LYP-D3 SMD	M062X-D3 SMD	ωB97X-D PCM

Table 2
Semi-empirical computational approaches.

A	B	C	D	E	F	G	H	I	J	K	L
PM7 In vacuo	PM7-D3 In vacuo	PM6 In vacuo	PM6-D3 In vacuo	PM7 PCM	PM7-D3 PCM	PM6 PCM	PM6-D3 PCM	PM7 SMD	PM7-D3 SMD	PM6 SMD	PM6 -D3 SMD

gradual release of the restraints in successive steps at 300 K in NVT ensemble and finally density equilibration in the isobaric-isothermal (NPT) ensemble for 250 ps. Subsequently, production runs of the system under the NPT ensemble using a Berendsen-type algorithm with coupling constants of 1.0 ps were carried out under physiological conditions until reaching 12 ns.

The MD outputs were processed through the cpptraj module [30] of AMBER 12, to calculate the structural analyses (RMSD, distances, H-bonding). Moreover, the total guest binding energy ΔGbind, including the entropic term (ΔS) (calculated with the nmode module of AMBER 12) and the analysis of its components ΔEMM (changes in the gas-phase molecular mechanics (MM) energy) and ΔGsolv (solvation free energy) were computed with the aid of MM/PBSA.py script [31] implemented in AMBER 12. VMD [32] was also used for visualization and structural analyses of the MD trajectories.

3. Results and discussion

3.1. HRMS

As it was already mentioned, in our previous article concerning the EST-βCD complex in the solid state, we reported that the guest:host molar ratio in a crystal form is 1:2 [15]. However, we were interested whether 1:2 is the only form present in the aqueous solution of the complex, as well. Hence, we have defined 3 possible scenarios. The first option was presence of just 1:1 molar ratio complex, as it was stated many times in the literature. The second option was presence of both 1:1 and 1:2 complexes, where the 1:2 ratio would be the effect of the increasing concentration happening due to the solvent evaporation. According to that hypothesis, the amount of 1:2 complex should increase with the decrease of water content, resulting in the solely 1:2 complex in the solid state. And finally, the third possibility was presence of the identical stoichiometry as observed in a solid state which is 1:2 molar ratio.

To verify which scenario is correct, we have used the high-resolution mass spectrometry (HRMS). So far, this method has not been used often to define the molar ratio of the cyclodextrin complexes, however there are already some examples of its successful application for this purpose, published in the recent years [33,34]. The advantage of this method over the UV or fluorescence spectroscopy is that HRMS delivers a direct answer about the complex's molar ratio.

The results of the HRMS measurement of EST-βCD complex are presented in Fig. 2 and Table 3. EST forms stable associate with two βCD which can be detected in the mass spectra as ions corresponding to the protonated complex $[EST\text{-}2\beta CD + H]^+$ at *m/z* 2541.92 and the sodium adduct of complex $[EST\text{-}2\beta CD + Na]^+$ at *m/z* 2563.91. These observations suggest an interaction with 1:2 stoichiometry. No peaks corresponding to the association of one EST and one βCD were detected revealing the absence of 1:1 stable noncovalent complexes.

3.2. Phase solubility studies

The phase-solubility profile of EST in aqueous solution of successively increased β-CD concentrations at 25 °C (Fig. 3) indicates a B_S-type system which is usually observed with natural CDs, especially β-CD [34]. EST solubility increases linearly with increasing β-CD concentration in the range of 0.2 - 1.4 mM due to the formation of 1:1 EST:β-CD molar ratio inclusion complexes at the first stage. At the end of this linear portion, the maximum solubility S_{max} of EST is achieved. The solubility of the 1:1 EST:β-CD complex $S_{1:1}$ can be calculated as: $S_{1:1} = S_{max} - S_0$, where S_0 is the solubility of EST in pure water determined by the intercept of the phase solubility diagram.

Additional CD does not further increase the EST solubility and a first plateau is observed at β-CD concentration of 2 to 10 mM. This is due to the limited 1:1 complex solubility and/or the formation of 1:2 EST:β-CD molar ratio complexes. As the Gibbs phase rule indicates [35] only one

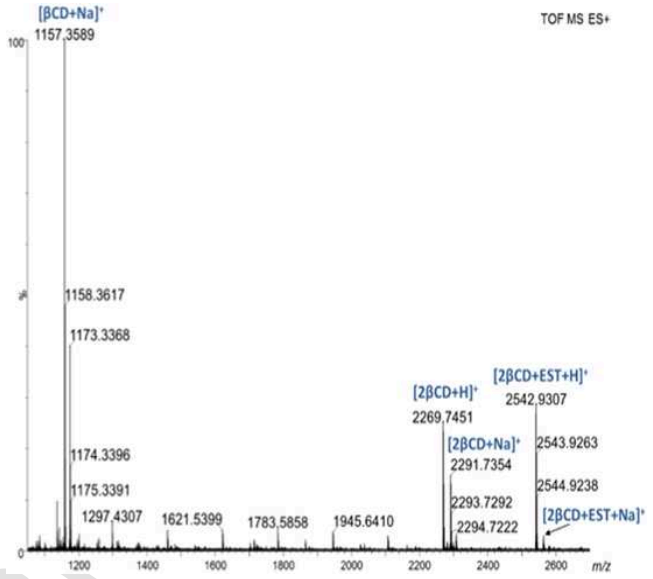


Fig. 2. HRMS spectrum of EST-βCD complex.

Table 3
The major peaks (*m/z* values and molecular formulas) from the HRMS measurement of the EST-βCD complex.

	Elemental composition	*m/z* $[M + H]^+$ or $[M+Na]^+$	
		Calculated	Found
$[\beta CD + Na]^+$	$C_{42}H_{71}O_{35}$	1157.3595	1157.3589
$[2\beta CD + H]^+$	$C_{84}H_{141}O_{71}$	2269.7474	2269.7451
$[2\beta CD + Na]^+$	$C_{84}H_{140}O_{70}Na_1$	2291.7293	2291.7354
$[2\beta CD + EST + H]^+$	$C_{102}H_{165}O_{72}$	2541.9250	2541.9248
$[2\beta CD + EST + Na]^+$	$C_{102}H_{164}O_{72}Na_1$	2563.9069	–

discrete complex may precipitate at the plateau segment of the diagram. Thus, at even higher β-CD concentrations (12 – 20 mM), where a second plateau is observed, the solubility approximates that of the pure 1:2 complex $S_{1:2}$. The scheme of the probable complex formation is given below as eq. 4.

For the two-step association process of EST complexation with β-CD, the apparent stability constants, K_1, K_2 of the following equilibria:

$$EST + CD \overset{K_1}{\leftrightarrow} EST/CD \tag{3a}$$

$$EST/CD + CD \overset{K_2}{\leftrightarrow} EST/CD2 \tag{4a}$$

and the $K_{overall} = K_1 \cdot K_2$ where estimated according to Liu et al. [35]

$$K_1 = \frac{slope}{S_0\,(1 - slope)} = 21,599\ M^{-1} \tag{3b}$$

$$K_2 = \frac{S_{1:2} \cdot\ slope}{{S_{1:1}}^2\,(1 - slope)} = 757\ M^{-1} \tag{4b}$$

where the slope was obtained from the linear part of the diagram and S_0, $S_{1:1}$, $S_{1:2}$ as described above. Thus, $K_{overall} = 1.6 \pm 0.4 \cdot 10^7\ M^{-2}$

This value is in the same order of magnitude to those estimated for progesterone, testosterone and cortisone by Liu et al. [36]. In that work the formation of inclusion complexes of steroids with β-CD at stoichiometric ratio of 1:2 was shown and its dependence on the steroid structure was discussed.

The *K* can be used to calculate Gibbs free energy (ΔG) according to the Eq. (5):

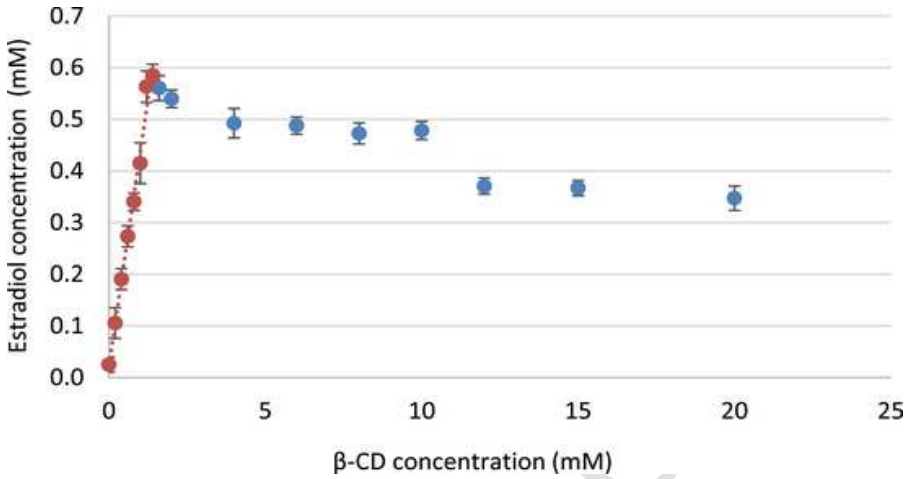


Fig. 3. Phase solubility diagram of EST/β-CD system in water at 25 °C ($n = 3$). The linear portion of the diagram (red line) was used for the calculation of K_1.

$$\Delta G = -RT\ln K \quad (5)$$

where R is gas constant (8.314 J mol^{-1} K^{-1}) and T is temperature (298 K).

This calculation reveals that the obtained here ΔG is equal to −9.92 kcal/mol.

3.3. QM calculations

As described above, the determination of this complex's stoichiometry is experimentally difficult and not straightforward task. Therefore, having in mind a huge applicability of the molecular modelling in the analysis of structure and properties of the CD complexes reviewed by us in the last year [16] we decided to apply these techniques also for this purpose. Therefore, the next step of our work was assessment of the QM approaches to check if they can properly foresee the host-guest molecular ratio and the complex association constant.

Knowledge about the crystal structure of a studied system significantly facilitates the calculations as it can be used to set the initial geometry of the complex. Thankfully, in our previous work we have determined the crystal structure of the EST-βCD complex [15]. In the current study it has been used as a starting point for all of the calculations.

Since we wanted to validate whether QM calculations can be used to predict the molar ratio and structure of the most stable complex, we have prepared 3 types of systems. The first one was 1:2 molar ratio complex and the experimental crystallographic structure of the hydrate of EST-βCD complex, after removing the water molecules, was used as an input for the computations. As EST is a molecule with quite limited conformational space, but also it is characterized by a structural anisotropy, to represent 1:1 molar ratio we needed two models called ''head up'' and ''head down'', as presented in Fig. 4. ''Head up'' is the case when EST's 5-carbon ring (steroidal D ring) goes through the wider CD's rim and the ''head down'' option is the case when it is the EST's 6-carbon ring (steroidal A ring) that protrudes through CD's wider rim. Those structures were based directly on the 1:2 experimental crystallographic data. We obtained them by removing each time different CD molecule from the 1:2 system.

3.4. DFT calculations

We have decided to apply Density Functional Theory (DFT) and semi-empirical methods. The chosen computational approaches were already presented in Table 1 and Table 2. In contrast to the majority of previously published cases of the CD complexes analysis by the means

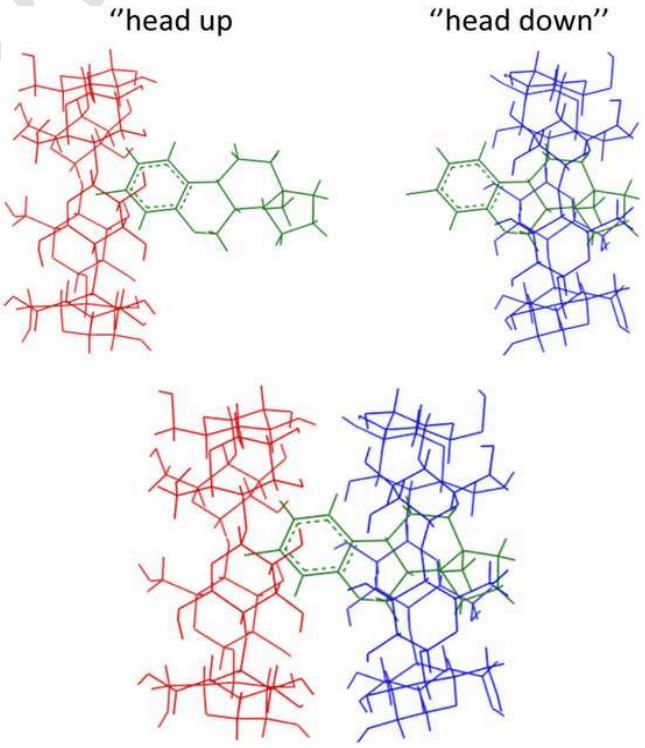


Fig. 4. Initial, non-optimized structures of 1:1 and 1:2 complexes, based on the experimental crystallographic data.

of DFT [16] where usually either 6–31 G or 6–31G(d) basis sets have been used, in this study for all DFT calculations we have used a relatively large 6–311G(d,p) basis set. Even if such approach has significantly elongated the computational time, it was a conscious choice. This way we have eliminated a risk of the influence of the basis set of the insufficient size on the obtained results.

Except of the choice of the functional in DFT (here: B3LYP and M062X) and type of the semi-empirical method (here: PM6 and PM7), one of the crucial decisions was about the type of the solvent representation. We have tested all 3 most common options: lack of solvent representation, PCM and SMD implicit solvent models.

PCM is the most often used solvent model in the computational analysis of CD complexes [37]. A different approach is presented by SMD. This model defines the free energy of solvation via two components: the one is electrostatic contribution arising from the self-consistent reaction field, the other comes from the short-range interactions between the solute and solvent molecules [20].

Second choice of the calculations parameters was the implementation or lack of the dispersion correction. Noncovalent forces like London and van der Waals interactions are crucial for the formation and stability of the CD inclusion complexes. This aspect is not included in the calculations using exchange-correlation functionals, as the long-range electron correlation effect, known as the London part of the dispersion energy term, is not included in the Kohn–Sham DFT equation. For years this was a real issue influencing the accuracy of DFT calculations. Nowadays, different dispersion corrections are available. According to our literature research [16], in almost all studies concerning CD complexes only Grimme dispersion correction (here: D3) was used. Hence, in this study we also apply only this type of dispersion correction and compare the results to the calculations where no dispersion correction was implemented.

The results of the DFT calculations are presented in Table 4. It is clearly visible that each of three variables used in this study and described above: type of functional, solvation scheme, dispersion correction had an influence on the results.

From the energetic point of view, in 6 out of 13 cases clearly the preferred structure is the '1:1, either "head up'' (or "head down". As we know from the HRMS studies, this is not consistent with the reality. The approaches which properly predict the 1:2 complex stoichiometry are models numbered 3, 5–8, 11 and 13. On this example we can see that applying dispersion correction improve the accuracy of results when the system is treated as the in vacuo one (models 1 and 3). Simultaneously, PCM solvent model works well also when no dispersion correction has been used (Models 5 and 6) [38].

Among them there is one scheme, the scheme number 5 (B3LYP-PCM, without dispersion correction), that most significantly favours the stability of the 1:2 molar ratio complex over the 1:1 ones, both in terms of energy and free enthalpy of complexation.

3.5. Semi-empirical calculations

From the energetic point of view, the results of the semi-empirical calculations are diversified, Table 5. In 10 out 12 cases, the 1:2 molecular ratio complex has been defined as the most stable one, what stays in accordance with the experimental results. However, in those two cases where the 1:1 complex stoichiometry has been favoured, difference between the 1:1 most preferred complex and 1:2 complex is small and has the value of 0.59 kcal/mol and 3.15 kcal/mol for C and F schemes, respectively. In other words, in general, when taking into account the energetic aspect, the applied semi-empirical approaches properly predicted which molar ratio describes a complex of the highest stability.

11 out 12 models show that all three options: 1:1 head up, 1:1 head down and 1:2, are energetically stable. Only in model K the value of ΔE is positive for 1:1 head down complex. In some computational schemes differences between these three possible options are almost neglectable, for instance in model C. Whereas in other cases, the differences are much bigger, like in model L, where this difference reached up to around 37 kcal/mol.

On the contrary, from thermodynamic point of view, the majority of models pointed out the 1:1 complex as the most stable. The exceptions are models A, B, I and K. However, similarly to the ΔE results, in almost all cases, all three structural options have been defined as probable. Only in model K some of the ΔG values are positive. Model K is also the only computational scheme which shows a distant difference between both 1:1 stoichiometries and 1:2 stoichiometry and distinctively favours the 1:2 complex molar ratio. Hence, we can assume that the K computational model (PM6-SMD, without dispersion correction) predicted the experimental results in a most accurate way.

Except for this one K model, in rest of the cases the differences in values between 1:1 head up and 1:1 head down options within one computational method are similar. This suggests that there is a similar probability of creation of 1:1 head up and 1:1 head down complexes.

All those findings may be taken as a guide to create the hypothesis on the 1:2 molar ratio complex formation path. Most probably, the complexation happens in two steps. The first one, is creation of the 1:1 complex. The second step is association of the second βCD to the already existing 1:1 EST-βCD system. Which complex out of those two, 1:1 head up or 1:1 head down, is formed at the beginning can be deducted from the ΔG values. The complex characterized by a lower value should be created as the first one. According to both the most accurate DFT approach, 5 (B3LYP PCM), as well as to the MD MMGBSA results, the 1:1 head up complex is the more stable one. Therefore, it is highly probable that the complex formation occurs according to the scheme (Eq. (6)) presented below.

$$EST + \beta CD \rightleftharpoons \mathbf{1:1\ up} \overset{\beta CD}{\rightarrow} \mathbf{1:2} \tag{6}$$

With the first step being reversible and second irreversible reaction. Such predictions about the mechanism of the complex's formation are possible only thanks to application of the molecular modelling approach. The structural analysis of the results can be find in the Supporting Information.

3.6. Computational thermodynamic results vs experimental data

As described in the previous section, obtention of the experimental complex stability constant allowed to define ΔG, which in case of EST-βCD complex is equal to −9.92 kcal/mol. Among the tested QM approaches, the ones which favour the most the 1:2 molar ratio are those in which the PCM correction has been applied However, the calculated ΔG overestimate the experimental ones Interestingly, the values close to the experimental one (−9.92 kcal/mol) have been obtained using M062X in vacuo (−9.18 kcal/mol) and PM7-D3 PCM approaches (−9.13 kcal/mol). Unfortunately, at the same time, both of them suggest that the 1:1 is more stable ratio than 1:2. Here, it should also be noted that β-CD is surrounded by hydration waters, which could be included in the equation of the complex formation. It was shown previously that an agreement between theoretical and experimental entropy data for inclusion complex formation was only attained when explicit water molecules were included [39].

3.7. MD simulations

MD simulations in aqueous media were carried out for the 3 types of systems, i.e. the crystallographically determined 1:2 guest:host complex and the two 1:1 monomeric complexes of opposite EST accommodation in the β-CD cavity (noted as "head up" and "head down" in Fig. 7). By monitoring the frames during the time interval of the simulations the following observations were made:

In the case of the 1:2 complex, where the starting model was retrieved from the crystal structure (Fig. 5a), the β-CD dimer encapsulating an EST molecule is preserved in the time frame of the simulation. In the absence of crystal contacts and in the presence of the surrounding water molecules, the guest EST rotates around and moves along the 7-fold molecular β-CD axis. However, it was observed a clear tendency of the guest's steroidal A-ring to be accommodated near the narrow rim of β-CD and its D-ring closer to the dimeric interface region. The measured distance between the center of mass (COM) of the A-ring and the O4n atoms mean plane of the host β-CD1, in whose cavity the A-ring is located, fluctuates around 1 Å, whereas that between the D-ring COM and the O4n plane of the other host (β-CD2), fluctuates in the range of 1 to 3 Å (Fig. 5b and 5c).

Table 4

DFT calculations results. Yellow colour indicates the lowest value of ΔE within the given method (within the column). Blue colour indicates the lowest value of ΔG within the given method (within the column). ΔG and ΔE "1:2 from 1:1 up" – the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "up" is used as an initial structure for the second step; ΔG and ΔE "1:2 from 1:1 down" - the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "down" is used as an initial structure for the second step.

		B3LYP In vacuo	M062X In vacuo	B3LYP -D3 In vacuo	M062X -D3 In vacuo	B3LYP PCM	M062X PCM	B3LYP -D3 PCM	M062X -D3 PCM	B3LYP SMD	M062X SMD	B3LYP -D3 SMD	M062X -D3 SMD	ωB97X-D PCM
model	kcal/mol	1	2	3	4	5	6	7	8	9	10	11	12	13
1:1 up	ΔE	-6.47	-37.35	-54.67	-49.14	-6.93	-31.49	-30.41	-45.89	-5.06	-30.35	-48.12	-44.56	-36.80
	ΔG	8.00	-12.28	-31.73	-34.28	6.12	-8.18	-9.64	-19.64	7.44	-10.08	-24.81	-21.35	-17.91
1:1 down	ΔE	0.44	-27.64	-42.20	-47.06	1.87	-30.82	-30.88	--43.37	1.99	-30.88	-49.20	-44.17	-41.70
	ΔG	15.40	-12.10	-21.32	-23.36	16.93	-8.22	-10.9	-19.09	14.24	-11.01	-26.97	-19.61	-19.90
1:2	ΔE	2.12	-26.46	-64.81	-47.99	-61.42	-81.56	-41.75	-118.54	5.59	-20.69	-65.69	-36.78	-120.68
	ΔG	13.33	-9.18	-38.67	-26.28	-47.69	-28.08	-49.66	-34.58	15.18	-0.75	-44.76	-19.26	-65.87
	ΔE	8.59	10.89	-10.14	1.15	-54.49	-50.07	-11.34	-72.65	10.65	9.66	-17.57	7.78	-83.88
1:2 from 1:1 up	ΔG	5.33	3.1	-6.94	8	-53.81	-19.9	-40.02	-14.94	7.74	9.33	-19.95	2.09	-47.96
	ΔE	1.68	1.18	-22.61	-0.93	-63.29	-50.74	-10.87	-161.91	3.6	10.19	-16.49	7.39	-78.98
1:2 from 1:1 down	ΔG	-2.07	2.92	-17.35	-2.92	-64.62	-19.86	-38.76	-15.49	0.94	10.26	-17.79	0.35	-45.97

Table 5

Semi-empirical calculations results. Yellow colour indicates the lowest value of ΔE within the given method (within the column). Blue colour indicates the lowest value of ΔG within the given method (within the column). ΔG and ΔE "1:2 from 1:1 up" – the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "up" is used as an initial structure for the second step; ΔG and ΔE "1:2 from 1:1 down" - the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "down" is used as an initial structure for the second step.

		PM7 In vacuo	PM7 -D3 In vacuo	PM6 In vacuo	PM6 -D3 In vacuo	PM7 PCM	PM7 -D3 PCM	PM6 PCM	PM6 -D3 PCM	PM7 SMD	PM7 -D3 SMD	PM6 SMD	PM6 -D3 SMD
model	kcal/mol	A	B	C	D	E	F	G	H	I	J	K	L
1:1 up	ΔE	-59.50	-40.53	-11.68	-37.84	-54.53	-35.59	-7.45	-34.91	-51.21	-32.38	-5.08	-24.27
	ΔG	-37.13	-27.15	-38.64	-54.75	-30.07	-30.93	-10.23	-43.58	-34.98	-17.14	7.82	-24.52
1:1 down	ΔE	-54.76	-38.20	-11.05	-35.41	-43.27	-38.76	-8.96	-35.52	-47.85	-34.58	86.58	-27.98
	ΔG	-37.14	-43.81	-25.83	-31.87	-25.31	-42.85	-5.93	-53.45	-41.29	-18.23	9.60	-54.32
1:2	ΔE	-84.25	-61.17	-11.09	-61.45	-60.84	-35.61	-18.59	-54.94	-61.54	-40.77	-19.91	-64.91
	ΔG	-71.81	-48.48	5.34	-41.83	-43.04	-9.13	0.62	-34.81	-48.49	-23.39	-3.33	-39.68
1:2 from 1:1 up	ΔE	-24.75	-20.64	0.59	-23.61	-6.31	-0.02	-11.14	-20.03	-10.33	-8.39	-14.83	-40.64
	ΔG	-34.68	-21.33	43.98	12.92	-12.97	21.8	10.85	8.77	-13.51	-6.25	-11.15	-15.16
1:2 from 1:1 down	ΔE	-29.49	-22.97	-0.04	-26.04	-17.57	3.15	-9.63	-19.42	-13.69	-6.19	-106.49	-36.93
	ΔG	-34.67	-4.67	31.17	-9.96	-17.73	33.72	6.55	18.64	-7.2	-5.16	-12.93	14.64

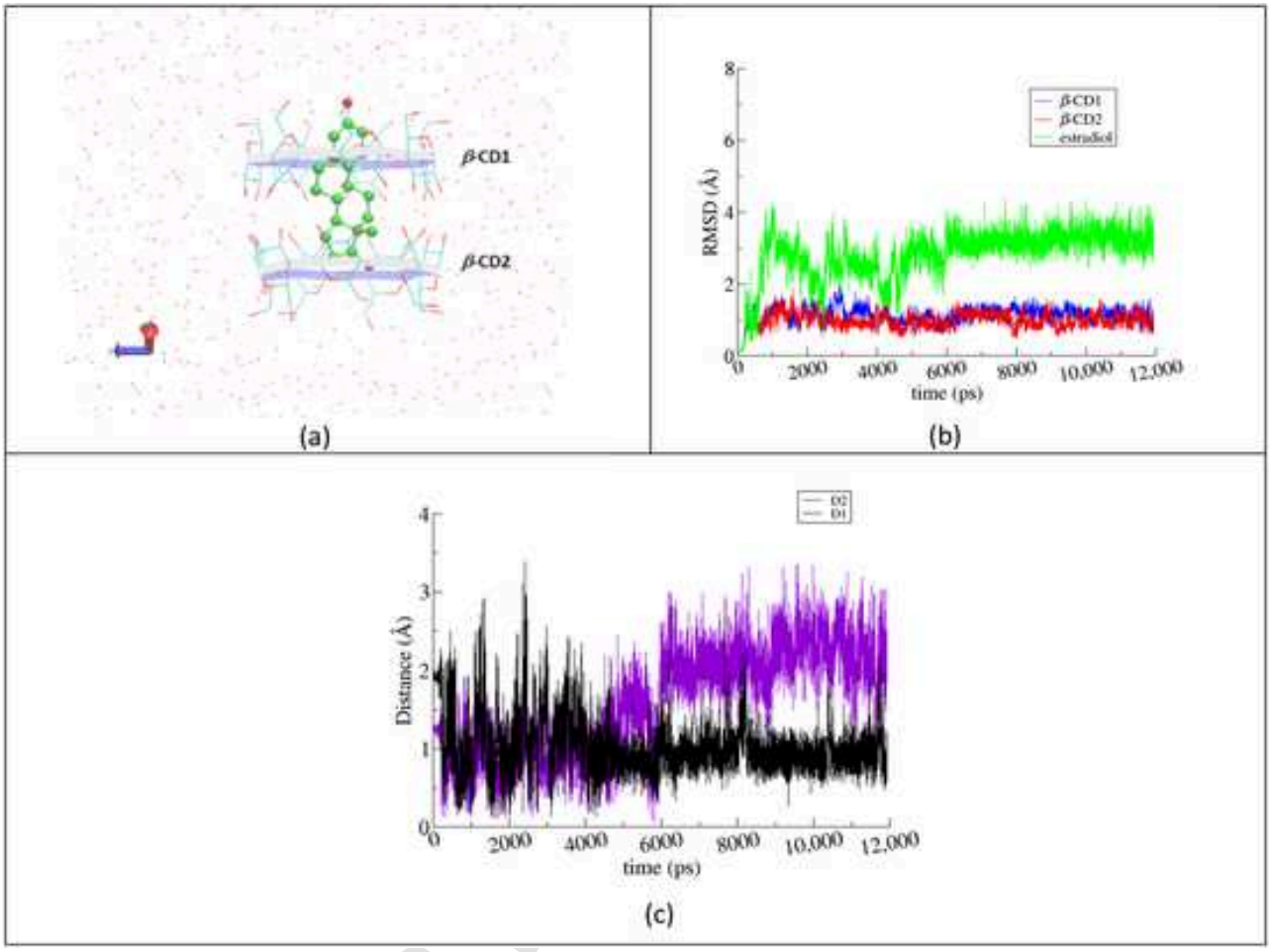


Fig. 5. (a) The starting model of the 1:2 inclusion complex based on the crystallographically determined atomic coordinates (CSD Refcode: OFANUI). (b) RMSD evolution of the host and guest molecules of the complex between the simulated states and the first frame of the simulations. (c) Distance D1 between the center of mass (COM) of the steroidal A-ring of EST and the O4*n* atom mean plane of the host β-CD1; distance D2 between COM of the steroidal D-ring of EST and the O4*n* atom mean plane of the other host of the dimeric cavity, β-CD2.

For the 1:1 molar ratio complex, both "head up" and "head down" models were examined. In the case of the "head down" binding mode, the A-ring of the guest cannot be stabilized in the host's wide rim with the rest part of EST protruding from the narrow rim of the host. Thus, EST is swiftly displaced from its initial location, exposing its A-ring to the solvent and accommodating the D-ring in the cavity (Fig. 6a). This behavior is also reflected in the high EST mobility displayed in the respective RMSD plot (Fig 6b).

On the other hand, in the case of the "head up" binding mode, the A-ring of EST which is initially exposed to the solvent by protruding from the narrow β-CD rim, it is accommodated quickly in the narrow β-CD rim where it remains relatively stable in the time frame of the simulation (Fig. 7a). The respective RMSD plot for the molecules of the system clearly shows a lower mobility of the guest compared to that of the "head down" system.

From all the above, it is concluded that the formation of a 1:1 inclusion complex of the "head up" binding mode is favored over the "head down" binding mode. Moreover, the dynamic behavior of the examined 1:2 complex, showed a very stable complex that tends to retain the accommodation of the guest (with its A-ring near the narrow rim and the D-ring near the interface of the hosts' dimer) according to that of the "head up" binding mode. These findings support the proposed complex formation described in the scheme of eq. 4.

Molecular Mechanics/Generalized Bohrn surface area (MM/GBSA) calculations [40], performed for the 3 examined types of complexes, further verify the above conclusions. The estimated host–guest binding affinities, as listed in Table 6, were extracted from 10,000 snapshots over the last 10-ns of the MD simulations. As expected, the lower ΔG-bind value is estimated for the 1:2 complex mainly due to the extended van der Waals interactions between the guest and the dimeric host, that significantly decrease the averaged change of van der Waals energies (ΔEvdW) upon EST inclusion in the host dimeric cavity. By comparing the ΔGbind values estimated for the systems of common 1:1 guest:host stoichiometry ("head down" and "head up"), the considerably lower ΔGbind value of the "head up" system indicates a more stable inclusion complex.

3.8. Computational stability results vs experimental data

As described in the previous section, acquisition of the experimental complex stability constant allowed to define ΔG, which in case of EST-βCD complex is equal to −9.92 kcal/mol. This value is within the limits of uncertainty of ΔGbind for 1:2 complex, −13.66 ± 4.20 kcal/mol (Table 6). However, the values of ΔGbind obtained for 1:1 complex, either head-up or head-down orientations, −2.27 ± 2.39 and −6.25 ± 2.40 respectively, although not strictly within the uncertainty limits, are also close to the experimentally determined one. Therefore,

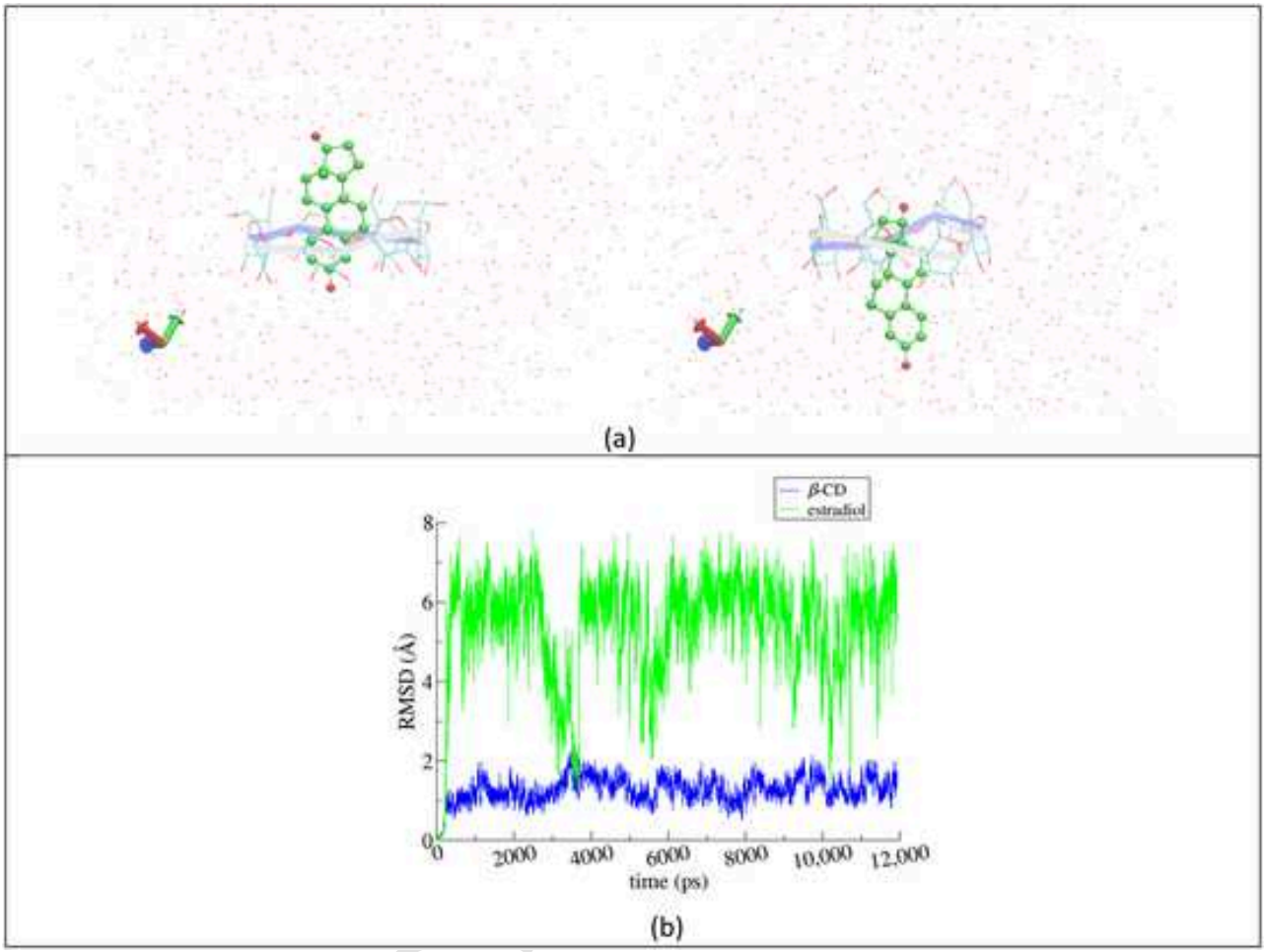


Fig. 6. (a) Two representative snapshots, at 0 and the 10th ns of the "head down" system simulation.
(b) RMSD plot for the host and guest molecule of the complex. The guest cannot be tightly stabilized in this orientation, thus exhibiting high mobility.

while the MD/MMGBSA method allows to properly indicate the order of the magnitude of ΔGbind, it should be supported by the experimental analysis such as HRMS to confirm the complex ratio.

4. Conclusions

Even though the existence of the EST-βCD complex in water solution has been known for decades, and despite quite some studies performed to define its molar ratio, this complex's structure remained not properly determined until now. The hypothesis that the knowledge in this context might be not complete, has occurred after determination of the EST-βCD crystal structure where the molar ratio was found to be 1:2 (EST:βCD).

In this work, thanks to application of the HRMS approach, it has been indisputably proven that the EST-βCD complex molar ratio is 1:2 and not as previously assumed 1:1. Moreover, the phase solubility studies confirmed these results. This type of experiments has been performed before, however, never the 1:2 molar ratio has been taken into account as a possible description of this system. In other words, the indisputable HRMS measurement results prompted the revision of the phase solubility studies. This allowed to properly define complex stability constant K, what in turn delivered information about the Gibbs free energy value of the complex formation.

The structure and thermodynamics of the complex were further analyzed using various QM (DFT, semi-empirical) and MM (MD/MMGBSA) approaches. Tests on the application of different computation parameters such as presence/absence of dispersion correction, choice of implicit solvent model or DFT functional, have been performed.

Possession of the credible experimental data allowed to assess the computational approaches. While some of the "static" QM methods properly indicated the correct host: guest ratio at the same time they failed to accurately predict the Gibbs free energy of complexation. On the other hand, QM methods that properly described the value of ΔG of 1:2 complex formation, such as M062X in vacuo, favored the 1:1 stoichiometry, which was experimentally excluded. The MD/MMGBSA method, although performed at the lower level of theory, accurately predicted the stability constant of the complexed but was not conclusive to indicate the formation of either 1:1 or 1:2 complex.

This leads to the conclusion that among tested computational approaches, there are some which are able to properly predict the composition of such complex and some that can assess the stability of the studied system. However, there is not single method that would allow to reproduce both the stoichiometry and the thermodynamic stability of the complex at the same time.This study finally describes the structure and thermodynamics of the EST-βCD complex in aqueous solution and delivers experiment-based information about the formation of this complex.

CRediT authorship contribution statement

Anna Helena Mazurek: Writing – original draft, Software, Investigation, Funding acquisition, Data curation, Conceptualization. **Łukasz**

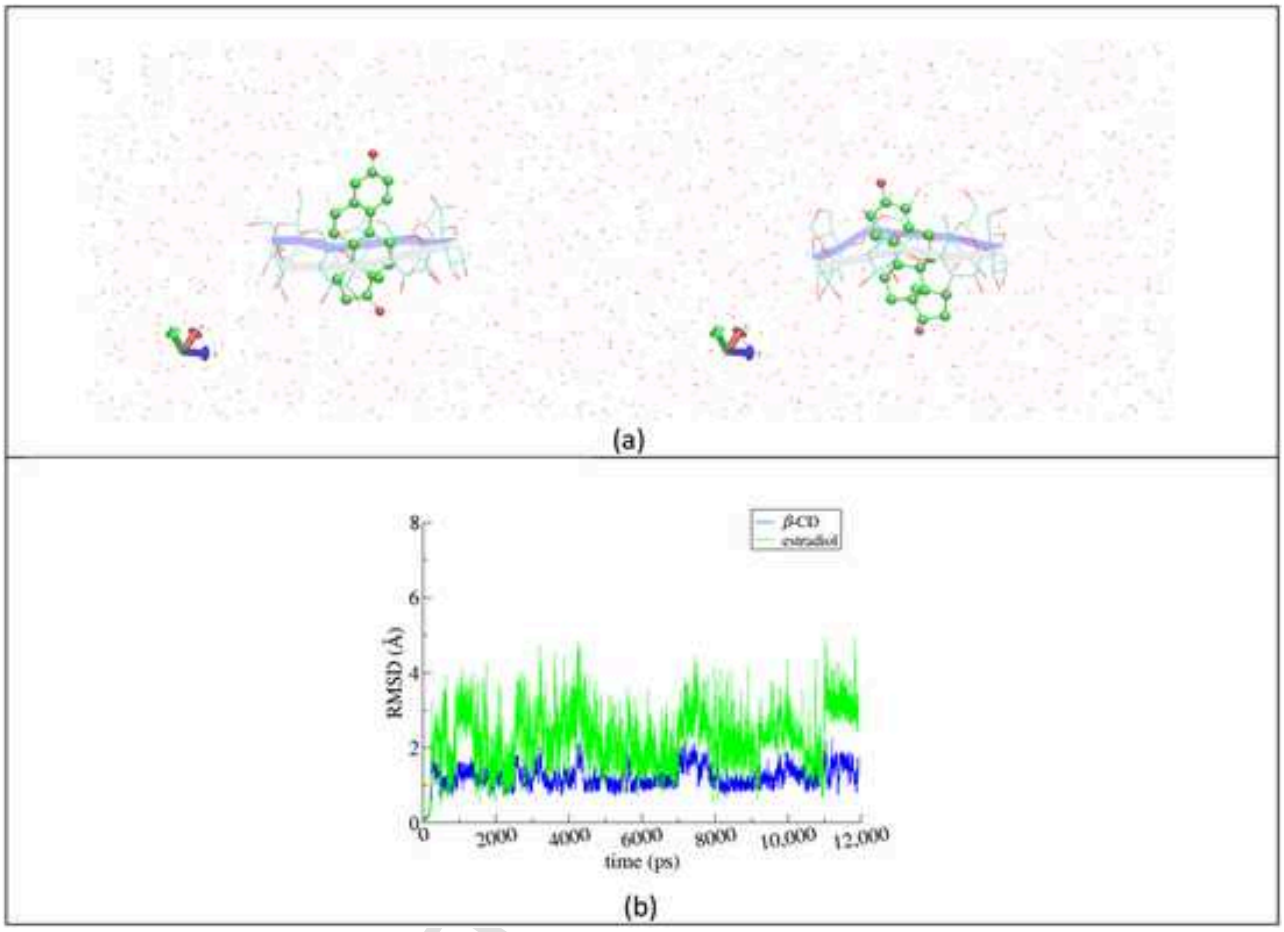


Fig. 7. (a) Two representative snapshots, at 0 and the 10th ns of the "head up" system simulation.
(b) RMSD plot the for host and guest molecule of the complex. The mobility of the guest is clearly lower than that of EST in the "head down" system simulation.

Table 6
Binding free energies and their standard deviations (kcal/mole) resulting from MM/GBSA analysis of the inclusion compounds of EST in β-CD with guest: host ratios of 1:2, 1:1 ("head down" mode) and 1:1 ("head up" mode), respectively.

	EST/β-CD (1:2)	EST/β-CD (1:1 "head down")	EST/β-CD (1:1 "head up")
ΔE_{vdW}	−47.31 ± 2.41	−29.10 ± 2.31	−30.46 ± 1.68
ΔE_{ele}	−4.15 ± 2.78	−2.12 ± 2.13	−2.09 ± 1.72
ΔE_{MM} [a]	−51.46 ± 3.58	−31.23 ± 3.21	−32.55 ± 2.38
ΔG_{GB}	23.22 ± 2.98	15.63 ± 2.80	13.45 ± 1.99
$\Delta G_{nonpolar}$	−4.45 ± 0.22	−2.94 ± 0.16	−3.03 ± 0.12
$\Delta G_{solvation}$ [b]	18.77 ± 2.94	12.69 ± 2.72	10.42 ± 1.94
ΔH [c]	−32.69 ± 4.63	−18.54 ± 2.06	−22.13 ± 1.97
T·ΔS [d]	−19.03 ± 3.00	−16.26 ± 1.21	−15.88 ± 1.36
ΔG_{bind} [e]	−13.66 ± 4.20	−2.27 ± 2.39	−6.25 ± 2.40

ΔE_{vdW} = van der Waals contribution from molecular mechanics; ΔE_{ele} = electrostatic energy as calculated by the molecular mechanics force field; ΔG_{GB} = the electrostatic solvation energy (polar contribution) calculated using the GB model; $\Delta G_{nonpolar}$ = nonpolar contribution to the solvation free energy, calculated by the solvent-accessible surface area (SASA) method;.

[a] $\Delta E_{MM} = \Delta E_{vdW} + \Delta E_{ele}$;.
[b] $\Delta G_{solvation} = \Delta G_{GB} + \Delta G_{nonpolar}$;.
[c] $\Delta H = \Delta G_{solvation} + \Delta E_{MM}$;.
[d] $T{\cdot}\Delta S$ entropic term calculated by normal mode analysis;.
[e] $\Delta G_{binding} = \Delta H - T{\cdot}\Delta S$.

Szeleszczuk: Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Conceptualization. **Kostas Bethanis:** Writing – original draft, Methodology, Investigation, Conceptualization. **Elias Christoforides:** Software, Resources, Project administration, Methodology. **Marta Katarzyna Dudek:** Software, Methodology, Conceptualization. **Ewelina Wielgus:** Writing – review & editing, Validation, Project administration. **Dariusz Maciej Pisklak:** Writing – review & editing, Writing – original draft, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the Medical University of Warsaw, Poland [grant number WF7/1/F/MB/N/23].

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.molstruc.2024.138710.

References

[1] J. Szejtli, Introduction and general overview of cyclodextrin chemistry, Chem. Rev. 98 (1998) 1743–1754.
[2] B.G. Poulson, Q.A. Alsulami, A. Sharfalddin, E.F. El Agammy, F. Mouffouk, A.-H. Emwas, L. Jaremko, M. Jaremko, Cyclodextrins: structural, chemical, and physical properties, and applications, Polysaccharides 3 (2022) 1.
[3] G. Crini, Review: a history of cyclodextrins, Chem. Rev. 114 (2014) 10940–10975.
[4] T. Irie, K. Uekama, Pharmaceutical applications of cyclodextrins. III. Toxicological issues and safety evaluation, J. Pharm. Sci. 86 (1997) 147–162.
[5] V. Aiassa, C. Garnero, M.R. Longhi, A. Zoppi, Cyclodextrin multicomponent complexes: pharmaceutical applications, Pharmaceutics 13 (2021) 1099.
[6] S.S. Jambhekar, P. Breen, Cyclodextrins in pharmaceutical formulations I: structure and physicochemical properties, formation of complexes, and types of complex, Drug Discov. Today. 21 (2016) 356–362.
[7] S.S. Jambhekar, P. Breen, Cyclodextrins in pharmaceutical formulations II: solubilization, binding constant, and complexation efficiency, Drug Discov. Today. 21 (2016) 363–368.
[8] M.P. Thomas, B.V. Potter, The structural biology of oestrogen metabolism, J. Steroid Biochem. Mol. Biol. 137 (2013) 27–49.
[9] Salole E.G. Estradiol, Analytical profiles of drug substances, 1986, 15, 283–318.
[10] N. Sadlej-Sosnowska, Fluorometric determination of association constants of three estrogens with cyclodextrins, J Fluoresc 7 (1997) 195–200.
[11] C. Yañez, J. Basualdo, P.J. Jara-Ulloa, A. Squella, Inclusion complexes of estrone and estradiol with β-cyclodextrin: voltammetric and HPLC studies, J. Phys. Org. Chem. 20 (2007) 499–505.
[12] R.L. Pérez, G.M. Escandar, Spectrofluorimetric study of estrogen-cyclodextrin inclusion complexes in aqueous systems, Analyst 138 (2013) 1239–1248.
[13] D.H. Schwarz, A. Engelke, G. Wenz, Solubilizing steroidal drugs by β-cyclodextrin derivatives, Int J Pharm 531 (2017) 559–567.
[14] Z.Y. Lin, X.X. Wang, S.B. Kou, J.H. Shi, Exploring the inclusion interaction of estradiol with β-CD and HP-β-CD with the help of molecular dynamics simulation as well as multi-spectroscopic approaches, Spectrochim Acta A Mol Biomol Spectrosc 269 (2022) 120764.
[15] A.H. Mazurek, Ł. Szeleszczuk, K. Bethanis, E. Christoforides, M.K. Dudek, M. Zielińska-Pisklak, D.M. Pisklak, 17-β-Estradiol—B-Cyclodextrin complex as solid: synthesis, structural and physicochemical characterization, Molecules 28 (2023) 3747.
[16] A.H. Mazurek, Ł. Szeleszczuk, Current status of quantum chemical studies of cyclodextrin host–guest complexes, Molecules 27 (2022) 3874.
[17] T. Higuchi, K.A. Connors, Phase solubility techniques, Adv. Anal. Chem. Instrument. 4 (1965) 117–212 - Open Access Library Available online: http://www.oalib.com/references/7163685. (accessed on 26 January 2024).
[18] Software https://gaussian.com/products/ accessed on 24th January 2024
[19] M. Cossi, N. Rega, G. Scalmani, V. Barone, Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model, J. Comput. Chem. 24 (2003) 669–681.
[20] A.V. Marenich, C.J. Cramer, D.G Truhlar, Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions, J. Phys. Chem. B 113 (2009) 6378–6396.
[21] J.-D. Chai, M. Head-Gordon, Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections, Phys. Chem. Chem. Phys. 10 (2008) 6615.
[22] J. Eberhardt, D. Santos-Martins, A.F. Tillack, S. Forli, AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings, J. Chem. Inf. Model. 61 (2021) 3891–3898.
[23] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B.A. Shoemaker, P.A. Thiessen, B. Yu, et al., PubChem 2023 update, Nucleic Acids Res. 51 (2023) D1373–D1380.
[24] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the amber biomolecular simulation package, Wiley Interdiscip. Rev.: Comput. Mol. Sci. 3 (2012) 198–210.
[25] J. Wang, W. Wang, P. Kollman, D. Case, ANTECHAMBER: an accessory software package for molecular mechanical calculations, J. Chem. Inf. Comput. Sci. - JCISD (2000) 222.
[26] K.N. Kirschner, A.B. Yongye, S.M. Tschampel, J. Gonzalez-Outeirino, C.R. Daniels, B.L. Foley, R.J. Woods, GLYCAM06: a generalizable biomolecular force field. Carbohydrates, J. Comput. Chem. 29 (2008) 622–655.
[27] P. Mark, L. Nilsson, Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K, J. Phys. Chem. A 105 (2001) 9954–9960.
[28] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[29] V. Kräutler, W.F. van Gunsteren, P.H. Hünenberger, A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations, J. Comput. Chem. 22 (2001) 501–508.
[30] D.R. Roe, T.E. Cheatham 3rd, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theory Comput. 9 (2013) 3084–3095.
[31] B.R. Miller 3rd, T.D.J. McGee, J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBSA.Py: an efficient program for end-state free energy calculations, J. Chem. Theory Comput. 8 (2012) 3314–3321.
[32] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (33–38) (1996) 27–28.
[33] G. Zengin, Yildiztugay E. Nilofar, A. Bouyahya, H. Cavusoglu, R. Gevrenova, D. Zheleva-Dimitrova, A comparative study on UHPLC-HRMS profiles and biological activities of inula sarana different extracts and its beta-cyclodextrin complex: effective insights for novel applications, Antioxidants (Basel) 12 (2023) 1842.
[34] T.K. Špehar, M. Pocrnić, D. Klarić, B. Bertoša, A. Čikoš, M. Jug, J. Padovan, S. Dragojević, N. Galić, Investigation of praziquantel/cyclodextrin inclusion complexation by NMR and LC-HRMS/MS: mechanism, solubility, chemical stability, and degradation products, Mol. Pharm 18 (2021) 4210–4223.
[35] M.E. Brewster, T. Loftsson, Cyclodextrins as pharmaceutical solubilizers, Adv. Drug Deliv. Rev. 59 (2007) 645–666.
[36] F.Y. Liu, D.O. Kildsig, A.K. Mitra, Beta-cyclodextrin/steroid complexation: effect of steroid structure on association equilibria, Pharm. Res 7 (1990) 869–873.
[37] M. Cossi, N. Rega, G. Scalmani, V. Barone, Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model, J. Comput. Chem. 24 (2003) 669–681.
[38] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132 (2010) 154104.
[39] J.F. Lopes, Nascimento C.S. Jr, C.P.A. Anconi, H.F.D. Santos, W.B Almeida, Inclusion complex thermodynamics: the β-cyclodextrin and sertraline complex example, J. Mol. Graph. Model 62 (2015) 11–17.
[40] S. Genheden, U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, Expert Opin. Drug Discov. 10 (2015) 449–461.

Journal of Molecular Structure 1313 (2024) 138710

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: www.elsevier.com/locate/molstr

17-β-Estradiol—β-Cyclodextrin complex as an aqueous solution: Structural and physicochemical characterization supported by MM and QM calculations

Anna Helena Mazurek [a,b], Łukasz Szeleszczuk [a,*], Kostas Bethanis [c], Elias Christoforides [c], Marta Katarzyna Dudek [d], Ewelina Wielgus [d], Dariusz Maciej Pisklak [a]

[a] Department of Organic and Physical Chemistry, Faculty of Pharmacy, Medical University of Warsaw, Banacha 1 Str., 02-093 Warsaw, Poland
[b] Doctoral School, Medical University of Warsaw, Żwirki i Wigury 81 Str., 02-093 Warsaw, Poland
[c] Laboratory of Physics, Department of Biotechnology, Agricultural University of Athens, 11855 Athens, Greece
[d] Structural Studies Department, Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, Sienkiewicza 112 Str., 90-363 Łódź, Poland

ARTICLE INFO

Keywords:
Estradiol
Cyclodextrin
DFT calculations
Complex stability constant

ABSTRACT

17-β-estradiol (EST) is an Active Pharmaceutical Ingredient characterized by a low water solubility. Complexation with β-cyclodextrin (βCD) enhances its bioavailability, hence such complex is an interesting research object from pharmaceutical point of view. However, basic facts like description of complex's structure and definition of its molar ratio, were debatable already for decades. This work for the first time justifies the EST:βCD molar ratio as 1:2 using the HRMS (high-resolution mass spectrometry) and phase solubility studies. The latter are used to define complex stability constant, as well. The structure and stability is analyzed using a variety of computational approaches: Quantum Mechanics (QM) based methods (DFT, semiempirical approaches) and MD/MMGBSA approach. In case of the QM, for the first time in the computational analysis of cyclodextrin complexes, a thorough benchmarking test is presented. Different computational parameters (solvent model, presence/absence of dispersion correction etc.) are used. Obtained results are compared with the experimental data.

1. Introduction

Cyclodextrins (CDs) are oligosaccharides of a donut-like structure, which enables them to form inclusion complexes with non-polar substances. This characteristic is used by the pharmaceutical industry but at the same time it often poses a non-trivial questions about the structure and stability of the created complexes.

More precisely, CDs are cyclic structures composed of glucose subunits, joined by α−1,4 glycosidic bonds [1]. Because of the cyclic character of CDs, various chemical compounds can enter CD's void and this way inclusion complexes are created. Depending on the size of a chemical guest, different types of CDs are preferred. However, the most common one is a medium size beta-CD (βCD) which consists of 7 glucose units [1,2] (Fig. 1).

Due to the presence of hydroxyl groups, the external fragments of CDs are polar. When a non-polar substance enters the molecular hole of a CD, the formed host–guest complex is polar and more water soluble than a separate non-complexed guest molecule [2,3]. Therefore, knowing also that CDs are non-toxic for a human organism [4], CDs are commonly used in the pharmaceutical industry in order to increase the solubility of a complexed Active Pharmaceutical Ingredient (API) or protect it from external factors like light, humidity or heat [5-7]. An important API group characterized by a poor solubility in water are hormonal steroids. Encapsulation in CDs enhance their solubility in water and as a result also their bioavailability.

An example of the steroid hormones used as a medication is estradiol (EST) (Fig. 1). EST belongs to the estrogens group and is the most potent estrogen naturally produced by a human body [8]. Therefore, the first complexes between EST and a well-known βCD have been obtained already decades ago [9]. However, the structural analysis of such complexes happens not to be as straightforward as one would assume.

In 1997 in the article entitled ''Fluorometric Determination of Association Constants of Three Estrogens with Cyclodextrins'' the EST-βCD complex molar ratio has been defined as 1:1 [10]. This information has been used as a reference for instance in the following articles [11,12] by other scientists. Only two decades afterwards, this pre-defined EST-βCD

* Corresponding author.
E-mail address: lukasz.szeleszczuk@wum.edu.pl (Ł. Szeleszczuk).

https://doi.org/10.1016/j.molstruc.2024.138710
Received 19 February 2024; Received in revised form 3 May 2024; Accepted 21 May 2024
Available online 24 May 2024

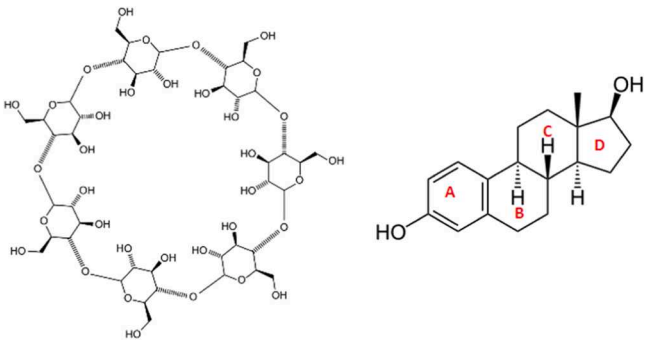


Fig. 1. Structures of β-CD and EST.

complex molar ratio and its stability constant have been revisited [13, 14]. However, again only either fluorescence or UV spectroscopy have been used for that purpose. Despite quite a few studies in this area, the 1:1 molar ratio of the complex in question has been accepted as a status quo and has never been questioned or verified by application of any other experimental methods such as high-resolution mass spectrometry (HRMS).

Recently, we have analyzed the EST-βCD complex in a solid state. For the first time, we have obtained and analyzed the structure of this complex's crystal structure. At the same time, it was one of the first ever obtained crystal structures of a steroid hormone complexed with any CD. The results have been published as [15] and the crystal structure has been deposited in the Cambridge Crystallographic Data Centre (CCDC). The results of the structural analysis in a solid state clearly indicate the 1:2 EST-βCD molar ratio, however they have also revealed significant structural disorder of this complex. Having obtained such results for a solid state, we were curious what is the molar ratio in the aqueous solution. Knowing the solid state structure of the complex, its 1:2 molar ratio after dissolution in water seems to be a scientifically sound hypothesis. Nevertheless, as we have gathered data on the previously performed experimental analyses repeatedly pointing out the 1:1 molar ratio, we decided against performing fluorescence or UV measurements and decided to concentrate on the molecular modelling techniques and HRMS spectrometry.

Complexes of various compounds with CDs have been analyzed using different molecular modelling approaches. At the beginning, computational methods used to predict the CD complexes structures and properties were the same as ones used for the analysis of much larger systems like receptors with ligands. Those were mainly Molecular Mechanics (MM) based methods and among them the most popular was molecular docking, sometimes followed by atomistic molecular dynamics (MD) simulations, at the same level of theory. However, in the recent years, with the increase of the computational power, such medium size systems like CD complexes started to be commonly analyzed using techniques based on the Quantum Mechanics (QM).

The conclusions of our recent review article [16] clearly show that the QM calculations era in the analysis of the CD complexes has already started more than a decade ago. The results show that the most commonly used are semi-empirical and Density Functional Theory (DFT) approaches. Among the former, PM6 and PM7 seem to be the most widely spread, also suggesting that they deliver the most appreciated results. When it comes to the DFT, a wider variety of computational options is used. The most frequent is application of B3LYP or M062X functionals both of the coming with or without dispersion correction. In terms of a solvent representation both possible approaches are practiced: either no solvent model is used or an implicit solvent model is applied.

As opposed to QM calculations, the Molecular Mechanics (MM) approach allows analysis of much larger systems but at the same time it delivers results of a significantly lower accuracy. However, for years CD complexes were too big for conducting QM calculations and Molecular Dynamics (MD) was the only option to obtain any information about the complexes' inner structure. Therefore, in the literature there are numerous examples of MD's application in the analysis of CD inclusion complexes. Even now this method is still in use. Therefore, we wanted to apply this well-described type of simulations in our study and compare them with the QM approaches.

Having gathered all the data on the previously conducted experimental analyses of the EST-βCD complex and having in mind our results concerning the solid state, we decided to try to determine the structure and the molar ratio of this complex. The aim of this study was to perform a benchmark studies on this topic using semi-empirical and DFT computational approaches. Additionally, we have managed to experimentally reveal the true complex's molecular ratio using a technique which has been never used in this particular case.

2. Materials and methods

2.1. Sample preparation

The EST-βCD complex was obtained by a method which is commonly used to obtain the CD inclusion complexes, a slow-cooling crystallization technique. This approach has already been used by us in a previous study describing the SCXRD analysis [15]. 60 mg of βCD was mixed in a flask with 1 mL distilled water and put into 70 °C water for 20 s to obtain a clear solution. Then the contents of the flask were poured into a beaker. In accordance with the molar mass of βCD and EST, the respective amount of EST was added to the beaker to maintain the 1:1 molar ratio. The beaker was put on a magnetic stirrer and left at room temperature for 15–20 min until a clear solution was obtained. Afterwards, the contents of the beaker were poured into a glass tube. The beaker was poured along with 0.5–1.0 mL water, which was also added to the glass tube. The glass tube was held in 70 °C water for 20 s to obtain a clear solution. Later, the tube was closed and put into a 70 °C water bath. A slow, gradual cooling process was performed over 10 days, reaching a temperature of 24 °C on the 10th day. At the end, a rotary evaporator was used.

2.2. HRMS

HRMS measurements were performed using Synapt G2-Si mass spectrometer (Waters) equipped with an ESI source and quadrupole-time-of-flight mass analyser. The mass spectrometer was operated in the positive ion detection mode. The optimized source parameters were: capillary voltage 3.0 kV, cone voltage 50 V, source temperature 110 °C, desolvation gas (nitrogen) flow rate 650 L/h with the temperature 450 °C, nebulizer gas pressure 6.5 bar. All samples were dissolved in water–methanol solution (1:1) and infused through a standard electrospray ion source into the instrument. The scan range was m/z 500–4000 and the acquisition method run time was 2 min. Mass calibration was performed using a cesium iodide solution. To ensure accurate mass measurements, data were collected in centroid mode and mass was corrected during acquisition using leucine enkephalin solution as an external reference (Lock-SprayTM) which generated reference ion at m/z 556.2771 Da ($[M + H]+$) in positive ESI mode. The results of the measurements were processed using the MassLynx 4.1 software (Waters) incorporated with the instrument.

2.3. Phase solubility studies

The UV–visible (UV–Vis) spectrophotometer (BioBase BK-S380, China) was utilized to assess the properties of EST and its inclusion complex. The EST showed a visible absorption peak at 280 nm. To generate a calibration curve, five standard solutions of EST in methanol were measured, each replicated three times. The concentrations of the standards used were 0.1, 0.25, 0.50, 0.75, and 1.00 mM.

Phase solubility studies were carried out following the procedure

outlined by Higuchi & Connors (1965) [17]. An excess amount of EST (50 mg) was added to 10 mL of deionized water containing various concentrations ranging from 0.20 to 20.00 mM for β-CD. The mixtures were then subjected to agitation using an orbital shaker (PHOENIX Instrument Laboratory Shaker RS-OS 5; Berlin, Germany) at 25 °C for 48 h to reach equilibrium. Afterwards, the collected samples were filtered through a 0.45 μm filter and assayed using a UV–Vis spectrophotometer at 280 nm.

2.4. QM calculations

All of the QM calculations were performed using the Gaussian 16 software [18]. All electron DFT computations were done employing the 6–311G(d,p) basis set and B3LYP or M062X functional, while the semiempirical calculations have been done using PM6 and PM7 approaches. Those four methods were used either with or without Grimme's dispersion force corrections (D3). All of the calculations were performed either in vacuo or using one of the implicit solvation models: Polarizable Continuum Model (PCM) [19] or SMD (Solvation Model Density) [20], each time choosing water as the solvent with dielectric constant 78.540. For the details of computational models, please see Table 1 and Table 2. At the review stage, additional calculations have been performed at the ωB97X-D/6–31G(d,p)-PCM-Water level [21].

Vibrational frequencies were calculated to estimate thermodynamic parameters, including Zero Point Vibrational Energy (ZPVE) and Gibbs free energy (ΔG) at 298.15 K and 101.325 kPa.

According to our recent review, different types of QM approaches are commonly used when modeling CDs inclusion complexes. In this work we decided to use those methods which are the most commonly encountered in the recent literature. The goal was to compare the geometrically optimized 1:2 and 1:1 systems from a quantitative and qualitative perspective. For this reason, we have analyzed the structural aspects as well as values of the energy ΔE and Gibbs free energy ΔG defined as

$$\Delta E = E_{com} - (E_{EST} + nE_{CD}) \quad (1)$$

where E_{com} is energy of the EST-βCD complex, E_{EST} is the energy of EST, E_{CD} is the energy of βCD and n is the number of βCD molecules forming the complex, in this case $n = 1$ or $n = 2$

$$\Delta G = G_{com} - (G_{EST} + nG) \quad (2)$$

where G_{com} is free enthalpy of the EST-βCD complex, G_{EST} is the free enthalpy of EST, G_{CD} is the free enthalpy of βCD and n is the number of βCD molecules forming the complex, in this case $n = 1$ or $n = 2$

2.5. Molecular dynamics (MD) simulations

In addition to the crystallographically determined structure of the EST/β-CD inclusion complex, which is characterized by a host: guest ratio of 2:1 and a head-to-head inclusion mode, two additional models of the same complex assuming host: guest ratio of 1:1 but two different inclusion modes (down and up) were investigated through MD simulations. The initial coordinates for these latter two complexes were generated using molecular docking with Autodock Vina [22], where the simulation boxes were defined around the coordinates of the CD centers with size 40 Å in each direction and a grid spacing of 0.375 Å. The Lamarckian genetic algorithm in AutoDockTools was utilized for this purpose, enabling effective management of numerous degrees of freedom. The 3D structure of estradiol was retrieved from the PubChem database (Compound CID: 5757) [23], while the coordinates of β-CD correspond to its crystal structure in complex with EST [15]. The docking runs were set to 10 and the produced models with the most favorable binding energies for each one of the two inclusion modes were chosen for subsequent analysis.

The AMBER 12 software package [24] was used for the simulation of the EST/β-CD inclusion complexes in a aqueous environment. Three simulations were performed. In the first case, the stating 3D model was provided by the crystallographically determined atomic coordinates of a β-CD dimer[15], which includes one EST guest molecule (site A) inside the formed dimeric cavity. Thus, the host: guest stoichiometry of the entire system in the simulation was 2: 1. In the other two cases, monomers of EST/β-CD inclusion complexes with different inclusion modes from docking analysis were used as the starting models.

The geometry of EST was optimised following the AM1BCC methodology with the program Antechamber [25]. xLeaP, the GUI version of AMBER's LeaP program, was utilized for system preparation. The GLYCAM-06j [26] force field, which is suitable for β-CD atoms' treatment and the generalized AMBER (GAFF) (for the guest molecule) were applied for the simulation. Additionally, the TIP3P water model [27] was used to solvate the CD dimer in a periodic, octahedral box forming a 12 Å thick water shell around the structure.

Minimization and MD calculations that resulted in a single trajectory of the hydrated inclusion complex system were performed with Sander. The particle mesh Ewald summation approach [28] was followed in order to handle the long-range electrostatic interactions with a 10 Å cut-off limit for the direct space sum. Hydrogen bonds were handled using the SHAKE algorithm [29]. The simulation protocol was as following: Energy minimization for hydrogens and waters using 1000 steps of steepest descent (SD) followed by 500 steps of conjugated gradient (CG) methods, while the rest non-hydrogen atoms were fixed with positional restraints of 50 kcal mol 1 Å 2. Heating equilibration up to 300 K of the water in the canonical (NVT) ensemble for 50 ps using positional restraints and the Berendsen thermostat algorithm with coupling constants of 0.5 ps to control temperature and pressure. Energy minimization of all system atoms with weak positional restraints (10 kcal mol 1 Å 2), gradual temperature increase from 5 to 300 K with 10 kcal mol-1 Å-2 restraints on the atoms of the system followed by gradual release of the restraints in successive steps at 300 K in NVT ensemble and finally density equilibration in the isobaric-isothermal (NPT) ensemble for 250 ps. Subsequently, production runs of the system under the NPT ensemble using a Berendsen-type algorithm with coupling constants of 1.0 ps were carried out under physiological conditions until reaching 12 ns.

The MD outputs were processed through the cpptraj module [30] of AMBER 12, to calculate the structural analyses (RMSD, distances, H-bonding). Moreover, the total guest binding energy ΔGbind, including the entropic term (ΔS) (calculated with the nmode module of AMBER 12) and the analysis of its components ΔEMM (changes in the gas-phase molecular mechanics (MM) energy) and ΔGsolv (solvation free energy) were computed with the aid of MM/PBSA.py script [31] implemented in AMBER 12. VMD [32] was also used for visualization and structural analyses of the MD trajectories.

Table 1
DFT computational approaches.

1	2	3	4	5	6	7	8	9	10	11	12	13
B3LYP In vacuo	M062X In vacuo	B3LYP-D3 In vacuo	M062X-D3 In vacuo	B3LYP PCM	M062 PCM	B3LYP-D3 PCM	M062X-D3 PCM	B3LYP SMD	M062X SMD	B3LYP-D3 SMD	M062X-D3 SMD	ωB97X-D PCM

Table 2
Semi-empirical computational approaches.

A	B	C	D	E	F	G	H	I	J	K	L
PM7	PM7-D3	PM6	PM6-D3	PM7	PM7-D3	PM6	PM6-D3	PM7	PM7-D3	PM6	PM6 -D3
In vacuo	In vacuo	In vacuo	In vacuo	PCM	PCM	PCM	PCM	SMD	SMD	SMD	SMD

3. Results and discussion

3.1. HRMS

As it was already mentioned, in our previous article concerning the EST-βCD complex in the solid state, we reported that the guest:host molar ratio in a crystal form is 1:2 [15]. However, we were interested whether 1:2 is the only form present in the aqueous solution of the complex, as well. Hence, we have defined 3 possible scenarios. The first option was presence of just 1:1 molar ratio complex, as it was stated many times in the literature. The second option was presence of both 1:1 and 1:2 complexes, where the 1:2 ratio would be the effect of the increasing concentration happening due to the solvent evaporation. According to that hypothesis, the amount of 1:2 complex should increase with the decrease of water content, resulting in the solely 1:2 complex in the solid state. And finally, the third possibility was presence of the identical stoichiometry as observed in a solid state which is 1:2 molar ratio.

To verify which scenario is correct, we have used the high-resolution mass spectrometry (HRMS). So far, this method has not been used often to define the molar ratio of the cyclodextrin complexes, however there are already some examples of its successful application for this purpose, published in the recent years [33,34]. The advantage of this method over the UV or fluorescence spectroscopy is that HRMS delivers a direct answer about the complex's molar ratio.

The results of the HRMS measurement of EST-βCD complex are presented in Fig. 2 and Table 3. EST forms stable associate with two βCD which can be detected in the mass spectra as ions corresponding to the protonated complex $[EST\text{-}2\beta CD{+}H]^+$ at *m/z* 2541.92 and the sodium adduct of complex $[EST\text{-}2\beta CD{+}Na]^+$ at *m/z* 2563.91. These observations suggest an interaction with 1:2 stoichiometry. No peaks corresponding to the association of one EST and one βCD were detected revealing the absence of 1:1 stable noncovalent complexes.

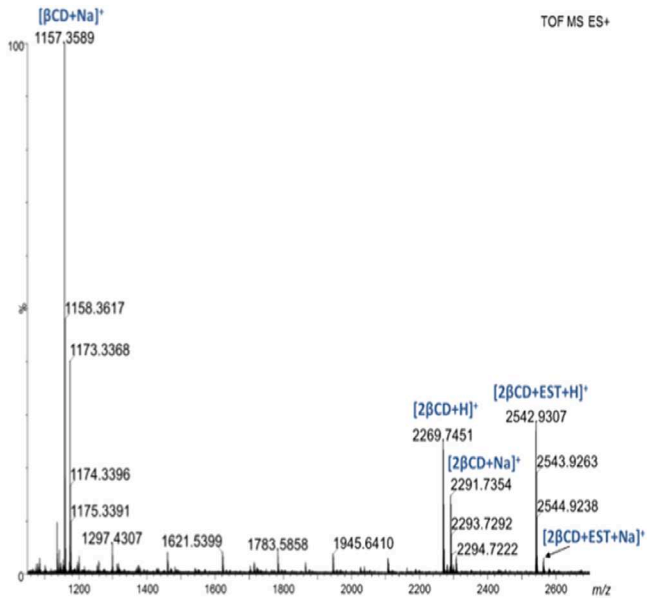


Fig. 2. HRMS spectrum of EST-βCD complex.

Table 3
The major peaks (*m/z* values and molecular formulas) from the HRMS measurement of the EST-βCD complex.

	Elemental composition	*m/z* $[M+H]^+$ or $[M{+}Na]^+$	
		Calculated	Found
$[\beta CD{+}Na]^+$	$C_{42}H_{71}O_{35}$	1157.3595	1157.3589
$[2\beta CD{+}H]^+$	$C_{84}H_{141}O_{71}$	2269.7474	2269.7451
$[2\beta CD{+}Na]^+$	$C_{84}H_{140}O_{70}Na_1$	2291.7293	2291.7354
$[2\beta CD\ {+}EST{+}H]^+$	$C_{102}H_{165}O_{72}$	2541.9250	2541.9248
$[2\beta CD\ {+}EST{+}Na]^+$	$C_{102}H_{164}O_{72}Na_1$	2563.9069	–

3.2. Phase solubility studies

The phase-solubility profile of EST in aqueous solution of successively increased β-CD concentrations at 25 °C (Fig. 3) indicates a B_S-type system which is usually observed with natural CDs, especially β-CD [34]. EST solubility increases linearly with increasing β-CD concentration in the range of 0.2 - 1.4 mM due to the formation of 1:1 EST:β-CD molar ratio inclusion complexes at the first stage. At the end of this linear portion, the maximum solubility S_{max} of EST is achieved. The solubility of the 1:1 EST:β-CD complex $S_{1:1}$ can be calculated as: $S_{1:1} = S_{max} - S_0$, where S_0 is the solubility of EST in pure water determined by the intercept of the phase solubility diagram.

Additional CD does not further increase the EST solubility and a first plateau is observed at β-CD concentration of 2 to 10 mM. This is due to the limited 1:1 complex solubility and/or the formation of 1:2 EST:β-CD molar ratio complexes. As the Gibbs phase rule indicates [35] only one discrete complex may precipitate at the plateau segment of the diagram. Thus, at even higher β-CD concentrations (12 – 20 mM), where a second plateau is observed, the solubility approximates that of the pure 1:2 complex $S_{1:2}$. The scheme of the probable complex formation is given below as eq. 4.

For the two-step association process of EST complexation with β-CD, the apparent stability constants, K_1, K_2 of the following equilibria:

$$EST + CD \overset{K_1}{\leftrightarrow} EST/CD \tag{3a}$$

$$EST/CD + CD \overset{K_2}{\leftrightarrow} EST/CD2 \tag{4a}$$

and the $K_{overall} = K_1 \cdot K_2$where estimated according to Liu et al. [35]

$$K_1 = \frac{slope}{S_0(1 - slope)} = 21,599\ M^{-1} \tag{3b}$$

$$K_2 = \frac{S_{1:2} \cdot slope}{{S_{1:1}}^2(1 - slope)} = 757\ M^{-1} \tag{4b}$$

where the slope was obtained from the linear part of the diagram and S_0, $S_{1:1}$, $S_{1:2}$ as described above. Thus, $K_{overall} = 1.6 \pm 0.4 \cdot 10^7\ M^{-2}$

This value is in the same order of magnitude to those estimated for progesterone, testosterone and cortisone by Liu et al. [36]. In that work the formation of inclusion complexes of steroids with β-CD at stoichiometric ratio of 1:2 was shown and its dependence on the steroid structure was discussed.

The *K* can be used to calculate Gibbs free energy (ΔG) according to the Eq. (5):

$$\Delta G = -RTlnK \tag{5}$$

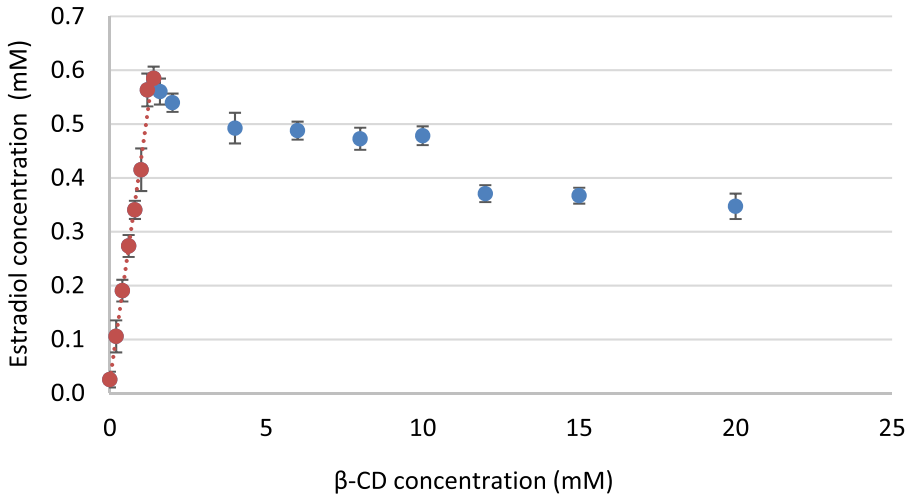


Fig. 3. Phase solubility diagram of EST/β-CD system in water at 25 °C ($n = 3$). The linear portion of the diagram (red line) was used for the calculation of K_1.

where R is gas constant (8.314 J mol^{-1} K^{-1}) and T is temperature (298 K).

This calculation reveals that the obtained here ΔG is equal to −9.92 kcal/mol.

3.3. QM calculations

As described above, the determination of this complex's stoichiometry is experimentally difficult and not straightforward task. Therefore, having in mind a huge applicability of the molecular modelling in the analysis of structure and properties of the CD complexes reviewed by us in the last year [16] we decided to apply these techniques also for this purpose. Therefore, the next step of our work was assessment of the QM approaches to check if they can properly foresee the host-guest molecular ratio and the complex association constant.

Knowledge about the crystal structure of a studied system significantly facilitates the calculations as it can be used to set the initial geometry of the complex. Thankfully, in our previous work we have determined the crystal structure of the EST-βCD complex [15]. In the current study it has been used as a starting point for all of the calculations.

Since we wanted to validate whether QM calculations can be used to predict the molar ratio and structure of the most stable complex, we have prepared 3 types of systems. The first one was 1:2 molar ratio complex and the experimental crystallographic structure of the hydrate of EST-βCD complex, after removing the water molecules, was used as an input for the computations. As EST is a molecule with quite limited conformational space, but also it is characterized by a structural anisotropy, to represent 1:1 molar ratio we needed two models called ''head up'' and ''head down'', as presented in Fig. 4. ''Head up'' is the case when EST's 5-carbon ring (steroidal D ring) goes through the wider CD's rim and the ''head down'' option is the case when it is the EST's 6-carbon ring (steroidal A ring) that protrudes through CD's wider rim. Those structures were based directly on the 1:2 experimental crystallographic data. We obtained them by removing each time different CD molecule from the 1:2 system.

3.4. DFT calculations

We have decided to apply Density Functional Theory (DFT) and semi-empirical methods. The chosen computational approaches were already presented in Table 1 and Table 2. In contrast to the majority of previously published cases of the CD complexes analysis by the means of DFT [16] where usually either 6–31 G or 6–31G(d) basis sets have been used, in this study for all DFT calculations we have used a relatively

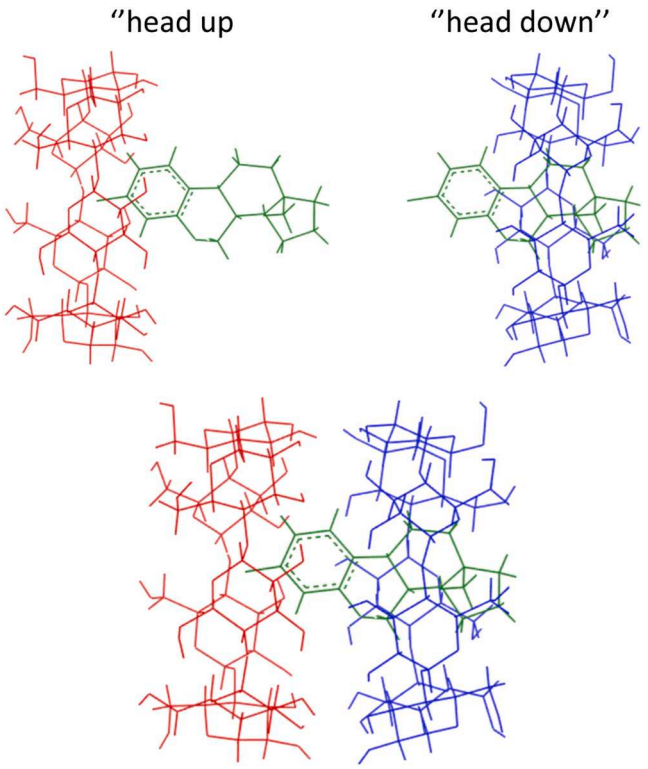


Fig. 4. Initial, non-optimized structures of 1:1 and 1:2 complexes, based on the experimental crystallographic data.

large 6–311G(d,p) basis set. Even if such approach has significantly elongated the computational time, it was a conscious choice. This way we have eliminated a risk of the influence of the basis set of the insufficient size on the obtained results.

Except of the choice of the functional in DFT (here: B3LYP and M062X) and type of the semi-empirical method (here: PM6 and PM7), one of the crucial decisions was about the type of the solvent representation. We have tested all 3 most common options: lack of solvent representation, PCM and SMD implicit solvent models.

PCM is the most often used solvent model in the computational analysis of CD complexes [37]. A different approach is presented by SMD. This model defines the free energy of solvation via two components: the one is electrostatic contribution arising from the

self-consistent reaction field, the other comes from the short-range interactions between the solute and solvent molecules [20].

Second choice of the calculations parameters was the implementation or lack of the dispersion correction. Noncovalent forces like London and van der Waals interactions are crucial for the formation and stability of the CD inclusion complexes. This aspect is not included in the calculations using exchange-correlation functionals, as the long-range electron correlation effect, known as the London part of the dispersion energy term, is not included in the Kohn–Sham DFT equation. For years this was a real issue influencing the accuracy of DFT calculations. Nowadays, different dispersion corrections are available. According to our literature research [16], in almost all studies concerning CD complexes only Grimme dispersion correction (here: D3) was used. Hence, in this study we also apply only this type of dispersion correction and compare the results to the calculations where no dispersion correction was implemented.

The results of the DFT calculations are presented in Table 4. It is clearly visible that each of three variables used in this study and described above: type of functional, solvation scheme, dispersion correction had an influence on the results.

From the energetic point of view, in 6 out of 13 cases clearly the preferred structure is the '1:1, either "head up'' (or "head down". As we know from the HRMS studies, this is not consistent with the reality. The approaches which properly predict the 1:2 complex stoichiometry are models numbered 3, 5–8, 11 and 13. On this example we can see that applying dispersion correction improve the accuracy of results when the system is treated as the in vacuo one (models 1 and 3). Simultaneously, PCM solvent model works well also when no dispersion correction has been used (Models 5 and 6) [38].

Among them there is one scheme, the scheme number 5 (B3LYP-PCM, without dispersion correction), that most significantly favours the stability of the 1:2 molar ratio complex over the 1:1 ones, both in terms of energy and free enthalpy of complexation.

3.5. Semi-empirical calculations

From the energetic point of view, the results of the semi-empirical calculations are diversified, Table 5. In 10 out 12 cases, the 1:2 molecular ratio complex has been defined as the most stable one, what stays in accordance with the experimental results. However, in those two cases where the 1:1 complex stoichiometry has been favoured, difference between the 1:1 most preferred complex and 1:2 complex is small and has the value of 0.59 kcal/mol and 3.15 kcal/mol for C and F schemes, respectively. In other words, in general, when taking into account the energetic aspect, the applied semi-empirical approaches properly predicted which molar ratio describes a complex of the highest stability.

11 out 12 models show that all three options: 1:1 head up, 1:1 head down and 1:2, are energetically stable. Only in model K the value of ΔE is positive for 1:1 head down complex. In some computational schemes differences between these three possible options are almost neglectable, for instance in model C. Whereas in other cases, the differences are much bigger, like in model L, where this difference reached up to around 37 kcal/mol.

On the contrary, from thermodynamic point of view, the majority of models pointed out the 1:1 complex as the most stable. The exceptions are models A, B, I and K. However, similarly to the ΔE results, in almost all cases, all three structural options have been defined as probable. Only in model K some of the ΔG values are positive. Model K is also the only computational scheme which shows a distant difference between both 1:1 stoichiometries and 1:2 stoichiometry and distinctively favours the 1:2 complex molar ratio. Hence, we can assume that the K computational model (PM6-SMD, without dispersion correction) predicted the experimental results in a most accurate way.

Except for this one K model, in rest of the cases the differences in values between 1:1 head up and 1:1 head down options within one computational method are similar. This suggests that there is a similar probability of creation of 1:1 head up and 1:1 head down complexes.

All those findings may be taken as a guide to create the hypothesis on the 1:2 molar ratio complex formation path. Most probably, the complexation happens in two steps. The first one, is creation of the 1:1 complex. The second step is association of the second βCD to the already existing 1:1 EST-βCD system. Which complex out of those two, 1:1 head up or 1:1 head down, is formed at the beginning can be deducted from the ΔG values. The complex characterized by a lower value should be created as the first one. According to both the most accurate DFT approach, 5 (B3LYP PCM), as well as to the MD MMGBSA results, the 1:1 head up complex is the more stable one. Therefore, it is highly probable that the complex formation occurs according to the scheme (Eq. (6)) presented below.

$$\boldsymbol{EST} + \boldsymbol{\beta CD} \rightleftharpoons \mathbf{1:1\ up} \xrightarrow{\beta CD} \mathbf{1:2} \tag{6}$$

With the first step being reversible and second irreversible reaction. Such predictions about the mechanism of the complex's formation are possible only thanks to application of the molecular modelling approach. The structural analysis of the results can be find in the Supporting Information.

Table 4
DFT calculations results. Yellow colour indicates the lowest value of ΔE within the given method (within the column). Blue colour indicates the lowest value of ΔG within the given method (within the column). ΔG and ΔE "1:2 from 1:1 up" – the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "up" is used as an initial structure for the second step; ΔG and ΔE "1:2 from 1:1 down" - the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "down" is used as an initial structure for the second step.

		B3LYP In vacuo	M062X In vacuo	B3LYP -D3 In vacuo	M062X -D3 In vacuo	B3LYP PCM	M062X PCM	B3LYP -D3 PCM	M062X -D3 PCM	B3LYP SMD	M062X SMD	B3LYP -D3 SMD	M062X -D3 SMD	ωB97X-D PCM
model	kcal/mol	1	2	3	4	5	6	7	8	9	10	11	12	13
1:1 up	ΔE	-6.47	-37.35	-54.67	-49.14	-6.93	-31.49	-30.41	-45.89	-5.06	-30.35	-48.12	-44.56	-36.80
	ΔG	8.00	-12.28	-31.73	-34.28	6.12	-8.18	-9.64	-19.64	7.44	-10.08	-24.81	-21.35	-17.91
1:1 down	ΔE	0.44	-27.64	-42.20	-47.06	1.87	-30.82	-30.88	--43.37	1.99	-30.88	-49.20	-44.17	-41.70
	ΔG	15.40	-12.10	-21.32	-23.36	16.93	-8.22	-10.9	-19.09	14.24	-11.01	-26.97	-19.61	-19.90
1:2	ΔE	2.12	-26.46	-64.81	-47.99	-61.42	-81.56	-41.75	-118.54	5.59	-20.69	-65.69	-36.78	-120.68
	ΔG	13.33	-9.18	-38.67	-26.28	-47.69	-28.08	-49.66	-34.58	15.18	-0.75	-44.76	-19.26	-65.87
1:2 from 1:1 up	ΔE	8.59	10.89	-10.14	1.15	-54.49	-50.07	-11.34	-72.65	10.65	9.66	-17.57	7.78	-83.88
	ΔG	5.33	3.1	-6.94	8	-53.81	-19.9	-40.02	-14.94	7.74	9.33	-19.95	2.09	-47.96
1:2 from 1:1 down	ΔE	1.68	1.18	-22.61	-0.93	-63.29	-50.74	-10.87	-161.91	3.6	10.19	-16.49	7.39	-78.98
	ΔG	-2.07	2.92	-17.35	-2.92	-64.62	-19.86	-38.76	-15.49	0.94	10.26	-17.79	0.35	-45.97

Table 5
Semi-empirical calculations results. Yellow colour indicates the lowest value of ΔE within the given method (within the column). Blue colour indicates the lowest value of ΔG within the given method (within the column). ΔG and ΔE "1:2 from 1:1 up" – the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "up" is used as an initial structure for the second step; ΔG and ΔE "1:2 from 1:1 down" - the energy and the Gibbs free energy change of the formation of 1:2 complex when the 1:1 complex orientation "down" is used as an initial structure for the second step.

		PM7 In vacuo	PM7 -D3 In vacuo	PM6 In vacuo	PM6 -D3 In vacuo	PM7 PCM	PM7 -D3 PCM	PM6 PCM	PM6 -D3 PCM	PM7 SMD	PM7 -D3 SMD	PM6 SMD	PM6 -D3 SMD
model	kcal/mol	A	B	C	D	E	F	G	H	I	J	K	L
1:1 up	ΔE	-59.50	-40.53	-11.68	-37.84	-54.53	-35.59	-7.45	-34.91	-51.21	-32.38	-5.08	-24.27
	ΔG	-37.13	-27.15	-38.64	-54.75	-30.07	-30.93	-10.23	-43.58	-34.98	-17.14	7.82	-24.52
1:1 down	ΔE	-54.76	-38.20	-11.05	-35.41	-43.27	-38.76	-8.96	-35.52	-47.85	-34.58	86.58	-27.98
	ΔG	-37.14	-43.81	-25.83	-31.87	-25.31	-42.85	-5.93	-53.45	-41.29	-18.23	9.60	-54.32
1:2	ΔE	-84.25	-61.17	-11.09	-61.45	-60.84	-35.61	-18.59	-54.94	-61.54	-40.77	-19.91	-64.91
	ΔG	-71.81	-48.48	5.34	-41.83	-43.04	-9.13	0.62	-34.81	-48.49	-23.39	-3.33	-39.68
1:2 from 1:1 up	ΔE	-24.75	-20.64	0.59	-23.61	-6.31	-0.02	-11.14	-20.03	-10.33	-8.39	-14.83	-40.64
	ΔG	-34.68	-21.33	43.98	12.92	-12.97	21.8	10.85	8.77	-13.51	-6.25	-11.15	-15.16
1:2 from 1:1 down	ΔE	-29.49	-22.97	-0.04	-26.04	-17.57	3.15	-9.63	-19.42	-13.69	-6.19	-106.49	-36.93
	ΔG	-34.67	-4.67	31.17	-9.96	-17.73	33.72	6.55	18.64	-7.2	-5.16	-12.93	14.64

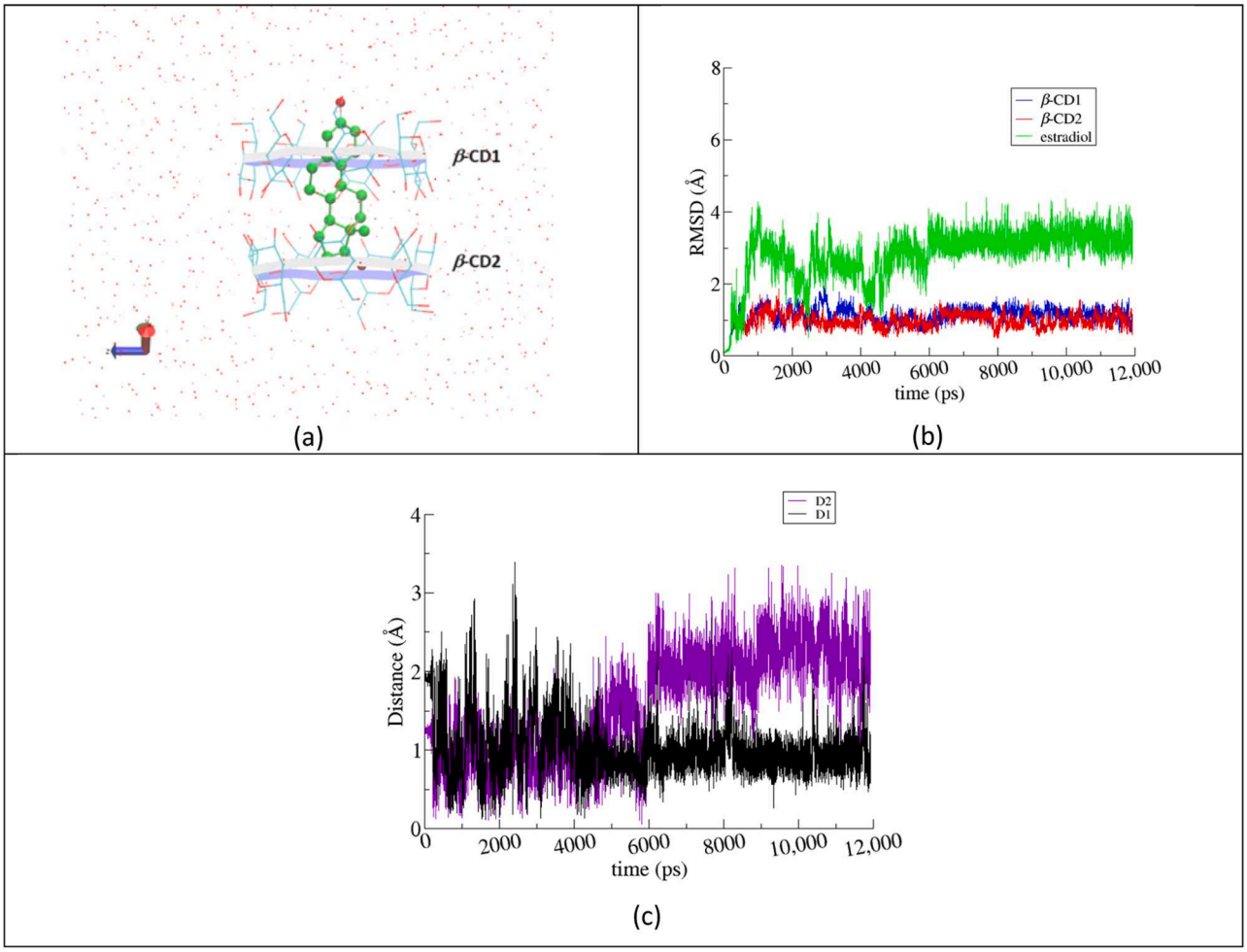


Fig. 5. (a) The starting model of the 1:2 inclusion complex based on the crystallographically determined atomic coordinates (CSD Refcode: OFANUI). (b) RMSD evolution of the host and guest molecules of the complex between the simulated states and the first frame of the simulations. (c) Distance D1 between the center of mass (COM) of the steroidal A-ring of EST and the O4*n* atom mean plane of the host β-CD1; distance D2 between COM of the steroidal d-ring of EST and the O4*n* atom mean plane of the other host of the dimeric cavity, β-CD2.

3.6. Computational thermodynamic results vs experimental data

As described in the previous section, obtention of the experimental complex stability constant allowed to define ΔG, which in case of EST-βCD complex is equal to −9.92 kcal/mol. Among the tested QM approaches, the ones which favour the most the 1:2 molar ratio are those in which the PCM correction has been applied However, the calculated ΔG overestimate the experimental ones Interestingly, the values close to the experimental one (−9.92 kcal/mol) have been obtained using M062X in vacuo (−9.18 kcal/mol) and PM7-D3 PCM approaches (−9.13 kcal/mol). Unfortunately, at the same time, both of them suggest that the 1:1 is more stable ratio than 1:2. Here, it should also be noted that β-CD is surrounded by hydration waters, which could be included in the equation of the complex formation. It was shown previously that an agreement between theoretical and experimental entropy data for inclusion complex formation was only attained when explicit water molecules were included [39].

3.7. MD simulations

MD simulations in aqueous media were carried out for the 3 types of systems, i.e. the crystallographically determined 1:2 guest:host complex and the two 1:1 monomeric complexes of opposite EST accommodation in the β-CD cavity (noted as "head up" and "head down" in Fig. 7). By monitoring the frames during the time interval of the simulations the following observations were made:

In the case of the 1:2 complex, where the starting model was retrieved from the crystal structure (Fig. 5a), the β-CD dimer encapsulating an EST molecule is preserved in the time frame of the simulation. In the absence of crystal contacts and in the presence of the surrounding water molecules, the guest EST rotates around and moves along the 7-fold molecular β-CD axis. However, it was observed a clear tendency of the guest's steroidal A-ring to be accommodated near the narrow rim of β-CD and its d-ring closer to the dimeric interface region. The measured distance between the center of mass (COM) of the A-ring and the O4n atoms mean plane of the host β-CD1, in whose cavity the A-ring is located, fluctuates around 1 Å, whereas that between the d-ring COM and the O4n plane of the other host (β-CD2), fluctuates in the range of 1 to 3 Å (Fig. 5b and 5c).

For the 1:1 molar ratio complex, both "head up" and "head down" models were examined. In the case of the "head down" binding mode, the A-ring of the guest cannot be stabilized in the host's wide rim with the rest part of EST protruding from the narrow rim of the host. Thus, EST is swiftly displaced from its initial location, exposing its A-ring to the solvent and accommodating the d-ring in the cavity (Fig. 6a). This behavior is also reflected in the high EST mobility displayed in the respective RMSD plot (Fig 6b).

On the other hand, in the case of the "head up" binding mode, the A-ring of EST which is initially exposed to the solvent by protruding from the narrow β-CD rim, it is accommodated quickly in the narrow β-CD rim where it remains relatively stable in the time frame of the simulation (Fig. 7a). The respective RMSD plot for the molecules of the system

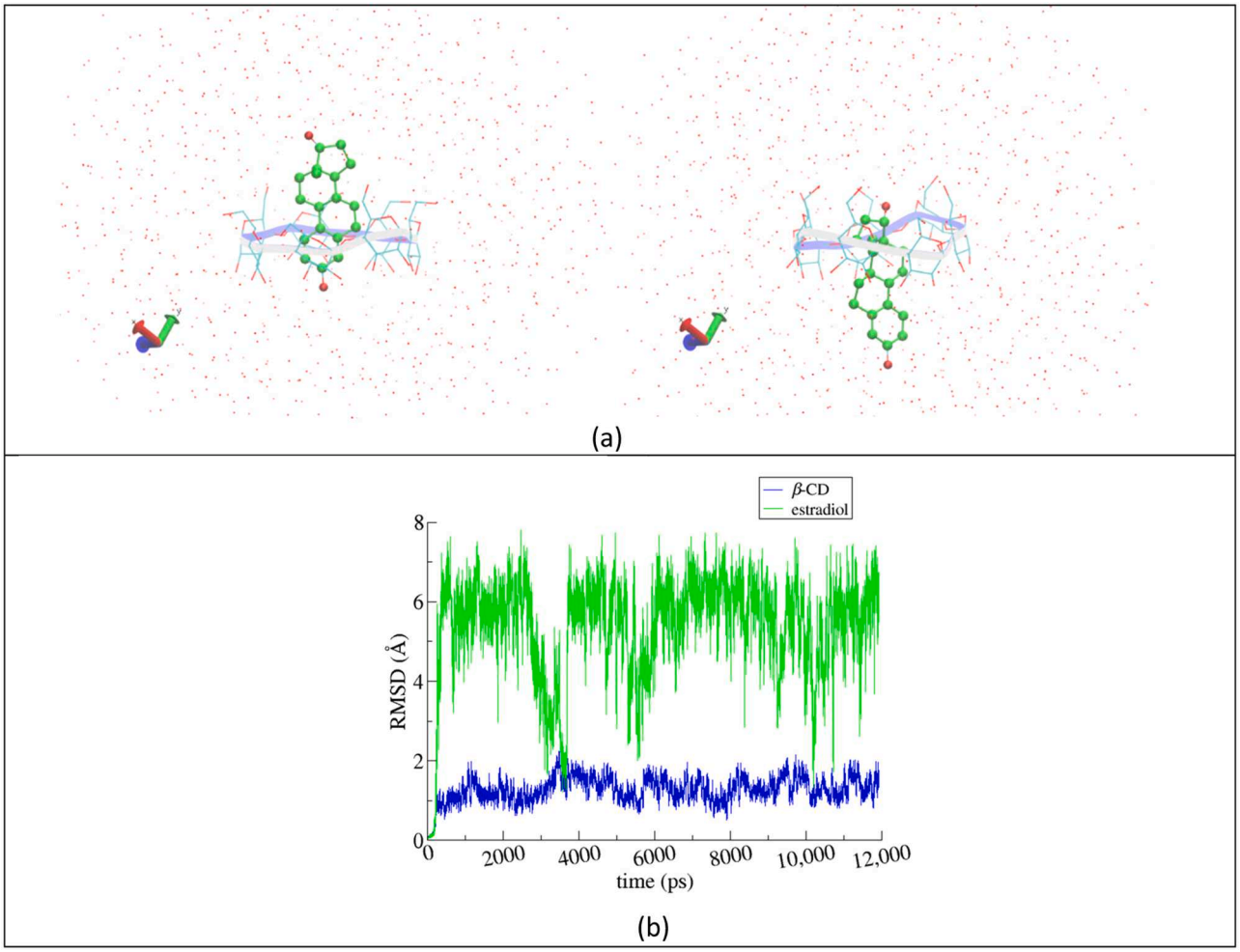


Fig. 6. (a) Two representative snapshots, at 0 and the 10th ns of the "head down" system simulation.
(b) RMSD plot for the host and guest molecule of the complex. The guest cannot be tightly stabilized in this orientation, thus exhibiting high mobility.

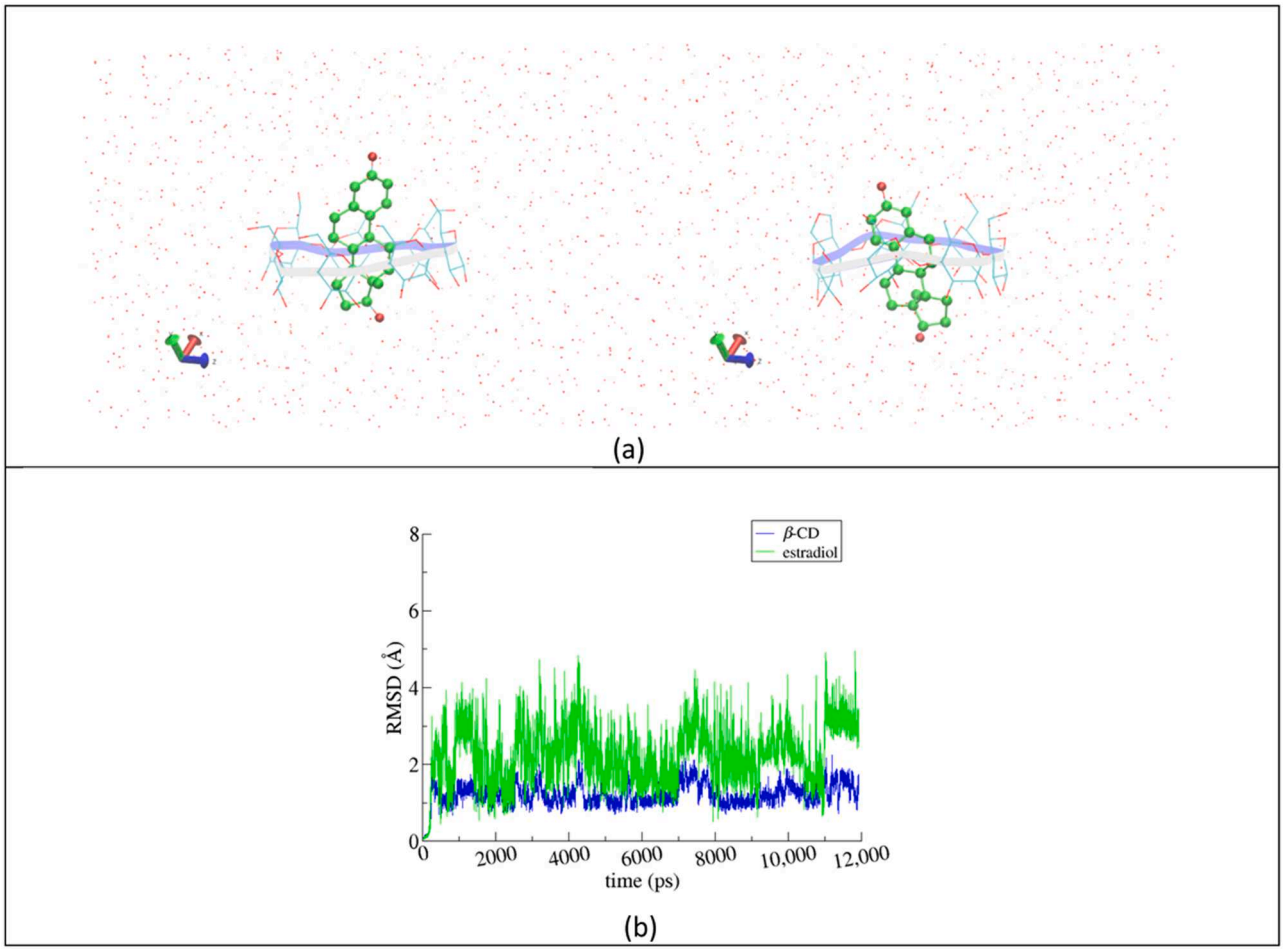


Fig. 7. (a) Two representative snapshots, at 0 and the 10th ns of the "head up" system simulation.
(b) RMSD plot the for host and guest molecule of the complex. The mobility of the guest is clearly lower than that of EST in the "head down" system simulation.

clearly shows a lower mobility of the guest compared to that of the "head down" system.

From all the above, it is concluded that the formation of a 1:1 inclusion complex of the "head up" binding mode is favored over the "head down" binding mode. Moreover, the dynamic behavior of the examined 1:2 complex, showed a very stable complex that tends to retain the accommodation of the guest (with its A-ring near the narrow rim and the d-ring near the interface of the hosts' dimer) according to that of the "head up" binding mode. These findings support the proposed complex formation described in the scheme of eq. 4.

Molecular Mechanics/Generalized Bohrn surface area (MM/GBSA) calculations [40], performed for the 3 examined types of complexes, further verify the above conclusions. The estimated host–guest binding affinities, as listed in Table 6, were extracted from 10,000 snapshots over the last 10-ns of the MD simulations. As expected, the lower ΔGbind value is estimated for the 1:2 complex mainly due to the extended van der Waals interactions between the guest and the dimeric host, that significantly decrease the averaged change of van der Waals energies (ΔEvdW) upon EST inclusion in the host dimeric cavity. By comparing the ΔGbind values estimated for the systems of common 1:1 guest:host stoichiometry ("head down" and "head up"), the considerably lower ΔGbind value of the "head up" system indicates a more stable inclusion complex.

Table 6
Binding free energies and their standard deviations (kcal/mole) resulting from MM/GBSA analysis of the inclusion compounds of EST in β-CD with guest: host ratios of 1:2, 1:1 ("head down" mode) and 1:1 ("head up" mode), respectively.

	EST/β-CD (1:2)	EST/β-CD (1:1 "head down")	EST/β-CD (1:1 "head up")
ΔE_{vdW}	−47.31 ± 2.41	−29.10±2.31	−30.46 ± 1.68
ΔE_{ele}	−4.15 ± 2.78	−2.12±2.13	−2.09 ± 1.72
ΔE_{MM} [a]	−51.46 ± 3.58	−31.23±3.21	−32.55 ± 2.38
ΔG_{GB}	23.22 ± 2.98	15.63±2.80	13.45 ± 1.99
$\Delta G_{nonpolar}$	−4.45 ± 0.22	−2.94±0.16	−3.03 ± 0.12
$\Delta G_{solvation}$ [b]	18.77 ± 2.94	12.69±2.72	10.42 ± 1.94
ΔH [c]	−32.69 ± 4.63	−18.54±2.06	−22.13 ± 1.97
$T{\cdot}\Delta S$ [d]	−19.03 ± 3.00	−16.26±1.21	−15.88 ± 1.36
ΔG_{bind} [e]	−13.66 ± 4.20	−2.27±2.39	−6.25 ± 2.40

ΔE_{vdW} = van der Waals contribution from molecular mechanics; ΔE_{ele} = electrostatic energy as calculated by the molecular mechanics force field; ΔG_{GB} = the electrostatic solvation energy (polar contribution) calculated using the GB model; $\Delta G_{nonpolar}$ = nonpolar contribution to the solvation free energy, calculated by the solvent-accessible surface area (SASA) method;.

[a] $\Delta E_{MM} = \Delta E_{vdW} + \Delta E_{ele}$;.
[b] $\Delta G_{solvation} = \Delta G_{GB} + \Delta G_{nonpolar}$;.
[c] $\Delta H = \Delta G_{solvation} + \Delta E_{MM}$;.
[d] $T{\cdot}\Delta S$ entropic term calculated by normal mode analysis;.
[e] $\Delta G_{binding} = \Delta H - T{\cdot}\Delta S$.

3.8. Computational stability results vs experimental data

As described in the previous section, acquisition of the experimental complex stability constant allowed to define ΔG, which in case of EST-βCD complex is equal to −9.92 kcal/mol. This value is within the limits of uncertainty of ΔGbind for 1:2 complex, −13.66 ± 4.20 kcal/mol (Table 6). However, the values of ΔGbind obtained for 1:1 complex,

either head-up or head-down orientations, −2.27±2.39 and −6.25 ± 2.40 respectively, although not strictly within the uncertainty limits, are also close to the experimentally determined one. Therefore, while the MD/MMGBSA method allows to properly indicate the order of the magnitude of ΔGbind, it should be supported by the experimental analysis such as HRMS to confirm the complex ratio.

4. Conclusions

Even though the existence of the EST-βCD complex in water solution has been known for decades, and despite quite some studies performed to define its molar ratio, this complex's structure remained not properly determined until now. The hypothesis that the knowledge in this context might be not complete, has occurred after determination of the EST-βCD crystal structure where the molar ratio was found to be 1:2 (EST:βCD).

In this work, thanks to application of the HRMS approach, it has been indisputably proven that the EST-βCD complex molar ratio is 1:2 and not as previously assumed 1:1. Moreover, the phase solubility studies confirmed these results. This type of experiments has been performed before, however, never the 1:2 molar ratio has been taken into account as a possible description of this system. In other words, the indisputable HRMS measurement results prompted the revision of the phase solubility studies. This allowed to properly define complex stability constant K, what in turn delivered information about the Gibbs free energy value of the complex formation.

The structure and thermodynamics of the complex were further analyzed using various QM (DFT, semi-empirical) and MM (MD/MMGBSA) approaches. Tests on the application of different computation parameters such as presence/absence of dispersion correction, choice of implicit solvent model or DFT functional, have been performed.

Possession of the credible experimental data allowed to assess the computational approaches. While some of the "static" QM methods properly indicated the correct host: guest ratio at the same time they failed to accurately predict the Gibbs free energy of complexation. On the other hand, QM methods that properly described the value of ΔG of 1:2 complex formation, such as M062X in vacuo, favored the 1:1 stoichiometry, which was experimentally excluded. The MD/MMGBSA method, although performed at the lower level of theory, accurately predicted the stability constant of the complexed but was not conclusive to indicate the formation of either 1:1 or 1:2 complex.

This leads to the conclusion that among tested computational approaches, there are some which are able to properly predict the composition of such complex and some that can assess the stability of the studied system. However, there is not single method that would allow to reproduce both the stoichiometry and the thermodynamic stability of the complex at the same time.This study finally describes the structure and thermodynamics of the EST-βCD complex in aqueous solution and delivers experiment-based information about the formation of this complex.

CRediT authorship contribution statement

Anna Helena Mazurek: Writing – original draft, Software, Investigation, Funding acquisition, Data curation, Conceptualization. **Łukasz Szeleszczuk:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Conceptualization. **Kostas Bethanis:** Writing – original draft, Methodology, Investigation, Conceptualization. **Elias Christoforides:** Software, Resources, Project administration, Methodology. **Marta Katarzyna Dudek:** Software, Methodology, Conceptualization. **Ewelina Wielgus:** Writing – review & editing, Validation, Project administration. **Dariusz Maciej Pisklak:** Writing – review & editing, Writing – original draft, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the Medical University of Warsaw, Poland [grant number WF7/1/F/MB/N/23].

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.molstruc.2024.138710.

References

[1] J. Szejtli, Introduction and general overview of cyclodextrin chemistry, Chem. Rev. 98 (1998) 1743–1754.
[2] B.G. Poulson, Q.A. Alsulami, A. Sharfalddin, E.F. El Agammy, F. Mouffouk, A.-H. Emwas, L. Jaremko, M. Jaremko, Cyclodextrins: structural, chemical, and physical properties, and applications, Polysaccharides 3 (2022) 1.
[3] G. Crini, Review: a history of cyclodextrins, Chem. Rev. 114 (2014) 10940–10975.
[4] T. Irie, K. Uekama, Pharmaceutical applications of cyclodextrins. III. Toxicological issues and safety evaluation, J. Pharm. Sci. 86 (1997) 147–162.
[5] V. Aiassa, C. Garnero, M.R. Longhi, A. Zoppi, Cyclodextrin multicomponent complexes: pharmaceutical applications, Pharmaceutics 13 (2021) 1099.
[6] S.S. Jambhekar, P. Breen, Cyclodextrins in pharmaceutical formulations I: structure and physicochemical properties, formation of complexes, and types of complex, Drug Discov. Today. 21 (2016) 356–362.
[7] S.S. Jambhekar, P. Breen, Cyclodextrins in pharmaceutical formulations II: solubilization, binding constant, and complexation efficiency, Drug Discov. Today. 21 (2016) 363–368.
[8] M.P. Thomas, B.V. Potter, The structural biology of oestrogen metabolism, J. Steroid Biochem. Mol. Biol. 137 (2013) 27–49.
[9] Salole E.G. Estradiol, Analytical profiles of drug substances, 1986, 15, 283–318.
[10] N. Sadlej-Sosnowska, Fluorometric determination of association constants of three estrogens with cyclodextrins, J Fluoresc 7 (1997) 195–200.
[11] C. Yañez, J. Basualdo, P.J. Jara-Ulloa, A. Squella, Inclusion complexes of estrone and estradiol with β-cyclodextrin: voltammetric and HPLC studies, J. Phys. Org. Chem. 20 (2007) 499–505.
[12] R.L. Pérez, G.M. Escandar, Spectrofluorimetric study of estrogen-cyclodextrin inclusion complexes in aqueous systems, Analyst 138 (2013) 1239–1248.
[13] D.H. Schwarz, A. Engelke, G. Wenz, Solubilizing steroidal drugs by β-cyclodextrin derivatives, Int J Pharm 531 (2017) 559–567.
[14] Z.Y. Lin, X.X. Wang, S.B. Kou, J.H. Shi, Exploring the inclusion interaction of estradiol with β-CD and HP-β-CD with the help of molecular dynamics simulation as well as multi-spectroscopic approaches, Spectrochim Acta A Mol Biomol Spectrosc 269 (2022) 120764.
[15] A.H. Mazurek, Ł. Szeleszczuk, K. Bethanis, E. Christoforides, M.K. Dudek, M. Zielińska-Pisklak, D.M. Pisklak, 17-β-Estradiol—B-Cyclodextrin complex as solid: synthesis, structural and physicochemical characterization, Molecules 28 (2023) 3747.
[16] A.H. Mazurek, Ł. Szeleszczuk, Current status of quantum chemical studies of cyclodextrin host–guest complexes, Molecules 27 (2022) 3874.
[17] T. Higuchi, K.A. Connors, Phase solubility techniques, Adv. Anal. Chem. Instrument. 4 (1965) 117–212. Open Access Library Available online: http://www.oalib.com/references/7163685 (accessed on 26 January 2024).
[18] Software https://gaussian.com/products/ accessed on 24th January 2024.
[19] M. Cossi, N. Rega, G. Scalmani, V. Barone, Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model, J. Comput. Chem. 24 (2003) 669–681.
[20] A.V. Marenich, C.J. Cramer, D.G Truhlar, Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions, J. Phys. Chem. B 113 (2009) 6378–6396.
[21] J.-D. Chai, M. Head-Gordon, Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections, Phys. Chem. Chem. Phys. 10 (2008) 6615.
[22] J. Eberhardt, D. Santos-Martins, A.F. Tillack, S. Forli, AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings, J. Chem. Inf. Model. 61 (2021) 3891–3898.
[23] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B.A. Shoemaker, P. A. Thiessen, B. Yu, et al., PubChem 2023 update, Nucleic Acids Res. 51 (2023) D1373–D1380.
[24] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the amber biomolecular simulation package, Wiley Interdiscip. Rev.: Comput. Mol. Sci. 3 (2012) 198–210.

[25] J. Wang, W. Wang, P. Kollman, D. Case, ANTECHAMBER: an accessory software package for molecular mechanical calculations, J. Chem. Inf. Comput. Sci. - JCISD (2000) 222.
[26] K.N. Kirschner, A.B. Yongye, S.M. Tschampel, J. Gonzalez-Outeirino, C.R. Daniels, B.L. Foley, R.J. Woods, GLYCAM06: a generalizable biomolecular force field. Carbohydrates, J. Comput. Chem. 29 (2008) 622–655.
[27] P. Mark, L. Nilsson, Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K, J. Phys. Chem. A 105 (2001) 9954–9960.
[28] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[29] V. Kräutler, W.F. van Gunsteren, P.H. Hünenberger, A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations, J. Comput. Chem. 22 (2001) 501–508.
[30] D.R. Roe, T.E. Cheatham 3rd, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theory Comput. 9 (2013) 3084–3095.
[31] B.R. Miller 3rd, T.D.J. McGee, J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBSA.Py: an efficient program for end-state free energy calculations, J. Chem. Theory Comput. 8 (2012) 3314–3321.
[32] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (33–38) (1996) 27–28.
[33] G. Zengin, Yildiztugay E. Nilofar, A. Bouyahya, H. Cavusoglu, R. Gevrenova, D. Zheleva-Dimitrova, A comparative study on UHPLC-HRMS profiles and biological activities of inula sarana different extracts and its beta-cyclodextrin complex: effective insights for novel applications, Antioxidants (Basel) 12 (2023) 1842.
[34] T.K. Špehar, M. Pocrnić, D. Klarić, B. Bertoša, A. Čikoš, M. Jug, J. Padovan, S. Dragojević, N. Galić, Investigation of praziquantel/cyclodextrin inclusion complexation by NMR and LC-HRMS/MS: mechanism, solubility, chemical stability, and degradation products, Mol. Pharm 18 (2021) 4210–4223.
[35] M.E. Brewster, T. Loftsson, Cyclodextrins as pharmaceutical solubilizers, Adv. Drug Deliv. Rev. 59 (2007) 645–666.
[36] F.Y. Liu, D.O. Kildsig, A.K. Mitra, Beta-cyclodextrin/steroid complexation: effect of steroid structure on association equilibria, Pharm. Res 7 (1990) 869–873.
[37] M. Cossi, N. Rega, G. Scalmani, V. Barone, Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model, J. Comput. Chem. 24 (2003) 669–681.
[38] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132 (2010) 154104.
[39] J.F. Lopes, Nascimento C.S. Jr, C.P.A. Anconi, H.F.D. Santos, W.B Almeida, Inclusion complex thermodynamics: the β-cyclodextrin and sertraline complex example, J. Mol. Graph. Model 62 (2015) 11–17.
[40] S. Genheden, U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, Expert Opin. Drug Discov. 10 (2015) 449–461.

molecules

MDPI

Article

17-β-Estradiol—β-Cyclodextrin Complex as Solid: Synthesis, Structural and Physicochemical Characterization

Anna Helena Mazurek [1,2], Łukasz Szeleszczuk [1,*], Kostas Bethanis [3], Elias Christoforides [3], Marta Katarzyna Dudek [4], Monika Zielińska-Pisklak [5] and Dariusz Maciej Pisklak [1]

1 Department of Organic and Physical Chemistry, Faculty of Pharmacy, Medical University of Warsaw, Banacha 1 Str., 02-093 Warsaw, Poland; anna.mazurek@wum.edu.pl (A.H.M.)
2 Doctoral School, Medical University of Warsaw, Żwirki i Wigury 81 Str., 02-093 Warsaw, Poland
3 Laboratory of Physics, Department of Biotechnology, Agricultural University of Athens, 11855 Athens, Greece
4 Structural Studies Department, Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, Sienkiewicza 112 Str., 90-363 Łódź, Poland
5 Department of Pharmaceutical and Biomaterials Chemistry, Faculty of Pharmacy, Medical University of Warsaw, Banacha 1 Str., 02-093 Warsaw, Poland
* Correspondence: lukasz.szeleszczuk@wum.edu.pl; Tel.: +48-501-255-121

Abstract: 17-β-estradiol (EST) is the most potent form of naturally occurring estrogens; therefore, it has found a wide pharmaceutical application. The major problem associated with the use of EST is its very low water solubility, resulting in poor oral bioavailability. To overcome this drawback, a complexation with cyclodextrins (CD) has been suggested as a solution. In this work, the host–guest inclusion complex between the ß-CD and EST has been prepared using four different methods. The obtained samples have been deeply characterized using ^{13}C CP MAS solid state NMR, PXRD, FT-IR, TGA, DSC, and SEM. Using SCXRD, the crystal structure of the complex has been determined, being to the best of our knowledge the first solved crystal structure of an estrogen/CD complex. The periodic DFT calculations of NMR properties using GIPAW were found to be particularly helpful in the analysis of disorder in the solid state and interpretation of experimental NMR results. This work highlights the importance of a combined ssNMR/SCXRD approach to studying the structure of the inclusion complexes formed by cyclodextrins.

Keywords: cyclodextrin; estradiol; DFT; SCXRD; solid state NMR

Citation: Mazurek, A.H.; Szeleszczuk, Ł.; Bethanis, K.; Christoforides, E.; Dudek, M.K.; Zielińska-Pisklak, M.; Pisklak, D.M. 17-β-Estradiol—β-Cyclodextrin Complex as Solid: Synthesis, Structural and Physicochemical Characterization. *Molecules* **2023**, *28*, 3747. https://doi.org/10.3390/molecules28093747

Academic Editors: Rosa Iacovino, Marina Isidori and Margherita Lavorgna

Received: 31 March 2023
Revised: 20 April 2023
Accepted: 24 April 2023
Published: 26 April 2023

1. Introduction

17-β-estradiol, EST (Figure 1), is the most potent form of naturally occurring estrogens [1]; therefore, it has found wide application in hormonal contraception, hormone replacement therapy (HRT), and treatment of menopausal and postmenopausal symptoms [2]. Oral administration of EST in a solid dosage form is the most favorable form of HRT [3]. While in the European Pharmacopoeia only the hemihydrate form of EST is described, recently its anhydrous form was successfully obtained [4]. Moreover, numerous cocrystals of EST have been designed [5,6] to solve one of the major problems associated with the application of EST: its poor oral bioavailability caused by very low water solubility—0.2–5 μg mL^{-1} [7].

Cyclodextrins (CDs) are cyclic oligosaccharides consisting of a macrocyclic ring formed by glucose subunits joined by α-1,4 glycosidic bonds. CDs are primarily used in pharmaceutical formulations due to their unique properties, resulting in their ability to form inclusion complexes [8]. The desirable properties of CDs in the pharmaceutical field can be explained at the molecular level. CD molecules resemble a "doughnut" ring, in which small, non-polar substances such as EST can be entrapped. The external fragments of CD molecules are polar due to the presence of hydroxyl groups. When a non-polar substance (e.g., an EST) enters the molecular hole of cyclodextrin, the formed host–guest complex

is polar (at outside) and, therefore, is more soluble than the separated guest molecule. Therefore, CDs are commonly used in pharmaceutical formulations as they are able to increase the solubility of APIs, protect them against external factors, such as light, humidity, and heat, or can even mask unpleasant smells or flavors of drugs. Currently, more than 100 original drugs are manufactured with CDs as excipients [9–11].

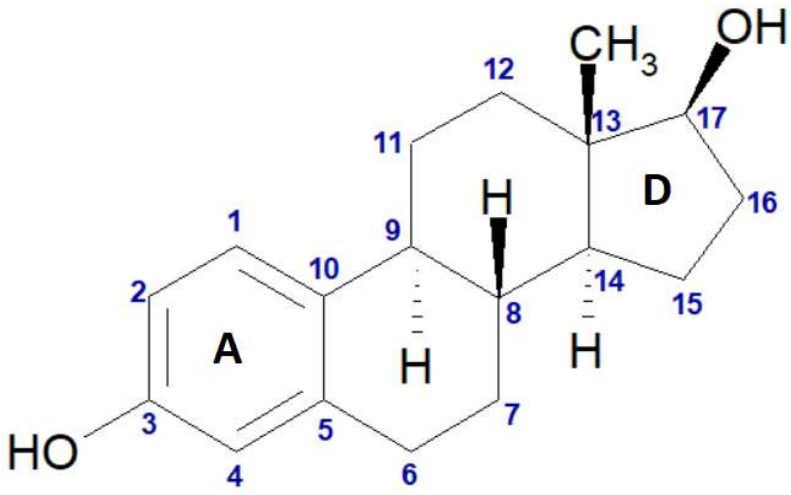


Figure 1. Chemical structure of 17-β-estradiol (EST) with atom numbering. "A" and "D" represent the symbols of the particular rings within the structure.

Multiple preparation methods for CD inclusion complexes are being exploited, such as the solvent evaporation method, grinding method, ultrasonic method, and freeze-drying method. It has been shown in many examples that the method of complex preparation may have a major impact on the obtained form of the final product [12–14].

Only a small amount of CD complexes have been reported with their crystal structures. This is caused by the fact that many of these complexes are either amorphous or polycrystalline, and even for the crystalline complexes, it is usually very hard to obtain a crystal of a size suitable for single-crystal X-ray measurements [15]. However, in order to fully understand the aforementioned changes resulting from complexation, knowledge of the molecular structure of CD complexes is crucial.

To achieve this goal—that is, to understand the structure and dynamics of the CD-based complexes—multiple computational and analytical methods are usually applied, some of which have been recently reviewed by us [16–18]. Among the analytical methods most commonly used to study these kinds of materials in a solid state are Fourier-transform infrared spectroscopy (FT-IR) and powder X-ray diffraction (PXRD), together with thermo-analytical techniques such as differential scanning calorimetry (DSC) and thermogravimetric analysis (TGA). Moreover, the application of solid-state nuclear magnetic resonance (ssNMR) can provide essential information, unobtainable by other methods [16]. From the theoretical approaches, the most important and accurate methods are obtained through the use of quantum chemical calculations, usually at the density functional theory (DFT) level [19]. A recent review [17] revealed that the application of quantum chemical calculations in studies of CD complexes can be essential, providing results unobtainable by any other method, both experimental and computational. In particular, density functional theory (DFT) methods are among the most accurate and most frequently used to model such systems, with the PBE functional being the method of choice when modeling solid state structures. However, to the best of our knowledge, such computations have not been performed on CD inclusion complexes yet.

The properties and structures of the complexes of EST with various CDs have been studied extensively for the last 30 years [20–27]. For example, in a recent work [28], single-crystal X-ray diffraction (SCXRD) results of the EST/ß-CD complex were presented, determining the unit cell and the crystallographic space group of the crystal structure. In

this work, only the atomic coordinates of the host molecule were determined, whereas the encapsulated hormone was not possible to be modeled due to its high disorder. Thus, although a 2:1 host:guest stoichiometry of the complex was estimated (based on the residual density calculated by that incomplete model), the structural information provided was limited to the host molecular arrangement in the crystalline state. On the other hand, in the present work, the crystal structure of the EST/ß-CD complex was fully determined and the atomic positions of both host and guest molecules reveal the orientation of the guest in the ß-CD host dimeric cavity and give valuable information about the intermolecular (host–guest and guest–guest) interactions and the arrangement of the full complex units in the crystalline state.

The aim of our work was to obtain the 17-β-estradiol–β-cyclodextrin complex (EST/ß-CD) by means of four different preparation methods. The obtained samples have been extensively characterized by means of various analytical (ssNMR, FT-IR, SCXRD, PXRD, SEM, DSC, TGA) and computational (periodic DFT) methods to extend knowledge of this complex formation and to study how the method of preparation influences the final results.

2. Results and Discussion

2.1. SCXRD Results

As was mentioned above, a single crystal analysis of the EST/ß-CD complex has been presented in a previous work [28] where only the host and the water oxygen atoms were located via the collected diffraction data using a Mo X-ray source, whereas the encapsulated hormone was not possible to be modeled as the residual electron density appearing within the host cavity was very low ($\Delta\rho \leq 1$ e·Å^3) due to the high disorder of the included estradiol. Similar to that work, the EST/ß-CD crystal structure presented here was found to belong to the monoclinic system with space group C2, with roughly the same lattice parameters (Table 1). However, in addition to the host and water oxygen atoms, the coordinates of the encapsulated estradiol atoms were successfully determined. This is likely due to the higher diffracted intensities collected by using a Cu *Ka* X-ray source, resulting in significant higher and discrete difference electron density peaks that allowed for the modeling of the disordered guest.

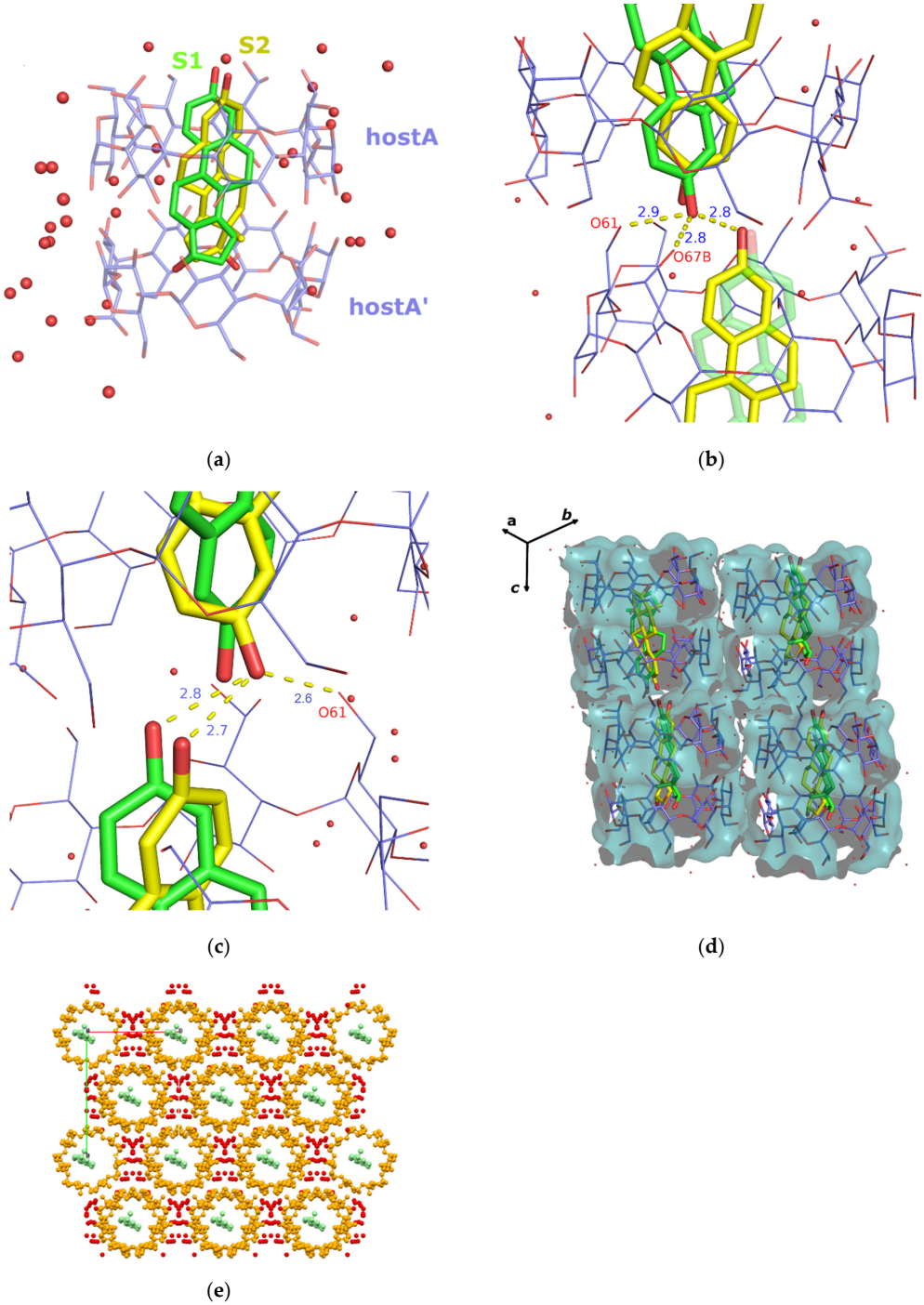
The determined asymmetric unit of the EST/ß-CD crystal structure contains one host, one guest (with s.o.f. of 0.5), and 10.5 water oxygens distributed over 17 sites (no water hydrogens were included). As the complex crystallizes in the C2 space group, the host:guest stoichiometry is 2:1, with two symmetry-related hosts (denoted as hostA and hostA') forming a classic "head-to-head dimer" stabilized by the well-known intermolecular hydrogen bonds between their secondary hydroxyls. The geometric features of the host molecule are reported in Supplementary Table S1, indicating that β-CD, upon complexation with estradiol, adopts the usual torus-like macrocycle shape and the round conformation due to the formation of the commonly observed intramolecular interglucose O3(n)-H···O2(n+1) hydrogen bonds. The encapsulated estradiol molecule is found disordered over two sites (site S1 and site S2, with s.o.f of 0.25 each). Both occupied sites have the same orientation: the guest is accommodated "axially" inside the dimeric β-CD cavity, with the hydroxyl of its A ring protruding from the primary rim of the host β-CD and its D ring being buried into the dimeric hydrophobic cavity (Figure 2a). More specifically, the mean plane of the estradiol aromatic ring system forms an angle of 95.795 (2)° and 82.404 (7)° in the case of the S1 and S2 occupied sites, respectively, with the mean plane of the glucosidic O4n atoms of the hosts. The oxygen atom of the guest's A ring hydroxyl is located at a distance of 0.638 (5) Å for S1 and 0.428 (3) Å for S2 above the mean plane of the O6n atoms of hostA, whereas the oxygen of the guest's D ring hydroxyl is found near the O4n atom plane of hostA' (the distance between the oxygen and the mean plane of the O4n atoms of hostA' being 0.1642 (12) Å and 0.112 (8) Å for S1 and S2, respectively). The protruding hydroxyl of the guest's A ring is at hydrogen bond distance from the primary hydroxyls of the host and the protruding hydroxyl of the guest of the consecutive complex unit. In particular, the A

ring hydroxyl of the guest occupying the S1 site can be hydrogen bonded with the fully occupied O(61)H, the 40% occupied O(67B)H of the host (1 − x, y, 1 − z), and the guest's A ring hydroxyl occupying the S2(1 − x, y, 1 − z) site (Table S2 and Figure 2b). However, as the distance between the two consecutive S1 and S1(1 − x, y, 1 − z) sites is just 2.4 Å, the guest is sterically forbidden to occupy the S1 site in two successive complex units. On the other hand, the A ring hydroxyl of the S2 site can be hydrogen bonded with the O(61)H of the host and the guest occupying the S2 site in the successive (1 − x, y, 1 − z) complex unit (Figure 2c). Thus, the arrangement of the encapsulated estradiol in the consequent dimers could be that of S1-S2(1 − x, y, 1 − z) or S2-S2(1 − x, y, 1 − z) but not S1-S1(1 − x, y, 1 − z) (left column of Figure 2d), whereas the S1-S1(x, y, −1 + z), S1-S2(x, y, −1 + z), and S2-S2(x, y, −1 + z) (right column of Figure 2d) arrangement is also possible.

Table 1. Crystallographic parameters of the EST/ß-CD inclusion complex [29].

Crystal Data	**EST/ß-CD**
CCDC No.	2250781
Complex formula in the asymmetric unit	$(C_{42}H_{70}O_{35})\cdot 0.5(C_{18}H_{24}O_2)\cdot 10.5H_2O$
Formula weight	1439.16
Crystal system, space group	Monoclinic, *C*2
Temperature (K)	100
a, *b*, *c* (Å)	19.1245 (15), 24.4180 (18), 15.6004 (11)
α, β, γ (°)	109.500 (5)
V (Å^3)	6867.2 (9)
Z	4
Radiation type	Cu *Ka*
μ (mm^{-1})	1.09
Crystal size (mm^3)	0.4 × 0.27 × 0.13
Data collection	
T_{min}, T_{max}	0.593, 0.754
No. of measured, independent, and observed [I > 2σ(I)] reflections	105,516, 11,945, 10,450
R_{int}	0.074
$(\sin \theta/\lambda)_{max}$ (Å^{-1})	0.595
Refinement	
R_1 [$F^2 > 2\sigma(F^2)$], $wR_2(F^2)$, GooF	0.095, 0.269, 1.04
No. of reflections	11,945
No. of parameters	950
No. of restraints	202
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.78, −0.44

The "head-to-head" β-CD dimers are arranged according to the channel (CH) packing mode (Figure 2) along the c-axis, the distance and the shift between the centroids of two successive dimers being 15.600(4) and 3.0647(7) Å, respectively. The adjacent channels are stacked via bridge water molecules and host–host intermolecular hydrogen bonds (Figure 2e).



(**a**) (**b**)

(**c**) (**d**)

(**e**)

Figure 2. (**a**) Estradiol disordered over two sites (S1 green and S2 yellow) is fully encapsulated inside the cavity of a classic "head-to-head" β-CD host dimer with a 2:1 host:guest stoichiometry. (**b**) Hydrogen bonds between the protruding oxygen atom of the A ring hydroxyl of the guest occupying the S1 site (green) and primary host–guest (occupying the S2 site) of the consecutive complex unit. (**c**) Hydrogen bonds between the guest occupying the S2 site (yellow) and the host–guest (both occupied sites) of the consecutive complex unit. (**d**) β-CD dimers forming channels along the c-axis. Left-hand side: S1-S2(1 − x, y, 1 − z) and S2-S2(1 − x, y, 1 − z) possible arrangements of the encapsulated estradiol molecules in the channel. Right-hand side: S1-S1(x, y, −1 + z), S1-S2(x, y, −1 + z), and S2-S2(x, y, −1 + z) possible arrangements of the encapsulated estradiol molecules in the channel. (**e**) Crystal packing of the EST/β-CD complex view perpendicular to the ab plane. In all figures, hydrogen atoms are omitted for clarity.

2.2. Periodic DFT Calculation Results

As described in Section 2.1 (SCXRD results), due to the disorder in the atomic positions of EST in the crystal structure, two significantly different orientations of the neighboring guest molecules exist in the solid state. Therefore, using the experimental crystal structure and choosing the proper guest molecules, we have created two model periodic structures for the DFT calculations. In the first one, the A-rings of neighboring EST molecules are located close to each other (Figure 3). For the purposes of this study, we have named this structure DAAD. This structure can be also found in the Supplementary Material ESI as DAAD.cif.

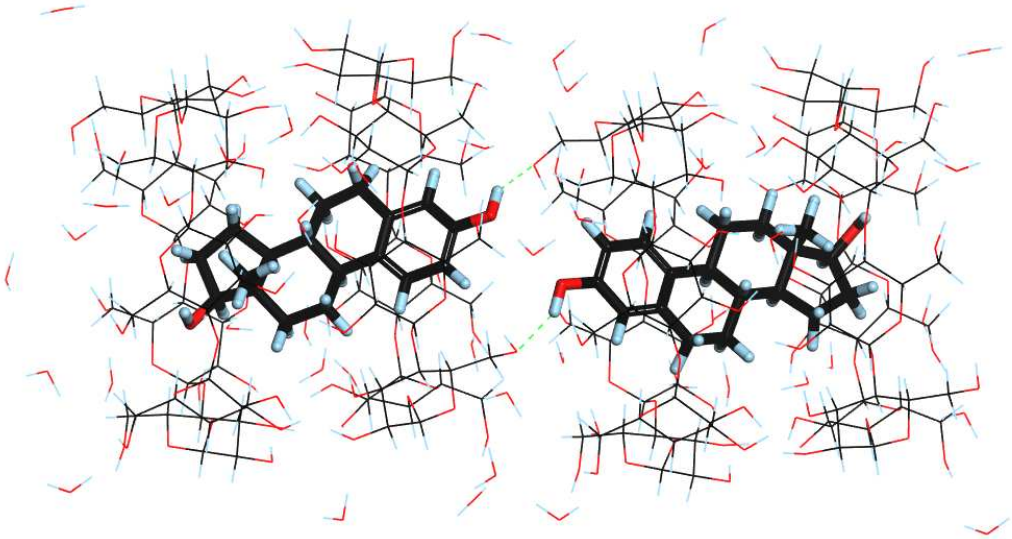


Figure 3. DFT-optimized structure of DAAD. Hydrogen bonds formed between EST and BCD are indicated as green dashed lines.

The other possible orientation was named ADAD, as in this one, the D ring of one molecule of EST is always located next to the A ring of the second EST molecule (Figure 4). This structure can also be found in the Supplementary Material as ADAD.cif. The unit cell dimensions of both DAAD and ADAD were exactly the same and can be found in Table 2.

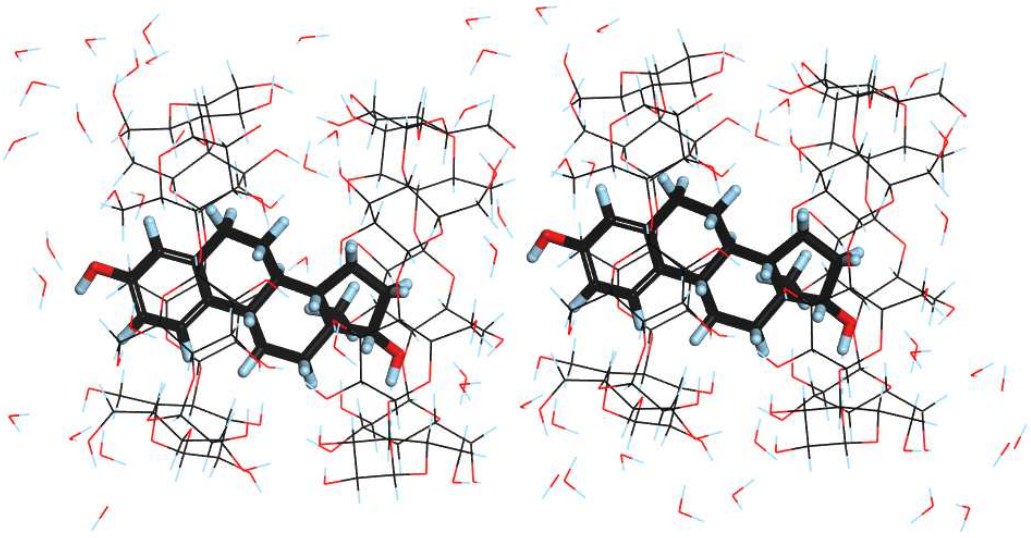


Figure 4. DFT-optimized structure of ADAD.

Table 2. Unit cell parameters and relative energy for the modeled structures, ADAD and DAAD, compared with the initial experimental structure used for geometry optimization (Exp.).

	Exp.	ADAD	DAAD
a [Å]	15.515011	15.742737	15.484059
b [Å]	15.515011	15.314397	15.484028
c [Å]	31.188400	31.718429	31.699391
α [°]	101.86701	100.93157	101.31461
β [°]	101.86701	99.83545	101.31466
γ [°]	103.84663	104.42976	101.44811
Relative energy [kcal/mol]		0	−5.152

It should be noted here that the DFT calculations performed in this study were done directly on the crystal structures, taking into account the periodicity of the studied system and, explicitly, water molecules. In addition, the initial structures for the geometry optimization calculations were taken directly from the SCXRD measurements, without adding any additional atoms. Therefore, the preparation of the structures for the calculations included solely removing chosen guest molecules and, in some cases, the water molecules if they were disordered over two neighboring positions. This resulted in the same stoichiometry of ADAD and DAAD, which enabled direct comparison of the energy of the studied systems. In addition, despite the symmetry found in both ADAD and DAAD, the structures were optimized without any constraints resulting from their corresponding crystal groups, with both of the structures being treated as P1 systems. This has been done purposely to enable the molecules to relax independently.

The results of the calculations (Table 2) show only slight changes in the unit cell dimensions resulting from the geometry optimization. Both the increase (i.e., 'a' for ADAD, 'c' for both of the structures) as well as the decrease (i.e., 'a' for DAAD and 'b' for both of the structures) of some unit cell lengths were observed. This is in agreement with the experimental SCXRD results, indicating that the structure was disordered and various guest orientations in the solid complexes exist. Despite similar unit cell dimensions, the optimized structures differed in their energies, indicating that DAAD is the more stable one by approximately 5 kcal/mol. Although no intermolecular interactions between guest molecules were observed in either of the structures, hydrogen bonds between the EST C3 hydroxyl group and primary hydroxyl groups of BCD were observed in ADAD (Figure 4). It should also be noted that after unit cell optimization, the symmetry in ADAD was no longer observed as the optimized 'a' and 'b' lengths differed significantly.

The next step in the DFT calculations was computation of the NMR chemical shielding constants for the optimized structures of ADAD and DAAD (which can be found in the Supplementary Material as ADADopt.cif and DAADopt.cif) using the GIPAW method. The isotropic chemical shielding values were then converted into chemical shifts to facilitate peak assignment of the NMR spectra and to compare the differences between the corresponding experimental and theoretical values obtained for the two optimized models.

2.3. ^{13}C CP MAS Solid State NMR Analysis

As described both in the introduction as well as in Section 4, in this study, the EST/ß-CD complexes were prepared using four different methods (LYS, STAND, MECH, STAND-SHORT). Detailed information on how exactly the samples were prepared can be found in Section 4.2.

To explore whether there are any structural differences between the complexes obtained in different ways, we have chosen the ^{13}C CP MAS solid state NMR analysis. The application of this method to the study of CD-based complexes has been recently reviewed

by us [16]. The spectra of the complexes (LYS, STAND, MECH, STANDSHORT) and reactants (EST, ß-CD) are presented in Figures 5–7. The spectra of the EST and ß-CD were recorded to facilitate the observation of changes in the chemical shifts and shapes of signals occurring upon complexation.

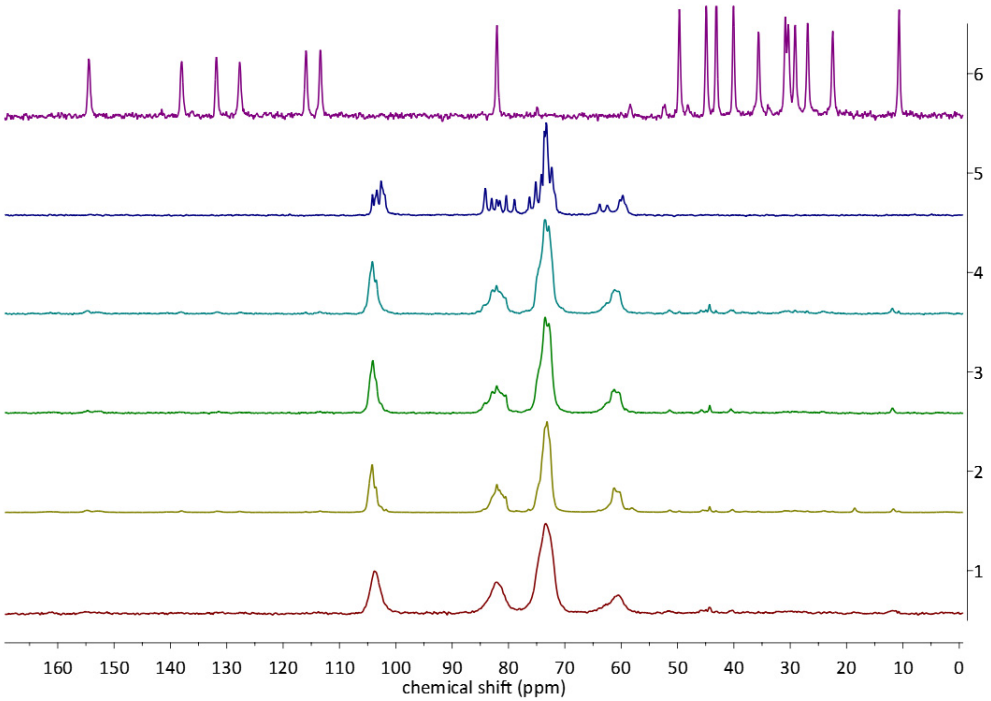


Figure 5. ^{13}C CP MAS NMR spectra of the EST (violet), ß-CD (dark blue), STAND (blue), STANDSHORT (green), MECH (olive green), and LYS (red).

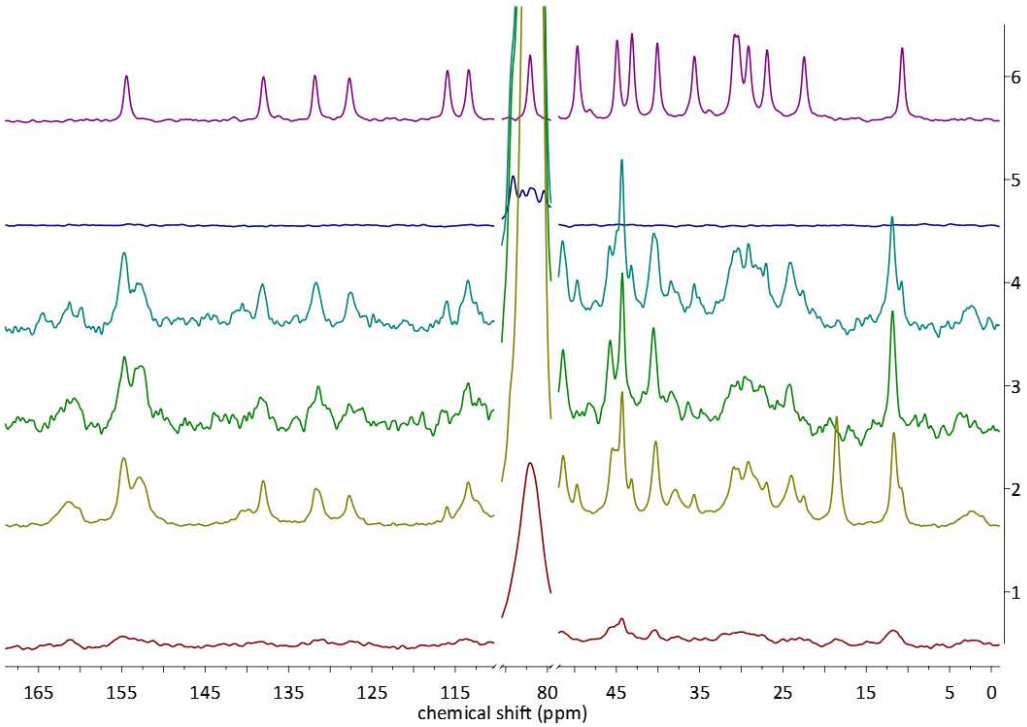


Figure 6. ^{13}C CP MAS NMR spectra of the EST (violet), ß-CD (dark blue), STAND (blue), STANDSHORT (green), MECH (olive green), and LYS (red). Only chosen regions of the spectra are presented for better visualization of the signals originating from EST.

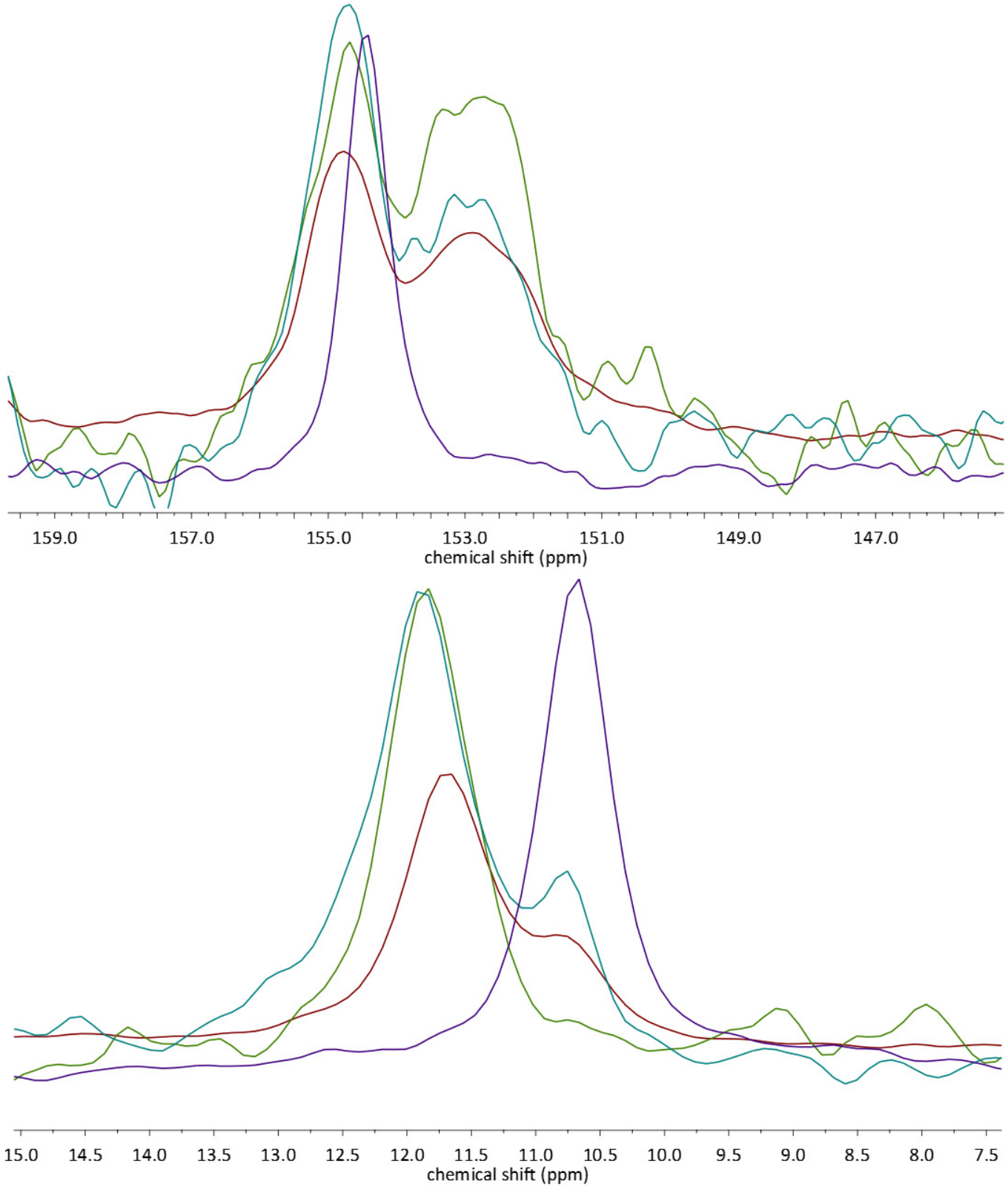


Figure 7. ^{13}C CP MAS NMR spectra of the EST (violet), STAND (blue), STANDSHORT (green), and MECH (red). Only chosen regions of the spectra are presented for better visualization of the signals originating from EST.

At first glance, in the stacked spectra, scaled in a way that the highest peaks have the same intensity (Figure 5), changes in the shape and number of signals originating from the

ß-CD are well visible. However, due to the significantly lower molecular mass of guest than host, the intensities of the signals of EST were very low. After increasing the intensity of the peaks (Figure 6), signals from EST have been revealed in all of the spectra of the complexes, with the exception of one obtained using the LYS method. In the LYS spectrum, signals originating from EST carbon atoms are either not visible or, in the best cases, very broad and low, i.e., in the 40–45 ppm region. This indicates that, as anticipated, the LYS method resulted in the amorphization of the sample, which was further confirmed by PXRD analysis (see Section 2.3).

In the discussion below, we will focus on the signals originating from the guest molecule, EST. This is justified for several reasons. First, the signals of EST, both in the non-complexed and in the EST/ß-CD forms, are usually well separated and sharp. Second, the chemical shifts of the EST carbon atoms occur in a wide range, 10–155 ppm, while all the signals from ß-CD can be found in a much wider range, 60–105 ppm. Moreover, the signals from the ß-CD are characterized by much larger FWHM and are highly overlapping; therefore, their analysis would not be possible without the ambiguous deconvolution.

We have started the NMR analysis from the 145–160 ppm range (Figure 7). In this region of the spectra, two overlapping peaks can be observed, with their maxima, respectively, at 152.9 and 154.8 ppm. Initially, after comparison with the spectrum of EST, in which a peak occurs at 152.8, we have assumed that these peaks originate from the complexed (154.42 ppm) and non-complexed (152.9 ppm) EST molecules. However, the intensities of the signals and areas under them were similar, which could mean that only around half of the EST was successfully complexed. Eventually, after careful analysis of the other regions of the spectra, 7.5–15 ppm (Figure 7), we have changed our initial assumptions. In this aliphatic region, the change in the chemical shift of the EST methyl group can be observed. The complexation resulted in the downfield shift of the single peak by c.a. 1 ppm, from 10.67 to 11.71 ppm. Still, even in some (MECH, STAND) of the spectra of complexes, the signal of the non-complexed EST methyl group could be observed. These observations allowed us to draw two conclusions. First, the two peaks in the spectra of complexes, located at 152.9 and 154.8 ppm, originate from the same carbon atom (C3) of the crystallographically nonequivalent EST molecules. This conclusion was also supported by the results of GIPAW calculations (Table 3), as the calculated chemical shifts for the C3 in ADAD and DAAD differ significantly. The other conclusion was that since a different amount of noncomplexed EST could be detected in the analyzed spectra, the yield of the complexation depends on the choice of the preparation method. As the ratio of the intensities of the peak at 10.67 to 11.71 was found to decrease in the order MECH → STAND → STANDSHORT, the yield was also decreasing in the same manner.

Upon complexation, the chemical shifts of some of the EST signals have only slightly changed (i.e., those from C1, C2, C4–C11). These carbon atoms form the A and B ring of EST and do not form any significant intermolecular interactions with other atoms; also, the conformation of these rings is highly rigid. The most apparent changes in the chemical shift values were observed for the signals occurring in the 21–53 ppm region (Figure 8). In the assignment of these signals, the results of GIPAW NMR calculations were found to be particularly useful. Additionally, the changes between the spectra of MECH, STAND, and STANDSHORT observed in the 10.67–11.71 ppm region (Figure 7) were found to be similar to those in the 21–53 ppm region. For example, the C14 signal has a 49.65 ppm shift in the spectrum of EST and 51.39 ppm in the spectrum of EST/ß-CD. In the spectra of MECH and STAND, the low-intensity signal from the noncomplexed EST can be observed, while in the spectrum of STANDSHORT, it is no longer visible. Similar observations were made for C15 signals, occurring at 22.45 and 23.98 ppm in the complexed and non-complexed forms, respectively (Figure 8).

Table 3. Experimental (exp.) and theoretically calculated (GIPAW) ^{13}C chemical shifts (δ) of EST and its complex with β-CD. Due to the presence of two EST molecules in the asymmetric unit of both ADAD and DAAD, two sets of values, (1) and (2), have been obtained for the first (1) and second (2) molecule present in the unit cell.

Atom Number	δ EST Exp.	δ EST GIPAW	δ EST Exp.—δ EST GIPAW	δ EST + ß-CD Exp.	δ ADAD GIPAW (1)	δ ADAD GIPAW (2)	(EST + ß-CD Exp.)—ADAD GIPAW (1)	(EST + βCD Exp.)—ADAD GIPAW (2)	δ DAAD GIPAW (1)	δ DAAD GIPAW (2)	(EST + ß-CD Exp.)—DAAD GIPAW (1)	(EST + ß-CD Exp.)—DAAD GIPAW (2)	δ EST Exp.—(EST/ß-CD Exp.)
1	127.65	129.59	−1.94	127.78	127.77	127.76	0.01	0.02	124.65	124.71	3.13	3.07	−0.13
2	113.33	111.18	2.15	113.43	109.86	109.82	3.57	3.61	109.61	109.65	3.82	3.78	−0.1
3	154.42	156.06	−1.64	154.8/152.9	157.21	157.23	−2.41	−2.43	155.72	155.71	−2.92	−2.91	1.62
4	115.89	118.35	−2.46	115.97	113.2	113.2	2.77	2.77	113.24	113.27	2.73	2.7	−0.08
5	137.94	139.1	−1.16	137.98	140.12	139.98	−2.14	−2.00	139.55	139.65	−1.57	−1.67	−0.04
6	30.83	30.75	0.08	30.92	28.79	28.81	2.13	2.11	28.58	28.56	2.34	2.36	−0.09
7	30.36	29.98	0.38	30.36	25.09	25.09	5.27	5.27	26.49	26.49	3.87	3.87	0.00
8	40.05	38.54	1.51	40.31	38	37.94	2.31	2.37	37.13	37.12	3.18	3.19	−0.26
9	44.89	43.52	1.37	45.5	45.24	45.2	0.26	0.30	43.73	43.74	1.77	1.76	−0.61
10	131.77	133.53	−1.76	131.84	132.39	132.31	−0.55	−0.47	130.46	130.52	1.38	1.32	−0.07
11	26.9	25.78	1.12	26.95	26.27	26.29	0.68	0.66	23.63	23.61	3.32	3.34	−0.05
12	35.59	36.07	−0.48	37.76	35.08	35.15	2.68	2.61	35.84	35.84	1.92	1.92	−2.17
13	43.09	42.62	0.47	44.3	43.5	43.38	0.8	0.92	42.08	42.07	2.22	2.23	−1.21
14	49.65	48.09	1.56	51.39	51.23	51.21	0.16	0.18	50.06	50.04	1.33	1.35	−1.74
15	22.45	21.06	1.39	23.98	20.61	20.61	3.37	3.37	21.45	21.43	2.53	2.55	−1.53
16	29.12	28.09	1.03	29.18	26.62	26.61	2.56	2.57	29.53	29.52	−0.35	−0.34	−0.06
17	82.04	82.63	−0.59	82.06	84.46	84.38	−2.4	−2.32	85.32	85.28	−3.26	−3.22	−0.02
18	10.67	7.54	3.13	11.71	8.46	8.5	3.25	3.21	8.54	8.52	3.17	3.19	−1.04

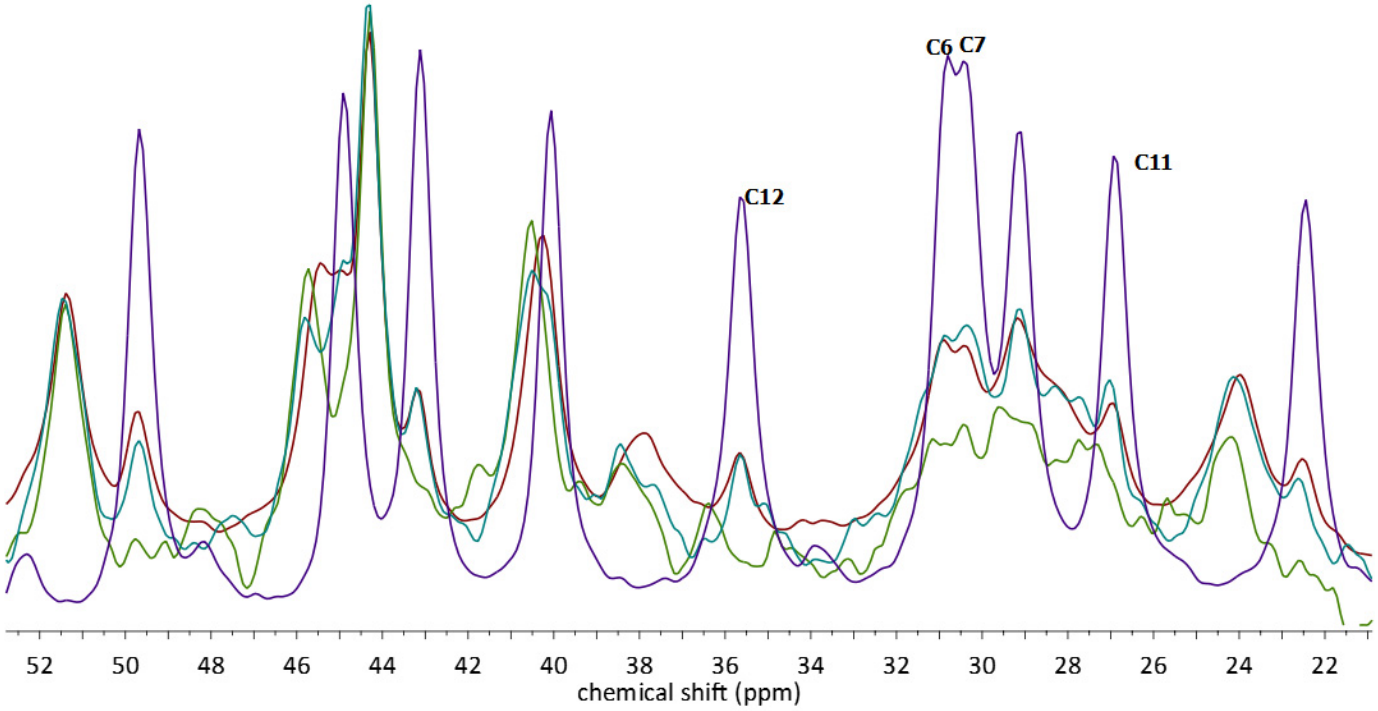


Figure 8. ^{13}C CP MAS NMR spectra of the EST (violet), STAND (blue), STANDSHORT (green), and MECH (red). Only chosen regions of the spectra are presented for better visualization of the signals originating from EST.

As mentioned above, the results of the GIPAW NMR calculations (Table 3) were found to be in very good agreement with corresponding experimental ones and facilitated proper signal assignment. The obtained differences between the experimental and theoretical values of δ for the complexes were found to be at a level similar to those of EST, not exceeding 4 ppm and, in most cases, below 3 ppm, with an exception for the C7 signal of ADAD. It should be noted, however, that in the spectra of the complexes, the peaks originating from C6, C7, C11, and C12 are overlapping and of a low intensity. This indicates the high level of dynamic disorder in this part of the guest molecule. All four atoms are chemically similar, as they are all secondary and form six-membered rings. Another explanation for this observation can be a dynamic of the C ring of EST. It has been reported previously that the C ring of EST can adopt either a chair or boat conformation, depending on its crystal form or, in the case of a solution, on the solvent [30]. It is therefore possible that the chair–boat conformational dynamics of the EST C ring can occur in the EST/ß-CD complex, which would explain the shape of the signals from C11 and C12.

The FT-IR spectra (Figure 9) of both ß-CD and EST have been found to be very similar to those reported previously [24]. The observed differences might have been caused by either the method of spectrum registration or the different degree of crystallinity. As in the case of the ^{13}C CP MAS NMR results, the FT-IR spectra of the complexes prepared by different methods have been found to be similar. However, there are also some noticeable differences among them. The signal at 3384 cm^{-1} is much narrower in LYS than in other cases. In addition, on the slope of the signal with a maximum at 2925 cm^{-1}, the small signals of EST, overlapped by the wide signal of ß-CD, are the most visible in the spectrum of STAND. Moreover, in the fingerprint area (500–850 cm^{-1}), the spectrum of LYS is flatter than the spectra of the complexes prepared by other methods. In all of the studied spectra of the complexes, broad signals in the range of 3000–3500 cm^{-1} can be found. These signals originate from multiple hydrogen bonds present in the studied system. These bonds differ in their length, energy, spectroscopic frequency, and intensity, as shown in previous works [31–34].

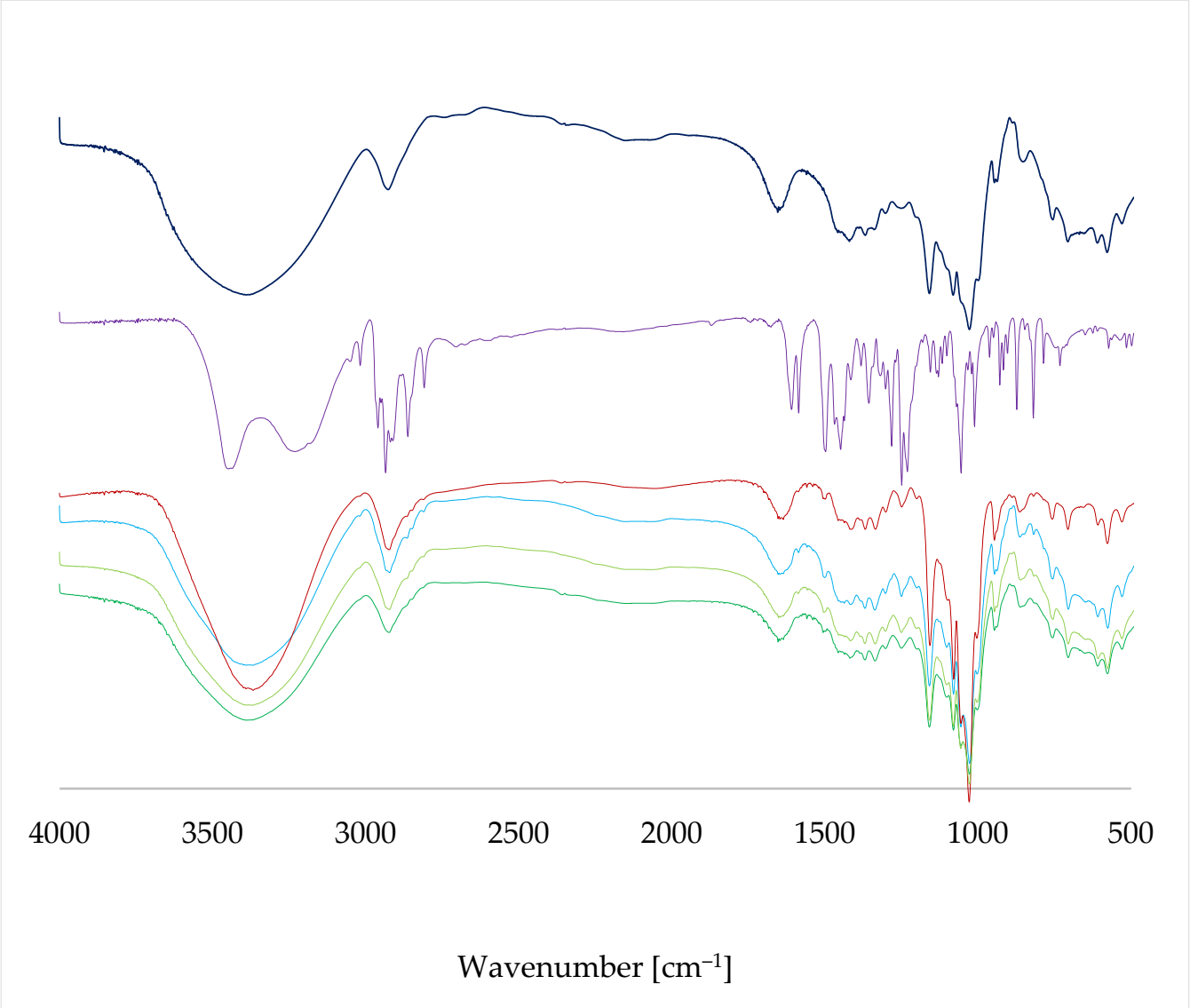


Figure 9. FT-IR spectra of ß-CD (dark blue), EST (violet), LYS (red), STAND (blue), MECH (olive green), and STANDSHORT (green).

Similar to the spectroscopic method results, the PXRD patterns (Figure 10) of the complexes are quite similar, with the exception of LYS. The lack of reflexes in the PXRD pattern of LYS and the characteristic halos indicate that the sample obtained by this method is amorphous. The PXRD patterns of STAND, MECH, and STANDSHORT are similar to the theoretical ones, simulated using the experimental crystal structure of EST/ß-CD. Characteristic reflexes can be found at the 2Θ values of 6.47, 7.25, 9.8, and 11.95 deg. In addition, two groups of signals in the ranges of 14.7–15.7 and 17.5–18.8 can be found both in the theoretical and experimental patterns. In the STAND, MECH, and STANDSHORT group, the pattern of STANDSHORT is slightly different than the other two. For example, in the PXRD pattern of STANDSHORT, there are no signals at 10.75 and 12.535, which are both present in the other two patterns. These reflexes are also present in the pattern of ß-CD, which indicates that in the samples of both STAND and MECH, some 'free' crystalline ß-CD can be found.

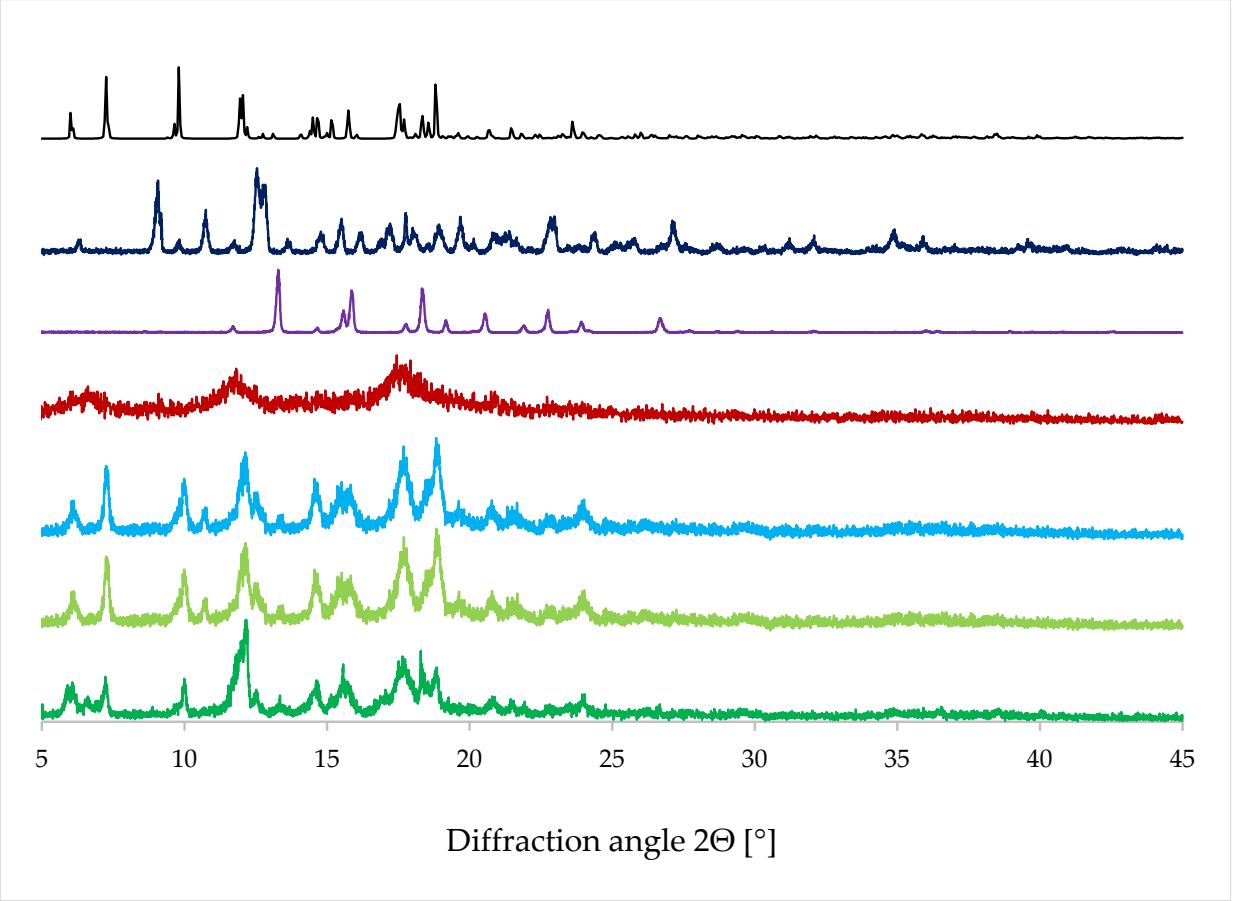


Figure 10. PXRD patterns: simulated for EST/ß-CD (black) and experimental PXRD prints for ß-CD (dark blue), EST (violet), LYS (red), STAND (blue), MECH (olive green), and STANDSHORT (green).

Comparison of the thermal analysis results (Table 4, Figures S2–S7) revealed additional differences between the complexes obtained using different preparation methods. The differences in mass loss during heating can be explained by the various ratios between the phases (EST:ß-CD:EST/ß-CD) in the analyzed samples. According to the SCXRD results, the total amount of water in EST and EST/ß-CD is 6.20% and 12.95%, respectively. The amount of water in BCD is variable [35], but usually within the 12.5–16% range. The TGA results for EST and ß-CD are in agreement with their corresponding crystal structures, indicating the total water loss in the analyzed temperature range. Lower than theoretically calculated values obtained for the EST/ß-CD complexes may indicate that some of the water molecules in the structure of EST/ß-CD are characterized by lower crystallographic occupancies and that the amount of crystal water in those complexes is variable. Moreover, the presence of non-complexed EST, with 6.20% water content, additionally decreases the anticipated values of water loss during the heating of the complexes. Significantly lower mass loss has been observed for the LYS sample, which can be explained by the final step of this method, lyophilization, which is used to decrease the amount of water in the sample. The DSC analysis of the complexes revealed two endothermic peaks in STANDSHORT and MECH, the first associated with the dehydration and the second with the decomposition of the sample. The higher enthalpy of dehydration was found in the sample with larger water loss, MECH. In the DSC thermograms of STAND and LYS, no clear endothermic peaks were observed.

Table 4. DSC-TGA analysis results.

STAND	DSC	(1) Onset temp. 29.26 °C Peak temp. 61.83 °C Enthalpy 83.82 J/g	(3) Onset temp. 152.0 °C Peak temp. 177.96 °C Enthalpy 10.74 J/g
		(2) Onset temp. 101.1 °C Peak temp. 111.27 °C Enthalpy 18.53 J/g	(4) Onset temp. 202.0 °C Peak temp. 212.69 °C Enthalpy 6.434 J/g
	TGA	Temp. range of dehydration: 27–200 °C Associated mass loss: 6.911%	
STANDSHORT	DSC	Onset temp. 26.38 °C Peak temp. 45.33 °C Enthalpy 185.1 J/g	
	TGA	Temp. range of dehydration: 25–200 °C Associated mass loss: 7.034%	
MECH	DSC	Onset temp. 30.55 °C Peak temp. 69.66 °C Enthalpy 209.8 J/g	
	TGA	Temp. range of dehydration: 20–200 °C Associated mass loss: 9.408%	
LYS	DSC	Onset temp. 41.0 °C Peak temp. 55.01 °C Enthalpy 47.53 J/g	
	TGA	Temp. range of dehydration: 25–200 °C Associated mass loss: 4.065%	
EST	DSC	(1) Onset temp. 81.88 °C Peak temp. 104.81 °C Enthalpy 19.70 J/g	
		(2) Onset temp. 175.91 °C Peak temp. 178.20 °C Enthalpy 91.06 J/g	
	TGA	Temp. range of dehydration: 20–200 °C Associated mass loss: 6.20%	
ß-CD	DSC	Onset temp. 56.31 °C Peak temp. 91.48 °C Enthalpy 380.8 J/g	
	TGA	Temp. range of dehydration: 20–100 °C Associated mass loss: 12.95%	

The surface morphology of EST/ß-CD complexes were assessed by SEM and the images are provided in Figure 11. As shown in Figure 11a, the SEM picture of STANDSHORT demonstrates a crystalline structure of this sample, dominated by cuboid-like crystals with average dimensions of about 100 µm. The SEM picture of MECH in Figure 11b presents an irregularly shaped crystalline structure. This sample is composed of crystals of different size, ranging from a few µm to a few hundred µm. Additionally, the crystals found in this sample are irregularly shaped. Meanwhile, the image of STAND (Figure 11c) shows a lot of large, prism-like crystals, larger even than those found in STANDSHORT, but with more irregular shapes. Finally, the LYS picture (Figure 11d) reveals the amorphous character of this sample, dominated by small particles of irregular shape, which is also in agreement with the PXRD results.

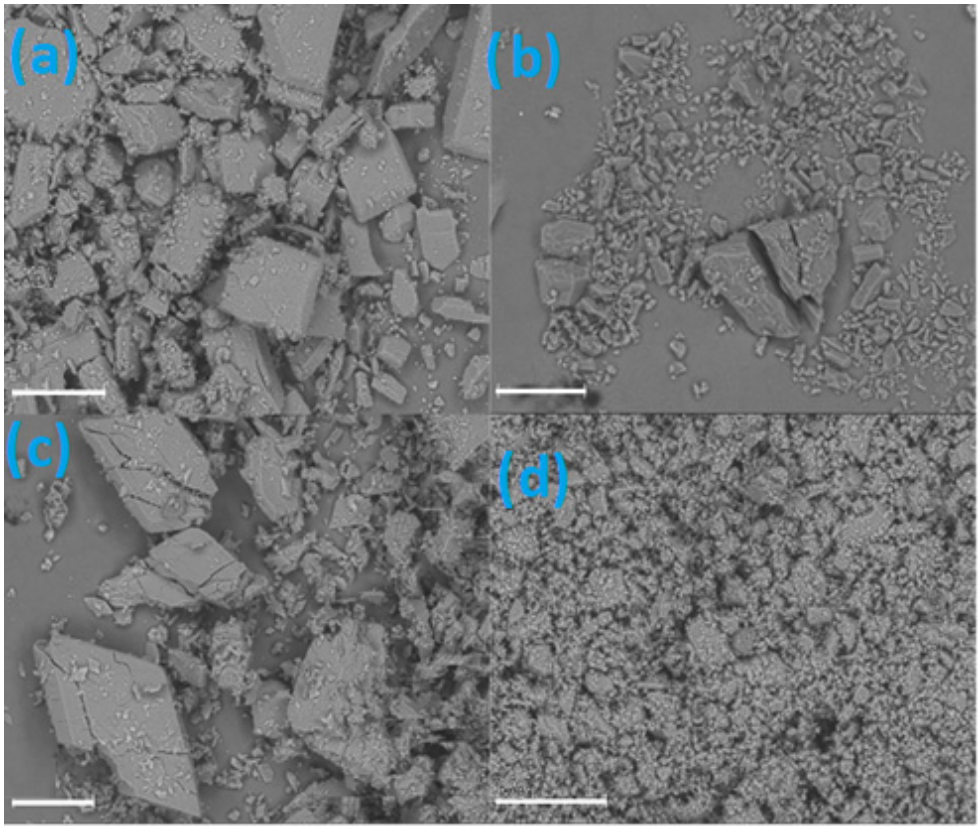


Figure 11. SEM images of STANDSHORT (**a**), MECH (**b**), STAND (**c**), and LYS (**d**). The length of the scale (white horizontal line) in each case is the same: 200 µm.

3. Conclusions

In this study, the inclusion complex of 17-β-estradiol and β-cyclodextrin has been prepared by four different methods. We have found that the method of complex preparation influences the final composition of the obtained sample as it affects the yield of complexation. However, regardless of the applied method, only one crystal form of the complex has been obtained, with an exception for a method involving lyophilization that resulted in the formation of an amorphous sample.

It should be noted that EST is an API with a long history of application in the treatment of various conditions, such as hormonal contraception, hormone replacement therapy (HRT), and treatment of menopausal and postmenopausal symptoms. ß-CD, on the other hand, is an excipient used in multiple original drugs currently on the worldwide market. It has also been shown that the complexation of EST with this particular cyclodextrin decreased toxicity of the studied hormone [23]. Moreover, the combination of EST and ß-CD was found to improve the bioavailability of the API by increasing its solubility [28].

The crystal structure of the 17-β-estradiol/β-cyclodextrin complex has been obtained for the first time by means of SCXRD. The 2:1 stoichiometry of host:guest has been determined. It has also been found that the highly disordered encapsulated 17-β-estradiol molecule has a unique orientation inside the dimeric host cavity. Using the solved crystal structure, periodic DFT calculations have been conducted to assess the energy differences between the two modeled structures. The GIPAW NMR calculations for the optimized structures facilitated peak assignment in the ^{13}C CP MAS NMR spectra. The ssNMR results confirmed that in the crystal structure of the studied complex, two orientations of 17-β-estradiol exist, as for the C3, two signals in the spectra were observed, indicating that only some of the hydroxyl groups of this carbon atom form hydrogen bonds.

SEM and thermal (TGA/DSC) analysis revealed noticeable differences between the complexes obtained using various methods. PXRD analysis confirmed the formation of the

complex in each case, with the exception of LYS, as this sample was proven to be amorphous. No major differences in the FT-IR spectra of the complexes obtained by different methods were observed.

The fact that, despite using numerous methods to obtain the studied complex, only one form has been received may indicate that only one stable crystal form is present in normal conditions. This, however, does not exclude the possibility of polymorphism at other temperature or pressure conditions, especially since in the studied structure significant disorder has been detected, which usually indicates the possibility of polymorphic phase transition.

This work highlights the importance of a combined ssNMR/SCXRD approach to studying the structure of the inclusion complexes formed by cyclodextrins, especially those characterized by a high level of structural disorder.

4. Materials and Methods

4.1. Materials

17-β-estradiol hemihydrate and β-cyclodextrin were purchased from BIOSYNTH; Biosynth AG, Staad, Switzerland) and used as received, without any further purification. For the mechanochemistry (MECH) method, freshly prepared, twice-distilled Milli-Q (Mq) water (Milli-Q water purification system, Millipore Corp., Waltham, MA, USA), with a conductivity of ~1 μS/cm, was used in the grinding process.

4.2. Methods of Complex Preparation

In this study, four different EST/ß-CD complex preparation methods were applied.

The first one was the **STANDARD [STAND]** method, commonly used in cases of CD inclusion complexes, when crystals of sufficient quality for SCXRD measurement are required. The STANDARD method is a slow-cooling crystallization technique: 60 mg of ß-CD was mixed in a flask with 1 mL distilled water and put into 70 °C water for 20 s to obtain a clear solution. Then the contents of the flask were poured into a beaker. In accordance with the molar mass of ß-CD and EST, the respective amount of EST was added to the beaker to maintain the 1:1 molar ratio. The beaker was put on a magnetic stirrer and left at room temperature for 15–20 min until a clear solution was obtained. Afterwards, the contents of the beaker were poured into a glass tube. The beaker was poured along with 0.5–1.0 mL water, which was also added to the glass tube. The glass tube was held in 70 °C water for 20 s to obtain a clear solution. Later, the tube was closed and put into a 70 °C water bath. A slow, gradual cooling process was performed over 10 days, reaching a temperature of 24 °C on the 10th day. At the end, a rotary evaporator was used.

The second one was the **STANDARD SHORT [STANDSHORT]** method: ß-CD and EST in a molar ratio of 1:1 were mixed with distilled water and put into a round bottom flask. The flask was closed and left on a magnetic stirrer for 24 h at a temperature of 44 °C. To obtain crystals, a rotary evaporator was used.

The third one was the **MECHANOCHEMICAL [MECH]** method: ß-CD and EST in a molar ratio of 1:1 were mixed in a mortar with 3–5 drops of Mq water and knitted for 5 min every day over 5 consecutive days.

The fourth one was the **LYOPHILIZATION [LYS]** method: the **STANDARD SHORT** crystallization method with the application of lyophilization instead of slow evaporation. A solution, which was obtained in accordance with the **STANDSHORT** method, was poured into smaller containers in which the lyophilization process took place. Firstly, the probes were frozen with the application of liquid nitrogen. Secondly, the probes were put into a lyophilizator for 48 h.

4.3. Powder X-ray Diffraction (PXRD)

For the PXRD measurements, a Panalytical Empyrean (Malvern, UK) diffractometer was used. The samples were analyzed in Bragg–Brentano reflection mode, using Cu-Kα radiation (λ = 1.54187 Å), a 2Θ range of 3–45°, and a 0.006565° step size. For the incident

beam a fixed divergence slit of 1/16°, an anti-scatter slit of 1/4°, and a fixed mask of 4 mm were used, and the diffracted beam path was equipped with a 7.5 mm anti-scatter slit.

4.4. Single-Crystal X-ray Diffraction (SCXRD)

A clear, light, colorless prism-like specimen, with dimensions of about 0.130 mm × 0.270 mm × 0.400 mm and coated with paraffin oil as cryo-protectant, was used for the X-ray crystallographic analysis. The X-ray intensity data were measured at 100 (2) K with a Bruker (Billerica, MA, USA) D8-VENTURE diffractometer using Cu *Kα* radiation (λ = 1.54178 Å). A low-temperature device (Oxford Cryosystems Ltd., Long Handorough, UK) provided a continuous stream of nitrogen vapor on the specimen during the data collection, while diffraction patterns were recorded using a CMOS-PHOTON III detector.

The total exposure time was 29.33 h. The frames were integrated with the Bruker SAINT software package [36] using a narrow-frame algorithm. The integration of the data using a monoclinic unit cell yielded a total of 133,843 reflections to a maximum θ angle of 76.27 (0.79 Å resolution). The final cell constants of a = 19.1245(15) Å, b = 24.4180(18) Å, c = 15.6004(11) Å, β = 109.500(5)°, and volume = 6867.2(8) Å^3 were based upon the refinement of the XYZ-centroids of 9073 reflections above 20 σ(I) with 6.094 < 2θ < 148.5°. Data were corrected for absorption effects using the Multi-Scan method (SADABS) [36].

The structure was solved by the intrinsic phasing method with SHELXT [37] and refined by full-matrix least squares against F^2 using SHELXL-2014/7 [38] through the SHELXLE GUI [39]. H-atoms were placed geometrically and refined in riding mode with isotropic displacement parameters fixed by SHELXL. Due to the structural complexity and disorder of the final model, soft restraints on bond lengths and angles, generated from the PRODRG2 webserver [40], were applied on the host and guest molecules of the inclusion complexes. Anisotropic displacement parameters were refined using soft restraints (SIMU) [41] implemented in the SHELXL program, where necessary. The final anisotropic full-matrix least-squares refinement on F^2 with 950 variables converged at R_1 = 9.5% for the observed data and wR_2 = 26.9% for all data. The goodness of fit (GoF) was 1.04. The crystallographic data along with the structural refinement details are summarized in Table 1. The data can be obtained from the Cambridge Crystallographic Data Centre under the reference number 2250781.

Geometric features of the crystal structure, e.g., interatomic distances, angles, dihedral angles, centroid coordinates, and mean plane equations through various groups of atoms, along with their e.s.d. estimations, were calculated via the full covariance matrix using the Olex2 program [42]. The final 3D model was drawn with Mercury 4.3.1 [43] and PyMoL [44].

4.5. Fourier-Transform Infrared Spectroscopy (FT-IR)

The studies were performed using a Perkin-Elmer Spectrum 1000 FT-IR spectrometer equipped with an MTEC 300 detector (MTEC, Ames, IA, USA). The samples were packed in ring cups with a diameter of 10 mm. The detector's chamber was purged with helium to reduce the effect of moisture evaporating from the samples during measurement. A spectrum obtained from the background sample was subtracted from the spectrum of each sample to eliminate the residual peaks of CO_2 and moisture. For each sample, 1024 scans were recorded and averaged in the infrared region between 4000 and 500 cm^{-1} at a resolution of 4 cm^{-1}. All spectral plots were prepared using the GRAMS/AI 8.0 Spectroscopy Software.

4.6. Cryo-Scanning Electron Microscopy (Cryo-SEM)

The analysis was performed using low-temperature scanning electron microscope ZEISS AURIGA (Warsaw, Poland) 60 coupled with a focused ion beam. Cryo-SEM allows sample observation without chemical fixing or drying. The procedure consists of sample freezing by immersion in liquid nitrogen, breaking, and etching.

4.7. Differential Scanning Calorimetry and Thermogravimetry Analysis (DSC-TGA)

Analysis was performed using apparat SDT Q600 (TA Instruments, New Castle, DE, USA) under nitrogen flow. The heating rate was equal to 10 °C/min and the sample mass was approximately 6–8 mg. Pierced aluminum sample pans were used in the analysis and the temperature range was set to 0–500 °C.

4.8. ^{13}C CP MAS NMR

Solid-state ^{13}C CP/MAS NMR spectra were recorded on a Bruker Avance III 600 spectrometer (Bruker BioSpin, Rheinstetten, Germany) operating at 600.15 MHz (^{1}H) and 150.91 MHz (^{13}C), and powder samples were spun at 12 kHz in a 4 mm ZrO_2 rotor using a double air-bearing probe head. Acquisition was performed with a standard CP pulse sequence with ramped CP scheme, 2 ms CP contact time, 4 s recycle delay, and a swept-frequency two-pulse phase modulation decoupling scheme, using a 3.2 µs proton 90° pulse. The decoupling field strength was set to 78 kHz. A total of 256 scans were acquired, 10.00 exponential apodization, a receiver gain equal to 2050, and 2048 acquired points. After zero filling and LP application, the spectrum size was 4096. ^{13}C chemical shifts were referenced to adamantane CH_2 at 38.48 ppm.

4.9. Periodic DFT Calculations

The density functional theory (DFT) calculations of geometry optimization and NMR parameters, under periodic boundary conditions, were carried out with the CASTEP program [45] implemented in the Materials Studio 2020 software [46] using the plane wave pseudopotential formalism. On-the-fly-generated ultrasoft pseudopotentials were generated using a Koelling–Harmon scalar relativistic approach [47]. The Perdew–Burke–Ernzerhof (PBE) [48] exchange-correlation functional, defined within the generalized gradient approximation, with Tkatchenko–Scheffler (TS) [49] dispersion correction, was used in the calculations.

4.9.1. Geometry Optimization

Geometry optimization was carried out using the limited memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS) [50] optimization scheme and smart method for finite basis set correction. The kinetic energy cutoff for the plane waves (E_{cut}) was set to 630.0 eV. The number of Monkhorst–Pack k-points during sampling for a primitive cell Brillouin zone integration [51] was set to 2 × 2 × 1 (for EST-ß-CD) and 1 × 1 × 2 (for 17-β-estradiol hemihydrate, refcode ESTDOL10), respectively.

The experimental X-ray structure of EST/ß-CD was used to create two initial periodic structures for calculations, containing two EST and four ß-CD molecules in the unit cell each. Details on the structural preparation can be found in Section 2.2.

During geometry optimization, all atom positions and cell parameters were optimized, with no constraints. The convergence criteria were set at 1×10^{-5} eV/atom for the energy, 3×10^{-2} eV/Å for the interatomic forces, 5×10^{-2} GPa for the stresses, and 1×10^{-3} Å for the maximum displacement. A fixed-basis set quality method for the cell optimization calculations and a 1×10^{-6} eV/atom tolerance for SCF were used.

4.9.2. NMR Parameter Calculations

The computation of shielding tensors was performed using the Gauge Including Projector Augmented Wave Density Functional Theory (GIPAW) method of Pickard et al. [52]. To compare the theoretical and experimental data, the calculated chemical shielding constants (σiso) were converted to chemical shifts (δiso) using the following equation: δiso = (σGly + δGly)—σiso, where σGly and δGly stand for the shielding constant and the experimental chemical shift, respectively, of the glycine carbonyl carbon atom (176.50 ppm).

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/molecules28093747/s1: SI.pdf file containing Table S1: Geometric

parameters of the ß-CD host molecule in the EST/ß-CD crystal structure; Table S2: Main H-bonds in the crystal structure of EST/ß-CD; Figure S1: Radar plots of some geometrical parameters of Table S1; Figure S2: DSC/TGA curve of LYS; Figure S3: DSC/TGA curve of EST; Figure S4: DSC/TGA curve of β-CD; Figure S5: DSC/TGA curve of STAND; Figure S6: DSC/TGA curve of MECH; Figure S7: DSC/TGA curve of STANDSHORT; DAAD.cif; ADAD.cif; DAADopt.cif; ADADopt.cif; estradiol_beta-CD.cif; and checkcif-1.pdf.

Author Contributions: Conceptualization, A.H.M., Ł.S., K.B. and D.M.P.; methodology, A.H.M., Ł.S., K.B., E.C., M.K.D., M.Z.-P. and D.M.P.; software, A.H.M., Ł.S., K.B., E.C. and D.M.P.; validation, A.H.M., Ł.S., K.B., M.K.D., M.Z.-P. and D.M.P.; formal analysis, A.H.M., Ł.S., K.B., E.C., M.K.D., M.Z.-P. and D.M.P.; investigation, A.H.M., Ł.S., K.B., E.C. and D.M.P.; resources, A.H.M., Ł.S. and K.B.; data curation, A.H.M., Ł.S. and K.B.; writing—original draft preparation, A.H.M., Ł.S. and K.B.; writing—review and editing, A.H.M., Ł.S., K.B., M.K.D. and D.M.P.; visualization, A.H.M., Ł.S., K.B. and E.C.; supervision, A.H.M. and Ł.S.; project administration, A.H.M. and Ł.S.; funding acquisition, A.H.M. and Ł.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Medical University of Warsaw, Poland, grant number F/MB/02/22.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data can be obtained from the corresponding author (Ł.S.) by email. Crystallographic data have been deposited with the Cambridge Structural Database (CSD) under deposition number CCDC: 2250781.

Acknowledgments: The authors acknowledge the support of the National and Kapodistrian University of Athens (NKUA) Core Facilities and Nikolaos Tsoureas (Department of Chemistry, NKUA) for the SCXRD data collection.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Sample Availability: Samples of the compounds are available from the corresponding author (Ł.S.) by email.

References

1. Thomas, M.P.; Potter, B.V. The structural biology of oestrogen metabolism. *J. Steroid Biochem. Mol. Biol.* **2013**, *137*, 27–49. [CrossRef] [PubMed]
2. Grandi, G.; Napolitano, A.; Cagnacci, A. Metabolic impact of combined hormonal contraceptives containing estradiol. *Expert Opin. Drug Metab. Toxicol.* **2016**, *12*, 779–787. [CrossRef]
3. Santoro, N.; Epperson, C.N.; Mathews, S.B. Menopausal Symptoms and Their Management. *Endocrinol. Metab. Clin. N. Am.* **2015**, *44*, 497–515. [CrossRef] [PubMed]
4. Daulbayev, C.; Kaidar, B.; Sultanov, F.; Bakbolat, B.; Smagulova, G.; Mansurov, Z. The recent progress in pitch derived carbon fibers applications. A Review. *S. Afr. J. Chem. Eng.* **2021**, *38*, 9–20. [CrossRef]
5. Guo, A.; Gong, X.; He, J.; Guo, Y.; Ning, L.; Chen, X.; Xu, J.; Guo, Y.; Wang, H. Study on Co-crystals of Estradiol. *Her. Med.* **2021**, *40*, 1716–1723.
6. Wang, J.R.; Wang, X.; Yang, Y.; Chen, X.; Mei, X. Solid-state characterization of 17β-estradiol co-crystals presenting improved dissolution and bioavailability. *CrystEngComm* **2016**, *18*, 3498–3505. [CrossRef]
7. Ning, L.; Gong, X.; Li, P.; Chen, X.; Wang, H.; Xu, J. Measurement and correlation of the solubility of estradiol and estradiol-urea co-crystal in fourteen pure solvents at temperatures from 273.15 K to 318.15 K. *J. Mol. Liq.* **2020**, *304*, 112599. [CrossRef]
8. Kovacs, T.; Nagy, P.; Panyi, G.; Szente, L.; Varga, Z.; Zakany, F. Cyclodextrins: Only Pharmaceutical Excipients or Full-Fledged Drug Candidates? *Pharmaceutics* **2022**, *14*, 2559. [CrossRef] [PubMed]
9. European Medicines Agency. Available online: https://www.ema.europa.eu/en (accessed on 21 February 2023).
10. U.S. Food & Drug Administration (FDA). Available online: https://www.fda.gov (accessed on 21 February 2023).
11. Pharmaceutical and Medical Devices Agencty. Available online: https://www.pmda.go.jp/english/index.html (accessed on 21 February 2023).
12. Sun, J.; Hong, H.; Zhu, N.; Han, L.; Suo, Q. Effect of preparation methods on tosufloxacin tosylate/hydroxypropyl-β-cyclodextrin inclusion complex. *Braz. J. Pharm. Sci.* **2022**, *58*, e18650. [CrossRef]

13. Cid-Samamed, A.; Rakmai, J.; Mejuto, J.C.; Simal-Gandara, J.; Astray, G. Cyclodextrins inclusion complex: Preparation methods, analytical techniques and food industry applications. *Food Chem.* **2022**, *384*, 132467. [CrossRef]
14. Deckmann Nicoletti, C.; de Sá Haddad Queiroz, M.; de Souza Lima, C.G.; de Carvalho da Silva, F.; Futuro, D.O.; Ferreira, V.F. An improved method for the preparation of β-lapachone:2-hydroxypropyl-β-cyclodextrin inclusion complexes. *J. Drug Deliv. Sci. Technol.* **2020**, *58*, 101777. [CrossRef]
15. Wdowiak, K.; Rosiak, N.; Tykarska, E.; Żarowski, M.; Płazińska, A.; Płaziński, W.; Cielecka-Piontek, J. Amorphous Inclusion Complexes: Molecular Interactions of Hesperidin and Hesperetin with HP-B-CD and Their Biological Effects. *Int. J. Mol. Sci.* **2022**, *23*, 4000. [CrossRef] [PubMed]
16. Mazurek, A.H.; Szeleszczuk, Ł. A Review of Applications of Solid-State Nuclear Magnetic Resonance (ssNMR) for the Analysis of Cyclodextrin-Including Systems. *Int. J. Mol. Sci.* **2023**, *24*, 3648. [CrossRef] [PubMed]
17. Mazurek, A.H.; Szeleszczuk, Ł. Current Status of Quantum Chemical Studies of Cyclodextrin Host–Guest Complexes. *Molecules* **2022**, *27*, 3874. [CrossRef] [PubMed]
18. Mazurek, A.H.; Szeleszczuk, Ł.; Gubica, T. Application of Molecular Dynamics Simulations in the Analysis of Cyclodextrin Complexes. *Int. J. Mol. Sci.* **2021**, *22*, 9422. [CrossRef]
19. Mazurek, A.H.; Szeleszczuk, Ł.; Pisklak, D.M. Periodic DFT Calculations—Review of Applications in the Pharmaceutical Sciences. *Pharmaceutics* **2020**, *12*, 415. [CrossRef]
20. Gallez, A.; Palazzo, C.; Blacher, S.; Tskitishvili, E.; Noël, A.; Foidart, J.M.; Evrard, B.; Pequeux, C.; Piel, G. Liposomes and drug-in-cyclodextrin-in-liposomes formulations encapsulating 17β-estradiol: An innovative drug delivery system that prevents the activation of the membrane-initiated steroid signaling (MISS) of estrogen receptor α. *Int. J. Pharm.* **2020**, *573*, 118861. [CrossRef]
21. Schwarz, D.H.; Engelke, A.; Wenz, G. Solubilizing steroidal drugs by β-cyclodextrin derivatives. *Int. J. Pharm.* **2017**, *531*, 559–567. [CrossRef]
22. Cai, W.; Yao, X.; Shao, X.; Pan, Z. Bimodal Complexations of Steroids with Cyclodextrins by a Flexible Docking Algorithm. *J. Incl. Phenom. Macrocycl. Chem.* **2005**, *51*, 41–51. [CrossRef]
23. Silva, M.C.G.D.; Silva, J.F.D.; Santos, T.P.; Silva, N.P.C.D.; Santos, A.R.D.; Andrade, A.L.C.; Souza, E.H.L.D.S.; Sales Cadena, M.R.; Sá, F.B.; Silva Junior, V.A.D.; et al. The complexation of steroid hormones into cyclodextrin alters the toxic effects on the biological parameters of zebrafish (Danio rerio). *Chemosphere* **2019**, *214*, 330–340. [CrossRef]
24. Haimhoffer, Á.; Vas, A.; Árvai, G.; Fenyvesi, É.; Jicsinszky, L.; Budai, I.; Bényei, A.; Regdon, G., Jr.; Rusznyák, Á.; Vasvári, G.; et al. Investigation of the Drug Carrier Properties of Insoluble Cyclodextrin Polymer Microspheres. *Biomolecules* **2022**, *12*, 931. [CrossRef]
25. Lin, Z.H.; Wang, X.X.; Kou, S.B.; Shi, J.H. Exploring the inclusion interaction of estradiol with β-CD and HP-β-CD with the help of molecular dynamics simulation as well as multi-spectroscopic approaches. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2022**, *269*, 120764. [CrossRef] [PubMed]
26. Sadlej-Sosnowska, N. Fluorometric determination of association constants of three estrogens with cyclodextrins. *J. Fluoresc.* **1997**, *7*, 195–200. [CrossRef]
27. van Uden, W.; Woerdenbag, H.J.; Pras, N. Cyclodextrins as a useful tool for bioconversions in plant cell biotechnology. *Plant Cell Tiss Organ Cult.* **1994**, *38*, 103–113. [CrossRef]
28. Vicatos, A.I.; Hoossen, Z.; Caira, M.R. Inclusion complexes of the steroid hormones 17β-estradiol and progesterone with β- and γ-cyclodextrin hosts: Syntheses, X-ray structures, thermal analyses and API solubility enhancements. *Beilstein J. Org. Chem.* **2022**, *18*, 1749–1762. [CrossRef]
29. Mentzafos, D.; Mavridis, I.M.; Le Bas, G.; Tsoucaris, G. Structure of the 4-It Tert-Butylbenzyl Alcohol–β-Cyclodextrin Complex. Common Features in the Geometry of β-Cyclodextrin Dimeric Complexes. *Acta Crystallogr. Sect. B* **1991**, *47*, 746–757. [CrossRef]
30. Commodari, F.; Sclavos, G.; Ibrahimi, S.; Khiat, A.; Boulanger, Y. Comparison of 17beta-estradiol structures from X-ray diffraction and solution NMR. *Magn. Reson. Chem.* **2005**, *43*, 444–450. [CrossRef]
31. Rekik, N.; Issaoui, N.; Ghalla, H.; Oujia, B.; Wójcik, M.J. Infrared spectral density of H-bonds within the strong anharmonic coupling theory: Indirect relaxation effect. *J. Mol. Struct.* **2007**, *844–845*, 21–31. [CrossRef]
32. Brela, M.Z.; Klimas, O.; Surmiak, E.; Boczar, M.; Nakajima, T.; Wójcik, M.J. Comparison of the Hydrogen Bond Interaction Dynamics in the Guanine and Cytosine Crystals: Ab Initio Molecular Dynamics and Spectroscopic Study. *J. Phys. Chem. A* **2019**, *123*, 10757–10763. [CrossRef]
33. Rekik, N.; Issaoui, N.; Ghalla, H.; Oujia, B.; Wójcik, M.J. IR spectral density of H-bonds. Both intrinsic anharmonicity of the fast mode and the H-bond bridge. Part I: Anharmonic coupling parameter and temperature effects. *J. Mol. Struct. Theochem.* **2007**, *821*, 9–21. [CrossRef]
34. Wójcik, J.M. Theoretical Modeling of Vibrational Spectra and Proton Tunneling in Hydrogen-Bonded Systems. *Adv. Chem. Phys.* **2016**, *160*. [CrossRef]
35. Pereva, S.; Nikolova, V.; Angelova, S.; Spassov, T.; Dudev, T. Water inside β-cyclodextrin cavity: Amount, stability and mechanism of binding. *Beilstein J. Org. Chem.* **2019**, *15*, 1592–1600. [CrossRef]
36. Bruker. *APEX 3, SAINT, SADABS*; Bruker AXS Inc.: Madison, WI, USA, 2012.
37. Sheldrick, G.M. It SHELXT—Integrated Space-Group and Crystal-Structure Determination. *Acta Crystallogr. Sect. A* **2015**, *71*, 3–8. [CrossRef] [PubMed]

38. Sheldrick, G.M. Crystal Structure Refinement with It SHELXL. *Acta Crystallogr. Sect. C* **2015**, *71*, 3–8. [CrossRef] [PubMed]
39. Hübschle, C.B.; Sheldrick, G.M.; Dittrich, B. It ShelXle: A Qt Graphical User Interface for It SHELXL. *J. Appl. Crystallogr.* **2011**, *44*, 1281–1284. [CrossRef]
40. Schüttelkopf, A.W.; van Aalten, D.M.F. PRODRG: A Tool for High-Throughput Crystallography of Protein–Ligand Complexes. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 1355–1363. [CrossRef]
41. Thorn, A.; Dittrich, B.; Sheldrick, G.M. Enhanced Rigid-Bond Restraints. *Acta Crystallogr. Sect. A Found. Crystallogr.* **2012**, *68*, 448–451. [CrossRef]
42. Dolomanov, O.V.; Bourhis, L.J.; Gildea, R.J.; Howard, J.A.K.; Puschmann, H. It OLEX2: A Complete Structure Solution, Refinement and Analysis Program. *J. Appl. Crystallogr.* **2009**, *42*, 339–341. [CrossRef]
43. Macrae, C.F.; Bruno, I.J.; Chisholm, J.A.; Edgington, P.R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P.A. It Mercury CSD 2.0—New Features for the Visualization and Investigation of Crystal Structures. *J. Appl. Crystallogr.* **2008**, *41*, 466–470. [CrossRef]
44. *The Pymol Molecular Graphics System*, Version 1.8; Schrödinger, Inc.: New York, NY, USA, 2015.
45. Clark, S.J.; Segall, M.D.; Pickard, C.J.; Hasnip, P.J.; Probert, M.J.; Refson, K.; Payne, M.C. First principles methods using CASTEP. *Z. Krist.-Cryst. Mater.* **2005**, *220*, 567–570. [CrossRef]
46. BIOVIA. Materials Studio. Available online: http://accelrys.com/products/collaborative-science/biovia-materials-studio (accessed on 23 February 2023).
47. Koelling, D.D.; Harmon, B.N. Technique for relativistic spin-polarized calculations. *J. Phys. C Solid State Phys.* **1977**, *10*, 3107–3114. [CrossRef]
48. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868. [CrossRef] [PubMed]
49. Tkatchenko, A.; Scheffler, M. Accurate Molecular van der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.* **2009**, *102*, 073005. [CrossRef]
50. Packwood, D.; Kermode, J.; Mones, L.; Bernstein, N.; Woolley, J.; Gould, N.; Ortner, C.; Csányi, G. A universal preconditioner for simulating condensed phase materials. *J. Chem. Phys.* **2016**, *144*, 164109. [CrossRef] [PubMed]
51. Monkhorst, H.J.; Pack, J.D. Special points for Brillouin-zone integrations—A reply. *Phys. Rev. B* **1977**, *16*, 1748–1749.
52. Pickard, C.J.; Mauri, F. All-electron magnetic response with pseudopotentials: NMR chemical shifts. *Phys. Rev. B* **2001**, *63*, 63–77. [CrossRef]

International Journal of
Molecular Sciences

MDPI

Review

Application of Various Molecular Modelling Methods in the Study of Estrogens and Xenoestrogens

Anna Helena Mazurek [1], Łukasz Szeleszczuk [1,*], Thomas Simonson [2] and Dariusz Maciej Pisklak [1]

[1] Chair and Department of Physical Pharmacy and Bioanalysis, Department of Physical Chemistry, Medical Faculty of Pharmacy, University of Warsaw, Banacha 1 str., 02-093 Warsaw Poland; annamazurek21@gmail.com (A.H.M.); dpisklak@wum.edu.pl (D.M.P.)
[2] Laboratoire de Biochimie (CNRS UMR7654), Ecole Polytechnique, 91-120 Palaiseau, France; thomas.simonson@polytechnique.edu
* Correspondence: lszeleszczuk@wum.edu.pl; Tel.: +48-501-255-121

Received: 21 July 2020; Accepted: 1 September 2020; Published: 3 September 2020

Abstract: In this review, applications of various molecular modelling methods in the study of estrogens and xenoestrogens are summarized. Selected biomolecules that are the most commonly chosen as molecular modelling objects in this field are presented. In most of the reviewed works, ligand docking using solely force field methods was performed, employing various molecular targets involved in metabolism and action of estrogens. Other molecular modelling methods such as molecular dynamics and combined quantum mechanics with molecular mechanics have also been successfully used to predict the properties of estrogens and xenoestrogens. Among published works, a great number also focused on the application of different types of quantitative structure–activity relationship (QSAR) analyses to examine estrogen's structures and activities. Although the interactions between estrogens and xenoestrogens with various proteins are the most commonly studied, other aspects such as penetration of estrogens through lipid bilayers or their ability to adsorb on different materials are also explored using theoretical calculations. Apart from molecular mechanics and statistical methods, quantum mechanics calculations are also employed in the studies of estrogens and xenoestrogens. Their applications include computation of spectroscopic properties, both vibrational and Nuclear Magnetic Resonance (NMR), and also in quantum molecular dynamics simulations and crystal structure prediction. The main aim of this review is to present the great potential and versatility of various molecular modelling methods in the studies on estrogens and xenoestrogens.

Keywords: molecular modelling; estrogens; xenoestrogens; estradiol; docking; Density Functional Theory (DFT)

1. Introduction

Successful applications of molecular modelling methods can be found in almost every branch of modern physics, chemistry, and biology. This versatility and popularity results from the constantly increasing computing power of both personal computers and specialized servers as well as the availability of molecular modelling software. The number of properties that can be accurately predicted and phenomena that can be explained as well as problems that can be solved using such calculations are enormous. Therefore, in this review, recent advances in molecular modelling applications to study the chemistry and biochemistry of estrogens [1] and xenoestrogens [2] will be presented. The aim of this article is not only to present the large volume of information concerning title compounds that have been obtained in recent years using in silico methods but also to present to

researchers who are not specialized in molecular modelling methods the possible applications of these compounds and, in that way, encourage them to use such calculations in their own studies.

This review is organized as follows: first the title compounds, estrogens, and xenoestrogens are briefly summarized, with particular focus on those that have already been objects of computational studies. Then, the most important biomolecules that are involved in the metabolism and action of estrogens and xenoestrogens are described. In the main part of this review, the computational methods that have been employed in the studies on estrogens and xenoestrogens are presented. Each of the methods is briefly described, without too many details, as there are many very good reviews referenced in this article focusing on the basics of particular methods. Wherever possible, the computational results were compared to the corresponding experimental ones; however, in many cases, such comparison was impossible either due to the lack of experimental results or the purely theoretical character of the published work. In the next section, a critical analysis of the reviewed methods is presented, supported by the presentation of some technical aspects of the reviewed studies. This was done to facilitate the choice of a certain method or its properties. From the authors' perspective, the number of studies on estrogens and xenoestrogens is constantly increasing; however, only in some of them are the experiments supported by suitable theoretical studies. Therefore, our aim was to present to researchers working with estrogens and xenoestrogens selected computational tools that can be used to facilitate and improve their studies.

1.1. Estrogens and Xenoestrogens: Types, Main Representatives, and Their Toxicity

Estrogens are a group of natural steroid sex hormones. There are four of them: estrone (E1), estradiol (E2), estriol (E3), and estretrol (E4) [1] (see the Supplementary Materials: Figure S1). The last one is produced only during pregnancy by the fetus liver [3]. Among these four compounds, E2 plays the most important role in the human organism and, therefore, is of high importance in breast or ovarian cancer progression. With regard to their relative binding affinity (RBA) to Estrogen Receptor (ER), with the exclusion of estretrol, estrogens are ranked in the following order: estradiol > estrone > estriol. In comparison to estradiol, the activity and potency of estrone and estriol are, respectively, 10 and 100 times lower than that of E2 [4]. E1 mainly functions as estradiol's metabolite and, at the same time, serves as its precursor (the estrone-to-estradiol transformation is reversible) [5]. On the other hand, in non-pregnant women, estriol levels in the blood are hardly detectable, whereas during pregnancy its amount distinctly grows because it is produced by the placenta as well [6]. All estrogens are used as medication in menopausal hormone therapy, although estradiol is the most applied [7,8]. In such external applications, they should be considered as xenoestrogens.

Apart from estrogens, other non-endogenous substances, called endocrine-disrupting chemicals (EDCs), can bind to the ER. According to the European Commission Regulation from 2018 [9], EDCs are substances that have adverse effects in non-target organisms, have an endocrine mode of action, and exhibit these adverse effects as a consequence of an endocrine mode of action. In the Guidance on the identification and studies regarding ECDs, published by the European Chemical Agency (ECA) and the European Food Safety Authority (EFSA) [10], the importance of in silico studies in the process of EDC research is clearly pointed out.

More detailed information on in silico examination regarding EDCs is provided in OECD (Organization for Economic Co-operation and Development) Guidance (Update v3, 2017) [11], where Quantitative Structure–Activity Relationship (QSAR) methodologies, prediction of the metabolic transformation (ADME), and CYP450 metabolism investigations are listed as Level 1 methods in the identification and study process. It is emphasized that applications of such methods can be used to identify the groups of chemicals and structural characteristics that are responsible for the observed in vivo effects and can serve as a tool to explain possible differences between in vitro and in vivo results, clarifying EDCs' mechanism of action. In the Conceptual Framework for Testing and Assessment of EDCs [12], at investigation Level 1, not only QSAR and ADME but also 'other in silico model predictions' have been listed. Additively, special QSAR Guidance on the topic has been published.

A large group of EDCs that are well-known and have been investigated for decades are xenoestrogens, which are xenobiotics that cause either an estrogenic or an antiestrogenic effect [13]. Nowadays, they are widespread and originate from different sources (see the Supplementary Materials: Figure S2). In the in silico studies, the main emphasis has been put, so far, on the investigation of pharmaceuticals, phytoestrogens, bisphenol A, and phthalates.

1.2. Estrogen-Related Biomolecules as the Molecular Modelling Study Objects

Numerous studies involving various in silico methods, but mainly ligand docking, have been performed to look for new inhibitors of the enzymes metabolizing estradiol. These proteins are a possible target for new drugs and are depicted in Table 1. Most of the research deals with aromatase [14–22], as its inhibitors, like letrozole, are already in medicinal use [23]. However, those inhibitors exhibit some detrimental side effects, which is the reason why research continues on this topic. 17β-Hydroxysteroid dehydrogenase (17β-HSD) [24,25], sulfatase (STS) [26], as well as sulfotransferase (SULT) [27] have also been taken under close consideration using molecular modelling in order to contribute to anti-cancer drug development. The first of the two enzymes transforms estrone into estradiol. The second inactivates estradiol by transforming it into a sulfated form.

Table 1. Reagents of the main metabolic processes regarding estradiol that have been investigated with in silico methods.

Substrate	Enzyme	Product	Ref.
Testosterone	Aromatase (CYP219A1)	Estradiol	[16]
Estradiol	17β-OH-dehydrogenase (17β-HSD)	Estrone	[24,25]
Estradiol	CYP1B1	4-OH-hydroxylated estradiol	[30]
Estradiol	CYP1A1, CYP1A2	2-OH-hydroxylated estradiol	[28,29]
Estradiol	Sulfotransferase (SULT)	Inactivated estradiol (sulfated)	[27]
Inactivated (sulfated) estradiol	Sulfatase (STS)	Activated estradiol	[26]

Even if conceptually estradiol is central to the above-mentioned studies, none of them used it as a target ligand. Nevertheless, in order to model estradiol activity metabolism in molecular studies, the most common objects are estradiol itself, its natural receptor—the estrogen receptor (ER)—and its main hydroxylating enzymes: CYP1A1 [28], CYP1A2 [29], CYP1B1 [30], and SHBG (sex hormone-binding globulin). The last one is a protein that can bond estradiol (E2); thus, it has direct impact on the amount of free E2 in plasma and, consequently, on the hormone's bioavailability [31]. Taking into account the interest of in silico research for the enzymes listed in this paragraph, only topics associated with these systems are covered in the main part of this review.

Two subtypes of the estrogen receptor are known, ERα and ERβ, with tissue-specific expression [32]. ERα is found mainly in the mammary gland, uterus, ovary (thecal cells), male reproductive organs, prostate, liver, and adipose tissue. ERβ is present in the prostate, bladder, ovary (granulosa cells), colon, adipose tissue, and immune system [33]. Despite being encoded by different genes, both estrogen receptors show high homology, and in both of them the E domain contains the ligand-binding domain (LBD) and C domain, which is the DNA binding domain (DBD) [34]. The homology between ERα and ERβ in DBD is more than 95% [35]. In LBD the homology is about 55% [35]. In ER structures, two transactivation functions are present, called AF-1 (located in N-terminal domain) and AF-2 (located in LBD) [36]. They contain the nuclear location signals and, after proper exposure of their surface, are

responsible for incorporation of the co-activators, which is a necessary step to induce activation of the intercellular signaling pathways.

Binding of 17β-estradiol and any agonist to the ER requires creating a hydrogen bond with His524 (in ERα) [37] or His475 (in ERβ) [38]. This leads to a unique agonist-bound conformation of the receptor's LBD, characterized by a specific repositioning of the H12 helix, which is the most C-terminal helix of the LBD (molecular switch) [39]. On the contrary, selective ER modulators (SERMs) such as raloxifene or tamoxifen induce relocation of H12 into the co-activator binding cleft, which blocks AF-2 activity. Finally, pure antagonists completely destabilize H12 [40].

All the above actions concerning estrogens' binding to ER are genomic actions. This means that they require translocation of the estrogen–ER complex to the nucleus and interaction with chromatin at specific sequences, known as estrogen response elements [41]. There are other estrogen signaling paths that are non-genomic and involve indirect regulation of gene expression [42]. They include activation of various protein-kinase cascades after binding of an estrogen molecule to a membrane receptor, usually GPER1 (G protein-coupled ER1) [43]. It has been proven that estrogen binding to GPER1 shows similarity to estrogen–ER binding; however, estrogen's affinity for GPER1 is significantly lower than for ER [43]. Nevertheless, all other steroid hormones are characterized by even lower affinities towards GPER1.

Experimental analysis of the GPER1 structure has been limited, as the protein is rather refractory to both X-ray crystallography and NMR due to its relatively high lipophilicity [44]. This explains the importance of computational homology modelling [45]. Homology models are later used to simulate binding with estrogens. There are only a few studies dealing directly with this subject [44,46]. Application of homology modelling, molecular docking, and molecular dynamics provides insight into the process of estrogen binding and helps to explain the induced non-genomic effects. As non-genomic estrogen actions through GPER1 and kinase cascades alter the cell membrane shape, further research on cell membranes and estrogens is also performed (see Section 2.1.3).

2. Application of Molecular Modelling Methods in the Study of Estrogens and Xenoestrogens

Molecular modelling could be of a great help to experimentalists. Ligand docking directs research to the most probable hit molecules; QSAR, an officially accepted OECD method, predicts toxicity and often can facilitate research; and MD allows one to observe the time-dependent mechanisms (e.g., in membranes) and, therefore, helps to explain the experimentally observed phenomena. What is more, two types of theories on which calculations are based deliver two levels of accuracy. These are QM and MM. The former is more accurate, but as a consequence, calculations are more time-consuming. The latter, on the contrary, allows one to obtain a general view on the topic in a shorter time. The diversity of methods highly contributes to more successful experimental studies as it saves time and points out these research approaches that have a high probability of success.

2.1. Application to Estrogens

2.1.1. Ligand Docking Using Force Field Methods

- Principles of docking and re-docking

Protein–ligand docking is a technique used to predict the orientation and active conformation of a molecule in an active center. Based on this prediction, the binding energy between protein and ligand is calculated [47]. Ligand docking can identify the chemical bonds crucial for activity and the specific atoms/residues that are responsible for them. Target proteins are usually acquired from the RCSB Protein Data Bank (PDB) [48] (Table 2). Presence of a ligand in the crystallographic receptor structure simplifies and speeds up researchers' work, as a proper region for docking is already plainly indicated.

Table 2. Selected proteins involved in the metabolism of estrogens present in the RCSB PDB (Protein Data Bank) [48]. All structures were obtained using X-ray diffraction.

Protein	RCSB PDB Reference Code	Resolution (Å)	Incorporated Ligands
βER	5TOA	2.5	Estradiol
βER-LBD	1QKM	1.8	Genistein
Phosphorylated βER-LBD	3OLL	1.5	Estradiol, N-peptide linking
αER-LBD	3UUC	2.1	Bisphenol C
αER-LBD mutant	4Q50	3.07	4-hydroxytamoxifen
αER-LBD mutant	2QXS	1.7	Raloxifene
αER-LBD	2R6Y	2.0	SERM
17β-HSD	1IOL	2.30	17β-estradiol
17β-HSD	6MNC	2.40	Estrone
17β-HSD	6MNE	1.86	Estrone, NADP+
17β-HSD	3DHE	2.30	Dehydroepiandrosterone (DHEA)
17β-HSD	4FJ0	2.2	3,7-dihydroxy flavone
17β-HSD	4FJ1	2.3	Genistein
SULT1E1	1AQU	1.6	Estradiol, PAP cofactor
SULT1E1	4JVM	1.994	Flame retardant, PAP cofactor
Placental E1/DHEA STS	1P49	2.6	L-octylglucoside, *N*-acetylo-D-glucosamine, Ca^{2+}, PO_4^{3-}
CYP1B1	6OyV	3.101	Estradiol

Accessibility of the X-ray protein structures with docked ligands gives also possibility to prove the quality of the chosen simulating docking method and parameters [49]. This is performed via extracting the ligand from the available in the PDB structure and re-docking it to this crystallographic measurements-based protein. The results are obtained in a form of the root mean square deviation (RMSD) of atom positions [49]. It is calculated as differences between the original experimental atom positions in the crystallized structure and the theoretically obtained molecular docking results [50]. It is acknowledged that the used modelling tools are able to identify the correct ligand pose, are repeatable and reliable when RMSD is smaller than 2Å [51].

Molecular modelling performed on estrogen-related receptors is not an exception. As the studies are performed mostly to search for new agonists or antagonists, the ligands re-docked into the ERs or SHBG are mainly E2 [52–57] and 4-hydroxytamoxifen [52,58,59]. RMSD value in most of the studies varies from 0.26 to 1.4Å, which proves the correctness of the docking methods in finding the proper orientation of the ligand in the active site.

- Enzymes and receptors used as targets in estrogen-related docking studies

A great many proteins have been used in molecular docking studies of estradiol. They include 17β-hydroxysteroid dehydrogenase [60], progesterone receptor [61], protein disulfide isomerase [62], SHBG [63], CYP1B1 [64], steroid sulfatase [65], and even the voltage- and Ca^{2+}-activated K^+ channel β1 subunit [66]. In all studies, an emphasis is put on the importance of OH-hydrogen binding [67,68]. The relevance of this emphasis is confirmed by experimental results. One example is a study where the estrogen analog with the highest affinity, measured in a fluorescence polarization displacement assay, appears to have the second highest predicted affinity [67].

Docking serves either to investigate the binding of estradiol to the ER or to dock other molecules, potentially new drugs, such as potent and highly selective estriol analogues [69]. In the second case, information from previous studies on estradiol behavior in an active site serves as reference data [70–73]. This is a starting point for the commonly applied research sequence: investigation of E2 binding mode, comparison with the calculation results for potential drugs, and confirmation of the hypothesis by analytical techniques. Such a three-step process has been performed, for example,

for estrogen-dependent MCF-7 cancer cells [70]. Eighteen compounds with antiproliferative activity have been docked into the cavity where E2 normally binds. It was predicted that, compared to E2, an additional aromatic ring is involved in the binding mechanism. The prediction was confirmed by site-directed mutagenesis.

In some cases, allosteric modulators have been studied. This means that estradiol must be present in the binding site during the docking process [74,75]. To test the prediction of allosteric activity, in vivo experiments are often performed. For example, in one study [74] it was shown that one particular compound can properly fit into the region of the binding pocket, along with E2. Afterwards, this ligand was investigated in vivo and was demonstrated, indeed, to be a new ER-beta-selective, negative allosteric modulator of E2 binding.

Including an estradiol molecule in the docking studies helps to properly score the resulting data [76] and rank the tested molecules according to their binding affinity [77]. For large ligand sets, high-throughput screening with a pure agonist (estradiol), an antagonist (tamoxifen), and decoys (known non-binders) can be performed. This can aid in setting a laboratory's experimental testing priority, reducing the cost and time of its research, and boosting its effectivity [77]. This explains why in vivo and in vitro investigations are often combined with in silico ones [78,79]. Altogether, it enables the discovery of new possible drug molecules that could influence estradiol's signaling pathways [80].

- Docking studies of plant-derived potential xenoestrogens

Docking plant-derived substances into ER and comparing their binding energies and interactions (above all hydrogen bonding) with those of estradiol is a common practice. Such studies deliver information on conformational flexibility [81], the ability of the investigated molecules to selectively modulate ERα/β ability [82,83], their possible reductive influence on breast cancer risk [84], or their applicability to reduce menopausal symptoms [85]. Most importantly, data derived in this way very often show good consistency with experiments [86]. For the obtained data, a correlation with estradiol, but sometimes also with tamoxifen [75] (ER antagonist) or genistein [87,88], is found. This last substance serves as an important reference, as genistein has a higher affinity for ERβ than 17β-estradiol.

Even if docking into ER is most widely used, modulation of CYP450 activity by plant-derived substances with regard to estrogenic effects has also been examined [89]. Indeed, the same signaling pathways are regulated by CYP450–estradiol interactions, and there are plenty of data available on this topic. Moreover, other calculation methods are also applied, including ADMET and molecular dynamics (MD) [90,91] (for MD description and examples, see Section 2.1.3). Nowadays, in silico methods are a standard tool in plant xenoestrogens studies. Thus, molecular modelling is often performed in parallel to either in vitro [92] or in vivo toxicity studies [93], and in most cases an agreement between data obtained from both sources is found. This was the case in the docking-based binding affinities of compounds derived from *C. elegans* and their measured reproductive toxicity [93]. Wide examination of many substances of a natural origin is possible thanks to the thorough structural knowledge on the estradiol molecule bound to the ER (Table 2).

2.1.2. Quantitative Structure–Activity Relationship (QSAR)

QSAR methods use mathematical models to correlate structural characteristics with the biological activity of a set of compounds that have closely related structures [94]. Empirical and theoretical molecular descriptors are used. Conceptually, three main types of QSARs are known: ones based on fragment analysis of a system (here, a pharmacophore [95] concept is used), ones that consider the given system as a whole (descriptors are computed from scalar quantities), and 3D-QSAR. In this last type, descriptors are obtained by application of a force field (3D approach). To achieve a 3D target structure, either software-based alignment or manual superimposition on the crystallographic data must be performed [96]. This is a necessary step, as different ligand-binding modes and bioactive conformations are possible. Examples of 3D-QSAR are Comparative Molecular Field Analysis (CoMFA) [97] and

Comparative Molecular Similarity Indices Analysis (CoMSIA) [98]. From the created QSAR models, predictions on the bio-activity and toxicity of other molecules are made.

A smaller group of structure–activity studies are ones that are non-quantitative, namely SAR ones (structure–activity relationship). They include virtual screening with a pharmacophore and a large set of molecules. Afterwards, in vitro evaluation of the data is performed. This methodology has been used to determine potential 17β-HSD inhibitors [99,100] that could be applied in the treatment of osteoporosis provoked by estradiol deficiency.

Among all molecular modelling approaches, QSAR is one of the most commonly used to examine estradiol's structure and activity. Often, data on estradiol serve only as a reference for information gathered on new possible drugs [14,101] such as raloxifene derivatives [102] or estradiol metabolites. Through comparison with experiments, it has been shown that QSAR models have a high sensitivity and specificity in providing relative binding affinities (RBAs), where estradiol's RBA equals 100% [103]. Among the most important descriptors are ones calculated with quantum mechanics at the DFT level [104,105].

While QSAR is mostly applied to ER-binding, other proteins like CYP1A2 [106] and CYP1B1 [107] have also been targeted. One study used the CoMFA approach to model estriol's influence on CYP1A1 [106]. It focused on the inhibition of estradiol to mutagenic 4-OH estradiol transformation. Out of 90 steroid candidates, thioestrone was selected and shown to have the desired inhibitory ability. Its mechanism of action was revealed by molecular modelling. It is desired because thioestrone's -SH group is closer to the iron atom in the CYP1B1 heme than the -OH group in natural ligands of the enzyme, namely estradiol and estrone.

OECD Guidance identifies QSAR as an important element in toxicity evaluations. As a result, QSAR plays an important role in substance risk assessment [108]. It has been applied in studies dealing with mutations in enzymes that metabolize estradiol [109] and to study estradiol oxidation and emerging contaminants [110]. In the latter study, the most accurate DFT descriptors were used. This enabled a better understanding of the degradation mechanisms.

2.1.3. Advanced Docking Using Combined Quantum Mechanics/Molecular Mechanics (QM/MM) or Molecular Dynamics (MD) Methods

Molecular modelling with QM is much more accurate than MM methods. However, the direct use QM approaches in drug design is limited due to the cost and size of protein structures. In recent years, QM/MM has been gaining attention, as it allows one to consider a whole protein–ligand complex and not only a binding site [111,112]. Due to computational limitations, MM calculations on the outer part of a receptor deliver only approximate data. Nevertheless, QM/MM provides more knowledge on the protein's influence than is obtained when only the LBD region is considered.

More commonly applied methodology is MD. It simulates time-dependent processes and provides data that are otherwise unavailable [52,54]. For example, in case of ERs or SHBG ligands MD can point out whether the examined substance is receptor's agonist or antagonist. This assumption is based on the RMSF (root mean square fluctuation) value which is extracted from the MD trajectories. RMSF represents the flexibility of the amino acid residues [50]. Both RMSF and RMSD depend on the interactions between the protein and the ligand and are the result of the ligand movements in the active site trying to achieve the appropriate position [63]. If the same ligand undergoes the MD process in both ERα and ERβ and its RMSF and RMSD values significantly differ for these two receptors, it indicates that the investigated molecule probably occupies more favorably one of the investigated receptors. Such data leads to the hypothesis of the ligand's selectivity.

Moreover, in terms of the ERs, RMSF and RMSD values suggest whether the analyzed molecule is the receptor's agonist or antagonist [50,52–54,59]. As already mentioned in Section 1.2, the positioning of the H12 helix is differently influenced by agonists and antagonists. If the estrogenicity of the compound is known, comparison of RMSF values with the known molecule's estrogenic effect, can serve as an evaluation of the applied docking parameters.

Nevertheless, MD requires a lot of computation time, which increases with the system's size and simulation length. In return, it describes the dynamics of the system as well as entropic effects associated with the protein–ligand interaction.

MD can also provide quantitative estimates of relative binding affinities through techniques known as "free energy calculations". The most rigorous is the "free energy perturbation," or FEP family of methods [113–117]. To compare two ligands, say A and B, one introduces a model where both ligands are present, each with a partial occupancy. This is closely analogous to a crystallographic refinement where a particular group (ligand or side chain, say) has two possible conformations. Each ligand is assigned a weight between 0 and 1, say w_A and $w_B = 1 - w_A$. These multiply terms in the energy function involving either ligand. By varying the weights gradually, one can effectively remove one ligand and introduce the other. Thus, when $w_B = 0$, ligand A is fully weighted while interactions of ligand B with its surroundings have a zero weight: ligand B is "invisible" to its environment. Usually, interactions within the ligand are not weighted. As w_B changes from zero to one, B is introduced and A is removed. Intermediate weight values correspond to an "alchemical" mixture, where both ligands are partially present. MD simulations (or Monte Carlo simulations) are performed for a few w_A values, typically around 10. This series of simulations mimics a gradual, reversible replacement of A by B. Energy statistics are collected from all simulations. From these, a free energy difference between A and B can be obtained. The same process is carried out for the unbound ligands, solvated by a box of water. Subtracting the bound and unbound free energy changes yields the binding free energy difference. The method requires force field parameters for each ligand but has no other adjustable parameters. The tradeoff is that not one, but several MD simulations are required, and these should be sufficiently long. Indeed, FEP accuracy is limited by the MM force field, but also the amount of conformational sampling that is carried out. The method can also be used to compute the binding free energy changes due to point mutations of the protein. In this case, a particular residue is modeled with two side chains, each having a partial occupancy. Nowadays, FEP can be applied to one ligand or mutation per day on a medium-sized computer cluster.

A simplified version of FEP is to use only two simulations per ligand: one bound to the protein and one in solution. From these, binding free energies can also be obtained, if one is willing to extrapolate from a w_A value of unity to a value of zero. To counter the use of such a large extrapolation, one introduces empirical weighting factors that multiply the interactions between the ligand and its surroundings (protein or solution). Usually one is applied to the electrostatic interactions and one to the Lennard–Jones interactions. The extrapolation and use of interaction energies have led to the name Linear Interaction Energy, or LIE method [118–120]. For a thorough review of its theoretical basis, see [121]. The tradeoff for its speed is that experimental data are needed to adjust the values of the empirical weights, which are not very transferable between different proteins and classes of ligands. Once the weights are optimized for the molecules of interest, predictions can be made.

In many studies, estradiol has been simulated in a complex with ER. The QM/MM approach seems to be the most accurate. It uses a QM description of the ligand and its binding pocket, but thanks to the MM description of more distant protein regions, it preserves information on the whole enzyme's impact [122]. Very recently, QM/MM elucidated the important role of estradiol's D-ring in the active site of ERα [123]. What is more, it helped to understand the influence of each enzyme segment on the ERα-agonist (estradiol, diethylsilbestrol) binding [124]. Most importantly, the binding energies of E2 and DES correlated well with experimental agonist binding affinities for the ER.

To analyze the changes in the investigated complexes upon DNA binding, not only MM, as with estradiol and DNA [125], but also MD has been applied [126]. The latter study identified specific bases within the aptamer (short-stranded DNA/RNA, binds only specific molecules [127]) and demonstrated the importance of water-mediated hydrogen bonds in the aptamer–estradiol complex.

MD is also often used to describe the effects of enzyme mutations on ligand binding. This methodology has been applied in estradiol studies, to compare wildtype and mutated CYP1B1 [128], which is mostly responsible for the 2-OH-hydroxylation of estradiol. MD serves also as a tool to explain

the results of molecular docking. For example, it has been stated that the interactions between human α-fetoprotein and agonists (estradiol, estrone, diethylsilbestrol) were caused by van der Waals forces, whereas binding of antagonists (tamoxifen and its analogues) was equally based on hydrophobic and electrostatic interactions [129]. Moreover, information from MD simulations can be used to construct a pharmacophore in order to screen protein databases for a desired type of ligand. This methodology was used to search for substrates and inhibitors of the estrone-SULT [130]. Nine selected molecules were consistent with the ones indicated by the experiment.

Another nuclear receptor, the farnesoid X receptor (FXR), was used as a test case for free energy calculations in 2018 [116,117,120]. FXR is involved in regulating bile acid, lipid, and glucose homeostasis [131], and it has been linked to hepatocarcinogenesis [132]. Its hormone binding site is hydrophobic with few conserved interaction motifs and strong induced fit effects. With FEP, mean errors were about 1.5 kcal/mol for relative binding free energies of around 30 ligands, and the largest errors were about 2.5 kcal/mol. In one of the studies [117], ligands were first docked to the receptor, then compared using FEP. LIE gave similar errors for the same protein and 47 ligands [120], at a lower computational cost, but required optimization of the two adjustable LIE parameters using a subset of the ligands.

An earlier LIE study [119] considered the ER binding of estradiol and a series of xenoestrogens. A training set of 19 ligands was used to optimize the LIE parameters. A mean unsigned error of 0.6 kcal/mol was then obtained for a test set of 13 ligands. Several binding poses (3–4) were considered for each ligand; this was not too expensive because only short MD simulations were run.

2.1.4. Other MD-Based Studies of Estrogens

- Membranes

As ERs are located in either the cytoplasm or the lipid bilayer (mER, membrane ER) [27,133], and estradiol itself is a steroid hormone, closer insight into the ligand's interaction with the cell membrane is an obvious research target. Since transfer through this cellular barrier is not a stable state, but a dynamic process, MD calculations could be seen as a method of choice. Recent studies revealed [134] that, regarding estradiol's long axis and the lipid acyl chains, E2 adopts a perpendicular position in the membrane. By having four rings located near the membrane interface, participation of the hormone's 3-OH and 17β-OH groups in hydrogen bonds and electrostatic interactions with the lipids are possible.

Combining MD and QM enables further research into E2 membrane crossing. It provides information not only about the E2 orientation in the membrane (MD) but also about the strength of the electrostatic potential mapped on the electron density surface (QM) [135]. These data, derived from a HDL disc model, enable deeper insight into the mechanism of E2 incorporation into lipid membranes and is an important step forward to develop tissue-specific discs encircled by a membrane, which would serve as transporters for E2 or its derivatives.

Similar methodology (QM, MD) was used to study the removal of hormonal pollutants from water. It has been shown that by using high levels of salinity, which increases the strength of hydrogen bonding and hydrophobic interactions, one can perform a membrane-based sorption of 17α-ethinyl estradiol on the polyethersulfone membrane [136].

- Nanotubes

Nowadays, estradiol is a relatively common water pollutant, and numerous studies have been performed to find a reliable tool for its removal. One possibility is to use nanotubes. For that purpose, within the molecular modelling approach, mostly single-walled carbon nanotubes (SWNTs) [137] are used. Free energies of adsorption have been calculated with a QM approach. In some cases, MD was also implemented [138]. The target ligands for these studies were 17β-estradiol and its medically useful derivative 17α-ethinyl-estradiol [139–141]. These simulations revealed the adsorption energy,

a preferential sorption among different nanotubes and estradiol derivatives, and provided a molecular explanation for the observed results [142].

2.1.5. Density Functional Theory (DFT) Calculations in the Study of Estrogens

DFT is a QM approach that determines the ground-state properties of a many-body system by applying the electron density concept. The underlying concept is the Hohenberg–Kohn theorem [143], later developed into the Kohn–Sham theory (KS-DFT) [144]. Firstly, the energy of the system for a non-degenerate stationary state is uniquely determined by its electron density, which depends on three spatial coordinates. For this reason, the energy is expressed as a functional of the (scalar) electron density function. Secondly, according to the H-K theorems, the minimum energy occurs for a unique, precise electron density in the ground state. KS-DFT includes the Coulomb interactions between electrons and considers the energy of the exchange and correlation interactions. For a long time, the dispersion energy (the energy of the long-ranged electron correlation) [145] represented a difficult problem, as it is a time-dependent phenomenon, and it was not included in KS-DFT. However, nowadays, dispersion corrections are available and can be included in DFT functions [146]. Therefore, application of DFT leads to the highest obtainable accuracy in calculations. The only existing drawbacks are the risk of underestimating the energy [147] and the time needed to acquire the results. DFT-based calculations are especially widely used in solid-state studies.

Although application of DFT is mostly concentrated on the investigation of single molecules (see the paragraphs below), it is also used to determine total binding energies between a ligand and protein. This is the case for systems composed of ER, SERMs, and two widely used ER antagonists: 4-hydroxy tamoxifen (4OH-T) and raloxifene (RAL) [148]. The results show that the 4OH-T-SERMs set binds more strongly to the ER than the RAL-SERMs set. This is fully in agreement with the experimental data and, once more, as many other studies, confirms the high accuracy of the DFT calculations.

- Crystal structure prediction

DFT is the theoretical basis for periodic calculations performed on solids, often pharmaceuticals, in order to find and depict new polymorphic forms of drugs or potentially bioactive molecules [149]. It can also be a part of the Crystal Structure Prediction (CSP) approach. Such methodology has been recently used to study crystal structures of 17β-estradiol. As a result, an estradiol hemihydrate has been computationally determined [150]. DFT calculations are also often necessary to refine a crystal structure obtained from powder X-ray diffraction (PXRD) experiments. These calculations are mostly consistent with experimental data, as in the case of estradiol ethinyl cocrystals [151].

DFT-based methodology has also been applied to examine the dissolution process in a study of estradiol cocrystals [152] and to calculate the free energy of solvation in the estradiol–ER complex [153].

- NMR and vibrational properties calculations

One of the most common DFT applications is to predict NMR and vibrational data. For NMR properties, GIAO [154] and GIPAW [155] methods are implemented. Estradiol being relatively complex (many atoms, presence of both aromatic and non-aromatic rings) could cause computational problems. For this reason, it has been part of a many-ligand study to prove the applicability of the GIAO method as well as its ability to calculate J_{HH}, J_{HC}, and J_{CC} NMR coupling constants [156,157]. Generally, the results obtained were satisfying. However, nowadays it is more common to study the solid state with the second method. GIPAW NMR calculations were also done recently for a new polymorph of 17β-estradiol [158]. The calculations helped to improve the assignment accuracy of chemical shifts obtained from the experiment and, therefore, to elucidate the structure of a new anhydrous estradiol form Figure 1.

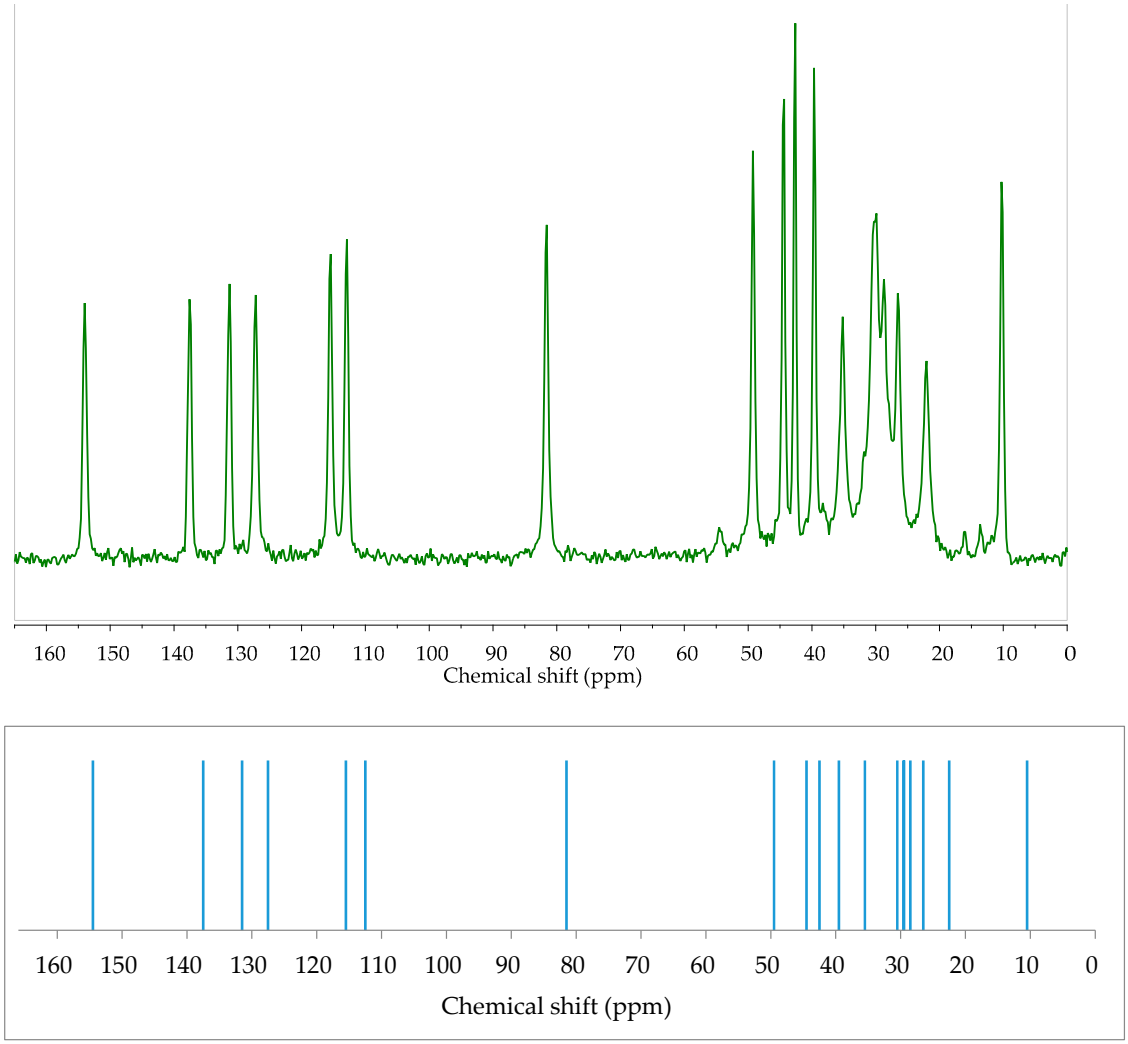


Figure 1. Experimental (top, green) Cross Polarization Magic Angle Spinning (CP MAS) and calculated (bottom, blue) GIPAW ^{13}C solid-state NMR spectra of E2. Very good agreement between calculated and experimental values proves the usefulness of DFT calculations in solid-state analysis of estrogens. More details in [158]. Source: Author's archive.

Another study showed the applicability of DFT-NMR to explain the nature of a more complex system: the transversal distribution of 17β-estradiol in lipid membranes [159]. NOESY 2D NMR spectra contained cross-peaks between the hormone and lipids. Here, too, DFT calculations helped to properly interpret the experimental data and, as a consequence, greatly aided in describing the position of the estradiol aromatic ring in the membrane. An implication of such study is to increase our understanding of estradiol's transfer through a lipid bilayer [159,160]. A second area where DFT methodology is widely applied is in calculating vibrational properties. The direct usefulness of the method is manifested through its contribution to accurate assignment of the vibrational modes in IR or Raman studies. This has been implemented, for example, in investigations of estrogens [160] and estradiol-17 valerate [161]. In most reported cases, computationally generated spectra were in very good agreement with the experimental data.

From a wider perspective, DFT calculations enable one to properly describe the examined subject. Additionally, in case of estrogens and estrogen derivatives DFT-derived spectra (both vibrational [162,163] and NMR [158]) stay in a good agreement with the experimental data and are often the only way to properly assign bands (Figure 2). In order to obtain theoretical spectra which ideally meet the experimental ones, scaling factors must be implemented [164]. Thanks to such

combination of the theoretical and experimental approach, the first full interpretation of estriol IR spectrum could have been published [164]. The analogical situation has been reported for E1, E2, E3, and ethynylestradiol Raman spectra [163]. Simulation was necessary to identify unique marker peaks in the finger-print region what was useful to differentiate between very similar estrogen structures.

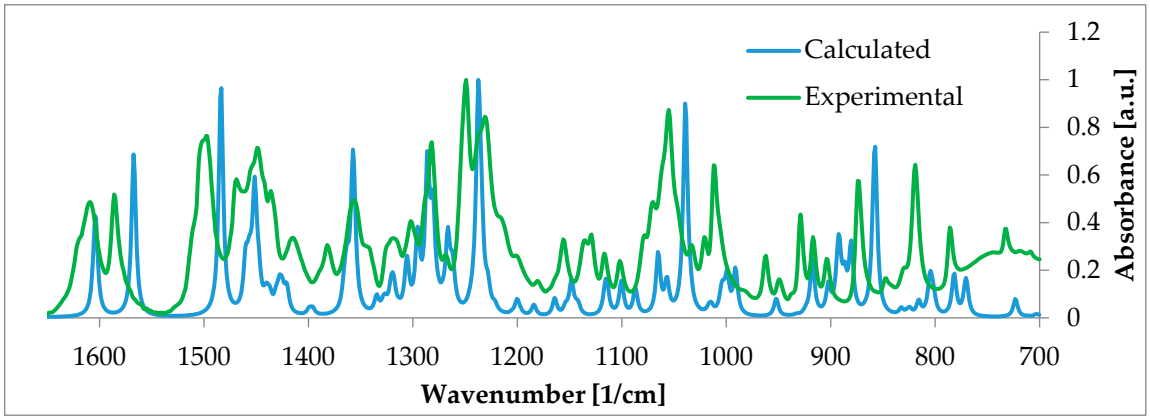


Figure 2. Experimental (green) and calculated (blue) IR spectra of β-estradiol hemihydrate, selected range 700–1650 cm^{-1}. Such calculations enable proper band assignments and thus facilitate the spectrum analysis. More information can be found in [158]. Source: Author's archive.

Another example shows that the vibrational frequency calculations of estradiol alone and in monohydrated form [165] has given insight into hydrogen bond formation by estradiol's D-ring. This has been used to discuss the relationship between the stability of hydrated clusters and the estradiol conformation. In another study [166], DFT-based IR and Raman frequencies were used to investigate estradiol and tamoxifen structures. This helped to understand the intermolecular interactions made by these two molecules and to interpret opposite estrogenic effects. However, it should be mentioned that in the studies of estrogens, Raman spectroscopy is used less frequently than IR. This is due to the inherently weaker signals and common presence of fluorescence interference from the contaminants [163]. Apart from IR and Raman, low-frequency vibrations could also be calculated with the help of DFT. A good example is the assignment of vibrational modes in terahertz spectra for testosterone, estradiol, and estrone [167].

- Removal of estrogenic pollutants

With regard to estrogen removal from the environment, QM calculations are performed not only on nanotubes but also on other sorbents (e.g., lignocellulosic material) [168]. The adsorption energy shows that this adsorbent could be used to remove all three main estrogens, E1, E2, and E3, from a solution by means of liquid phase extraction. Another extended study revealed the applicability of reduced graphene oxide modified with silver nanoparticles in electrochemical detection of estriol [169]. Firstly, MD simulations at 1000 K were performed in order to obtain 100 conformers of estriol. Later, these structures were used in a semi-empirical Hartree–Fock geometry that was pre-optimized with solvent simulated via the conductor-like Screening Model COSMO [170]. Afterwards, the most stable conformers were fully optimized by the DFT-based software. In this step, the solvent was included via application of the Polarizable Continuum Model [171]. Then, on the structure with the lowest energy, MD at 300 K was performed. Electric properties of the newly obtained conformers were determined with DFT calculations. This study is a good example of the wide range of in silico methods that can be applied to molecularly model the investigated subject (i.e., to develop a method to detect estriol in tap water and urine samples).

2.2. Application of Various Molecular Modelling Methods in the Study of Xenoestrogens

2.2.1. Various Molecular Modelling Methods Applied in Xenoestrogen Studies

The same methods as for estradiol, above, are also used to investigate xenoestrogens. They include QM/MM [172] and QSAR approaches [173–175] and confirm the importance of the hydrogen bond between xenoestrogen molecules and the His524 residue in the active site. A variety of xenoestrogens, such as bisphenol A and C, butylparaben, 4-octylphenol, DDE, phthalate, zaeranalol, estradiol, 4-OH-tamoxifen [176], and several proteins (ER [177], SHBG [178]), have been studied. These studies either looked for structural similarities between different ligands or focused on one specific molecule, such as zearalenone [69]. The latter study suggested a lack of agonistic activity against the ER due to the lack of any stable, functionally active conformation of the tested molecule in the LBD. On the other hand, according to QSAR analyses performed on zearalenone analogues [179] and metabolites [180], these zearalenone-related compounds show some estrogenicity due to the presence of a keto/hydroxyl group, a trans double bond in the macrolide ring, and two hydroxyl groups in the aromatic ring, which participate in binding to ER.

As endocrine-disrupting chemicals are present in the environment, research into their removal is constantly being performed. For example, to study dioxin adsorption on graphene, DFT [181,182] and MD [183] calculations have been undertaken.

2.2.2. Bisphenol A

- Bisphenol A (BPA)–ER complex studies

Bisphenol analogues are among the most studied xenoestrogens. Molecular modelling helps show how these closely related structures adopt agonist/antagonist orientations in the estradiol binding pocket [184] and delineate the binding modes of each bisphenol molecule [185]. A deeper study, concentrating not only on the pure ligand–receptor binding but also including the allosteric effects and application of MD calculations, showed that BPA causes changes in a full-size receptor, and its effect is not limited to the separate domains [186].

A separate set of studies investigated BPAs with halogen substituents on the phenolic rings. All studies showed that a hydrogen bond with His524 for an agonist and with Thr347 for an antagonist was created, exactly as in the case of estradiol and 4-OH-tamoxifen, respectively [187,188]. What is more, just as for estradiol, the stability of helix H12 is crucial in halogenated BPAs–ER complexes [188,189]. The in silico results have been confirmed by experimentally measured affinities.

Although BPA mimics estradiol's action in the LBD, the QM/MM study revealed that, in comparison to other tested EDCs, it exhibited lower estrogenic activity, probably due to the lack of interaction with His524 [190]. In turn, application of MD helped to elucidate mechanisms driving BPA–ER binding. According to that study, direct hydrogen bonds and hydrophobic interactions are responsible for the binding [173]. Like the previously mentioned experiment, this one confirmed that the ER binding affinity is slightly lower for BPA than for estradiol. MD not only helps to elucidate bound conformations and binding energies between LBD and BPA in the ER [191], but it also gives insight into the influence of that binding on the whole protein, including the DBD. One of the studies [185] showed that the allosteric effects in the LBD due to BPA binding could cause relaxation of the DBD and, therefore, alter ER's function. Other researchers reported the influence of bisphenol compounds on the protein's allosteric modulation, altering the Helix12 stability and reducing the recruitment potency of co-activators [187]. This knowledge can be useful in the process of estimating the toxicity of compounds.

- Risk assessment and removal attempts

In 2012, a protocol for in silico risk assessment of BPA on the ER was proposed [189]. Later, many studies dealing with BPA removal from water were performed [192,193]. Applied molecular

modelling techniques include DFT and MD. Since BPA results from the depolymerization of, for example, polycarbonates, both BPA and the initial polymer have been studied under periodic boundary conditions (pbc) [194]. The same DFT-pbc approach has been used to model the photocatalytic degradation of BPA caused by cobalt-doped BiOCl nanosheets [195] and by the effect of humidity combined with UV irradiation [196]. Another example is the evaluation of BPA's binding to microextraction coatings [197]. In this case, dispersion-corrected DFT was applied.

2.2.3. Phthalates

The second group of xenoestrogens most widely examined by molecular modelling are the phthalates [198,199]. In silico investigations showed that estradiol has a lower ER binding affinity than phthalates. The highest RBA is exhibited by monophthalates [200]. In the case of SHBG binding, score values suggest that short-chain phthalates are more potent than long-chain ones [201]. This agrees with known experimental data.

To look for an effective method of phthalate removal from water, as with estradiol, DFT and MD calculations on SWNT–pollutant complexes have been performed [202]. The adsorption energy has been calculated, and the adsorbent's chemical groups responsible for the binding have been determined. MD has also been used to examine polymer–solvent interactions while looking for a new, more eco-friendly substitute for plasticizers [203,204].

2.2.4. Technical Aspects of Calculations Performed on (xeno)Estrogens

The computational method most commonly applied in the analysis of (xeno)estrogens is molecular docking. Available publications show that, for this purpose, the most common software packages are Maestro Schrödinger and AutoDockTools. The former is also widely applied in Virtual Screening Workflow (Table 3, N° 9–11), which includes high-throughput virtual screening (HTVS) [205] and molecular docking with either standard or extra precision (SP, XP). In both cases, the OPLS 2005 force field is used. The mentioned (N° 9) consensus score is an effective score that enables ranking of the investigated ligands. It is a combination of different scores, like DockingScore (GlideScore + state penalties for protonation) [206], MM/GBSA Score (binding free energy calculations based on the MD trajectories) [207], and QSAR Score.

Table 3. Selected technical computation data in terms of ER and (xeno)estrogens regarding the publications cited in this article.

N°	Code/Software Used	Force Field or DFT Functional and Basis Set	Type of Calculation	Ref. Method	Ref. in Article
1	GOLD		Molecular docking	[208]	[57,69]
2	-Ghemical 2.95 -Swiss Dock	-Tripos 5.2 -CHARMM	-Geometry optimization -Molecular docking	[209–212]	[70]
3	-Swiss model -Hex 8.0, HADDOCK	-OPLS	**-Homology of receptors** -Molecular docking	[212–215]	[87]
4	-Swiss model -SybylX	-Tripos 5.2	**-Homology of receptors** -Molecular docking	[209,214,216]	[84]
5	Maestro Schrödinger	OPLS 2005, Glide SP, XP	Molecular docking	[217,218]	[83,85,86]
6	Maestro Schrödinger	MMFF94	Geometry optimization, molecular docking	[217–219]	
7	-Gaussian09W -AutoDockTools	-B3LYP/6-31G(d) -AutoDockZN	-Geometry optimization -Molecular docking	[220–223]	[169,184]

Table 3. *Cont.*

N°	Code/Software Used	Force Field or DFT Functional and Basis Set	Type of Calculation	Ref. Method	Ref. in Article
8	Gaussian03	B3LYP/6-311++g**, PCM	Hydration enthalpy	[220,224]	[183]
9	Maestro Schrödinger	ZINC database OPLS 2005, Glide SP eHiTS docking module consensus score	Energy minimization **HTVS** rank	[205,217,218, 225]	[75]
10	Maestro Schrödinger	OPLS 2005, Glide SP, XP	**HTVS**	[205,217,218]	[76]
11	Maestro Schrödinger	OPLS 2005, Glide SP, XP	**Segregation**: agonists/antagonists	[217,218]	[77]
12	Maestro Schrödinger	-OPLS 2005, Grid (Glide) -Desmond OPLS 2005	-Molecular docking, MD -ADMET parameters	[217,218]	[90]
13	-Maestro Schrödinger -AMBER14 -AMBER14	-OPLS 2005, Glide -FF03 (protein) GAFF (ligand) -MMPBSA, MMGBSA	-Docking -MD -Binding free energy, decomposition energy	[207,217,218, 226–228]	[91]
14	-MOPAC2016 -Gaussian09 -Gabedit package	-PM6 in HF, COSMO model -B3LYP, PCM -Verlet algorithm	-Pre-optimization, solvent model -Optimization (DFT), solvent model -MD	[156,220,224, 229,230]	[169]
15	-GOLD -GROMACS -Swiss Param Tool	-CHARMM27 -CHARMM27	-Molecular docking -MD -Ligand parametrization	[209–211,231]	[180]
16	-SybylX -AMBER11 -AutoDock 4.0	-Tripos 5.2 -AMBER -AutoDockZN	-Geometry optimization -MD -Molecular docking	[216,221–223, 226–228]	[186]
17	-Gaussian09 -LeDock -AMBER12 -AmberTools14	-B3LYP/-cc-pVTZ -CHARMM -AMBER -MM/GBSA	-Geometry optimization -Molecular docking -MD -Binding free energy	[207,220–223, 226–228]	[188]
18	-Gaussian09 -Molegro Virtual Docker -AMBER Tools	-B3LYP/6-311++G(d,p) -AMBER -AMBER03	-Molecular electrostatic potential -Molecular docking -MD	[220,226–228]	[187]
19	-Maestro Schrödinger -Gaussian09 -AMBER10	-OPLS 2005 -HF, 6–31G* -GAFF (ligand), ff03 (protein)	-Molecular docking -Geometry optimization -MD	[217,218,220, 226–228]	[189]
20	-VASP -GROMACS	-PBE GGA (DFT-D3) -GROMOS96	-Geometry optimization -MD	[230,232]	[192]
21	-GROMACS -AutoDock Tools -AutoDock Vina, Hex8.0.0 GROMACS	-AutoDockZN -AutoDock Vina, GROMOS96	-Energy minimization -Molecular docking -MD	[221–223,231]	[199]

Table 3. *Cont.*

N°	Code/Software Used	Force Field or DFT Functional and Basis Set	Type of Calculation	Ref. Method	Ref. in Article
22	-NAMD -Spartan04	-Charm CMAP FF -HF 3-21G	-MD -QM	[233,234]	[125]
23	-Gaussian 03 -AutoDock -AMBBER	-B3LYP/6311**G -AutoDockZN -PM3/Amberff14SB FF	-Geometry optimization -Molecular docking -QM/MM	[220–223,226–228]	[114]
24	-GROMACS -Gaussian 09	-CHARMM (MM) -GGA-D2 (QM)	-Geometry optimization -DFT calculations	[210,211,220–223,229]	[115]
25	-Maestro Schrödinger -AMBER Tools	-OPLS 2005 Glide -B3LYP/Amberff14SB	-Protein, ligand preparation (geometry optimization), molecular docking -QM/MM	[217,218,226–228]	[124]
26	-Crystal Predictor -Crystal Optimizer (Gaussian) -DMACRYS	-PBE0/631G(d,p)	-Conformations -Geometry optimization **CSP** -Intermolecular lattice energies	[220,235–237]	[150]
27	-GULP, DFTB+ -VASP	-optB88 level	-Geometry pre-optimization **CSP** -Geometry re-optimization	[232,235–237]	[181]
28	DMol3	DNP basis set, PBE GGA	Geometry, energy optimization	[238,239]	[182]
29	CASTEP	GGA PBE	DFT, NMR	[239,240]	[157]
30	CASTEP	GGA PBE	DFT, structure parameters calculation	[239,240]	[195]
31	Gaussian09W	B3LYP/631G(d)	DFT, IR	[238,241]	[164]
32	Gaussian09	M05-2X/6-311++G**	DFT, IR	[238,241]	[165]
33	Gaussian09W	B3LYP/6-31G (d,p)	DFT, Raman	[238,241]	[166]
34	Gaussian	B3LYP/6-31G(d,p)	DFT, IR	[238,241]	[167]

The most widely used are commercial codes: Maestro Schrödinger [217], CASTEP [240], GOLD [208], Gaussian [220], AMBER [226], SybylX, VASP [220]; academic codes: AutoDock [209], CHARMM [210], GROMACS [232]. (AMBER and CHARMM are names of both the codes and the force fields.) AMBER (Assisted Model Building with Energy Refinement), CHARMM (Chemistry at HARvard Macromolecular Mechanics), CSP (Crystal Structure Prediction), GAFF (General AMBER Force Field), Glide (Grid-based Ligand Docking with Energetics), GOLD (Genetic Optimization for Ligand Docking), GROMACS (GROningen MAchine for Chemical Simulations), HADDOCK (High Ambiguity Driven protein-protein DOCKing), HTVS (high throughput virtual screening), MM/GBSA (Molecular Mechanics/Generalized Born Surface Area), PCM (polarizable continuum model), VASP (Vienna ab initio Simulation Package), Glide SP (Standard Precision), XP (Extra Precision). References in the last column refer to articles already cited in this review. These are examples of application of the listed methods in (xeno)estrogens research. References in the fourth column refer to articles that describe the theoretical basis of the listed software and calculation methods.

Virtual Screening Workflow makes it possible to screen large sets of ligands. It helps to differentiate between 'actives' (compounds active against the target protein), 'decoys' (compounds of known non-activity against the target protein), and 'inhibitors/activators' (potential bio-active substances). As a consequence, the Virtual Screening Workflow approach guides future ligand synthesis and helps in setting a priority for in vitro testing.

Often, a computational step following molecular docking is MD. Here, the most applied codes are GROMACS and AMBER using CHARMM and AMBER force fields, respectively. In most cases, the TIP3 (transferable intermolecular potential with three points) solvent model is applied [242].

Computational methods applied in (xeno)estrogens studies that deal not with a solvent environment but with a solid state are Crystal Structure Prediction (CSP) (N° 26, 27) [235–237] and DFT-based calculation of spectroscopic (IR, Raman, NMR) properties (N° 28–34). For the latter, the most commonly applied codes are CASTEP and Gaussian with GGA PBE [239] and B3LYP [241] functionals, respectively. These two functionals seem to be the most reasonable for the investigated subjects. GGA PBE establishes the non-homogeneity in the electron density and leads to more precise results, which in turn is of high importance in NMR spectra calculations. B3LYP is a hybrid functional based on combining DFT and Hartree–Fock theories and finds its application in spectroscopic spectra generation.

On the contrary, CSP [235–237] is a multi-step and much more complicated methodology, as it implements both MM and QM and sometimes even MD. Firstly, MM calculations are performed to generate and rank possible compound conformations. Afterwards, the selected conformers are subjected either to ab initio calculations on a molecule or to DFT-D (dispersion corrected) [146] calculations performed on the whole crystal structure. This enables one to observe conformational polymorphisms in the first case and packing polymorphisms in the second. The lattice energies obtained could be adjusted if kinetic factors (like temperature) are included. For that purpose, time- and computational power-consuming MD must be applied.

In contrast to molecular docking or MD simulations, one of the most important calculation methodologies for (xeno)estrogens, QSAR, is independent of the protein and based solely on the ligand structure. This explains why different codes must be applied for QSAR. Their application to (xeno)estrogens has already been described in detail, and the available codes have been compared in [14].

The above-mentioned codes and parameters cover the most common calculations. However, it is impossible to point out the best ones due to the insufficient number of studies. We can only observe that, out of the gathered data, some standard methodologies emerge.

However, a comparison of the applied methodologies in terms of their usage as well as their advantages and drawbacks is possible. The most important aspects have been gathered in Table 4.

Table 4. Comparison of the calculation methods used in (xeno)estrogen investigations.

Calculation Method	Pros and Capabilities	Cons and Limitations
Molecular docking	-Explanation of a molecular basis for protein–ligand binding -Relatively short calculation time -Enables virtual screening for active compounds	-Lower accuracy when compared to QM methods -Significant increase in time and complexity of calculations when combined with QM (QM/MM)
QSAR	-Evaluation of estrogenicity -No protein preparation needed	-No receptor–ligand binding data -Large set of high-quality experimental data needed to obtain accurate results
QM (DFT-D)	-High accuracy of calculations -Simulation of IR, Raman, NMR spectra -Thermodynamic calculations	-Long calculation time -A lot of computational power needed -Usually limited to small molecules and systems such as estrogen complexes, salts, co-crystals, etc.
QM/MM	-High accuracy of calculations in the binding area (QM) -Consideration of a whole complex (protein–ligand) with emphasis on the binding pocket	-Calculation time elongated due to QM -Limitation of the QM-calculated area
MD	-Simulation of dynamical processes -Possibility to perform DFT-MD	-Significantly longer time required

To predict the binding affinities or the interactions between the (xeno)estrogens and biomacromolecules, either simple molecular docking or more sophisticated methods such as QM/MM, MD/MM, or FEP can be used. Notably, the more sophisticated methods require not only more specialized software but more computational time and power. Since, to the best of our knowledge, no study has been reported comparing the accuracy of various ligand docking methods applied to the particular group of (xeno)estrogens, no specific indications can be provided.

When focusing on the structural and physicochemical properties of estrogens and xenoestrogens, DFT-based methods have proven their high accuracy and reasonable calculation time. Therefore, such computations can be performed to obtain structural, spectroscopic (IR, Raman, NMR), and thermodynamic data of estrogens, xenoestrogens, their complexes, and solid-state forms such as solvates, salts, and co-crystals. Standard DFT functions (B3LYP for isolated compounds and PBE for periodic structures) have been found to be accurate in multiple studies.

3. Conclusions

In this review, it was clearly shown that molecular modelling methods are valuable tools in studies on estrogens and xenoestrogens. Their relatively low cost, requiring only certain specialized software licenses and computing servers, their increased personal and environmental safety, and their reasonable accuracy make molecular modelling methods unique and modern tools for these studies. In this article, the most common biomolecules studied using molecular modelling were presented, together with appropriate references to the published results. This group of molecules is composed mostly of enzymes participating in the metabolism of estrogens, along with estrogen receptors and even specific nucleic acid domains. While most studies focused on predicting the affinities of small molecules (ligands) to the chosen receptors, the computational research is not limited to this aspect. Another important role for modelling is to explain the conformational changes resulting from binding. Such in silico studies would not be possible without the very large number of already deposited, high-quality crystal structures of estrogen-related proteins that can be easily accessed and used in molecular modelling studies. An overview of those structures was presented in this review. The oldest studies in which molecular modelling was used to study the biochemistry of estrogens focused on molecular docking with molecular mechanics. More recent studies have used more sophisticated methods such as molecular dynamics or combinations of quantum mechanics and molecular mechanics. Further, in this review, it was shown that computational studies concern not only interactions between estrogens and biomacromolecules, but they also can be used to describe phenomena such as migration of estrogens through lipid bilayers or their adsorption on various materials. This can help predict the most efficient way to remove them from the environment when treated as pollutants. Further, it has been shown by multiple examples that quantum molecular modelling methods, such as those based on density functional theory, can be successfully used in structural studies on new solid forms of estrogens such as salts, co-crystals, hydrates, and polymorphs as well as on the complexes of estrogens with other molecules (i.e., cyclodextrins). In addition, the possibility to accurately calculate vibrational and NMR properties can be very helpful to explain spectroscopic results. Finally, we presented similar molecular modelling studies on xenoestrogens such as Bisphenol A and phthalates. Therefore, taking into consideration the versatility and confirmed accuracy of molecular modelling methods, it is not surprising that they have been listed in the specific guidance for studies on EDC published by international organizations such as ECA, EFSA, and OECD.

Supplementary Materials: The following are available online at http://www.mdpi.com/1422-0067/21/17/6411/s1.

Author Contributions: Conceptualization, A.H.M. and Ł.S.; investigation, A.H.M. and Ł.S.; writing—original draft preparation, A.H.M. and T.S.; writing—review and editing, A.H.M., Ł.S., T.S., and D.M.P.; visualization, A.H.M., Ł.S., and D.M.P.; supervision, A.H.M. and Ł.S.; project administration, Ł.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

17β-HSD	17β-hydroxysteroid dehydrogenase
ADMET	Adsorption distribution metabolism elimination toxicity
BPA	Bisphenol A
CoMFA	Comparative molecular field analysis
DBD	DNA binding domain
DFT	Density functional theory
E1	Estrone
E2	Estradiol
E3	Estriol
E4	Estretrol
ER	Estrogen receptor
EDCs	Endocrine-disrupting chemicals
FEP	Free energy perturbation
H-K	Hohenberg–Kohn theorems
K-S	Kohn–Sham theorems
LBD	Ligand-binding domain
LIE	Linear interaction energy
MD	Molecular dynamics
MM	Molecular mechanics
PDB	Protein Data Bank
QM	Quantum mechanics
QSAR	Quantitative structure–activity relationship
RBA	Relative binding affinity
SERMs	Selective estrogen receptor modulators
SHBG	Sex hormone-binding globulin
STS	Sulfatase
SULT	Sulfotransferase

References

1. Bennink, H.C.; Verhoeven, C.; Zimmerman, Y.; Visser, M.; Foidart, J.-M.; Gemzell-Danielsson, K. Clinical effects of the fetal estrogen estetrol in a multiple-rising-dose study in postmenopausal women. *Matur. Eur. Menopause J.* **2016**, *91*, 93–100. [CrossRef]
2. Singleton, D.W.; David, W.S. Xenoestrogen exposure and mechanisms of endocrine disruption. *Front. Biosci.* **2003**, *8*, 110–118. [CrossRef] [PubMed]
3. Bennink, F.C.; Holinka, C.F.; Visser, M.; Bennink, H.J.T.C. Maternal and fetal estetrol levels during pregnancy. *Climacteric* **2008**, *11*, 69–72. [CrossRef]
4. Schreiner, W.E. *The Ovary in Labhart, A. Clinical Endocrinology: Theory and Practice*; Springer: Berlin/Heidelberg, Germany, 2012; p. 548.
5. Kuhl, H. Pharmacology of estrogens and progestogens: Influence of different routes of administration. *Climacteric* **2005**, *8* (Suppl. 1), 3–63. [CrossRef]
6. Blackburn, S. *Maternal, Fetal, & Neonatal Physiology*; Elsevier: Amsterdam, The Netherlands, 2014.
7. Fait, T. Menopause hormone therapy: Latest developments and clinical practice. *Drugs Context* **2019**, *8*, 1–9. [CrossRef]
8. Tofovic, S.P.; Jackson, E.K. Estradiol Metabolism: Crossroads in Pulmonary Arterial Hypertension. *Int. J. Mol. Sci.* **2019**, *21*, 116. [CrossRef]
9. EC. Corrigendum to Commission Regulation (EU) 2018/605 of 19 April 2018 Amending Annex II to Regulation (EC) No 1107/2009 by Setting Out Scientific Criteria for the Determination of Endocrine Disrupting Properties. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R0605&rid=210 (accessed on 25 August 2020).

10. Ciubotaru, R.M.; Oyedele, J.; Zancanaro, G. Annual assessment of Echinococcus multilocularis surveillance reports submitted in 2018 in the context of Commission Regulation (EU) No 1152/2011. *EFSA J.* **2018**, *16*, 33–36. [CrossRef]
11. OECD. *Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption, No. 150, Update v3, Series on Testing and Assessment, ENV/JM/MONO(2012)22*; OECD: Paris, France, 2017; 988p.
12. OECD. Conceptual Framework for Testing and Assessment of Endocrine Disrupters. Available online: https://www.oecd.org/env/ehs/testing/oecdworkrelatedtoendocrinedisrupters.htm (accessed on 25 August 2020).
13. Safe, S.; Khan, S.; Wu, F.; Li, X. Chemical-induced estrogenicity in Xenoestrogens and phytoestrogens as SERMs and implications for risk assessment (Chapter 65). In *Veterinary Toxicology Basic and Clinical Principles*; Elsevier: Amsterdam, The Netherlands, 2007; pp. 811–822.
14. Cotterill, J.; Palazzolo, L.; Ridgway, C.; Price, N.; Rorije, E.; Moretto, A.; Peijnenburg, A.; Eberini, I. Predicting estrogen receptor binding of chemicals using a suite of in silico methods—Complementary approaches of (Q)SAR, molecular docking and molecular dynamics. *Toxicol. Appl. Pharmacol.* **2019**, *378*, 114630. [CrossRef]
15. Kerdivel, G.; Habauzit, D.; Pakdel, F. Assessment and Molecular Actions of Endocrine-Disrupting Chemicals That Interfere with Estrogen Receptor Pathways. *Int. J. Endocrinol.* **2013**, *2013*, 501851. [CrossRef]
16. Park, J.; Czapla, L.; Amaro, R.E. Molecular Simulations of Aromatase Reveal New Insights into the Mechanism of Ligand Binding. *J. Chem. Inf. Model.* **2013**, *53*, 2047–2056. [CrossRef]
17. Lokhande, K.; Mehere, A.S.; Yadav, A.K.; Swamy, K.V. Molecular Modeling and Docking Studies of Aromatase Inhibitors with Aromatase for ERP Breast Cancer. In Proceedings of the 86th Conference of Society of Biological Chemists: Emerging Discoveries in Health and Agricultural Sciences, New Delhi, India, 16–19 November 2017. [CrossRef]
18. Kang, H.; Xiao, X.; Huang, C.; Yuan, Y.; Tang, D.; Dai, X.; Zeng, X. Potent aromatase inhibitors and molecular mechanism of inhibitory action. *Eur. J. Med. Chem.* **2018**, *143*, 426–437. [CrossRef] [PubMed]
19. Awasthi, M.; Singh, S.; Pandey, V.P.; Dwivedi, U.N. Molecular docking and 3D-QSAR-based virtual screening of flavonoids as potential aromatase inhibitors against estrogen-dependent breast cancer. *J. Biomol. Struct. Dyn.* **2014**, *33*, 804–819. [CrossRef] [PubMed]
20. Suvannang, N.; Nantasenamat, C.; Isarankura-Na-Ayudhya, C.; Prachayasittikul, V. Molecular Docking of Aromatase Inhibitors. *Molecules* **2011**, *16*, 3597–3617. [CrossRef]
21. Rampogu, S.; Son, M.; Park, C.; Kim, H.-H.; Suh, J.-K.; Lee, K. Sulfonanilide Derivatives in Identifying Novel Aromatase Inhibitors by Applying Docking, Virtual Screening, and MD Simulations Studies. *BioMed Res. Int.* **2017**, *2017*, 1–17. [CrossRef] [PubMed]
22. Narayana, B.L.; Kishore, D.P.; Balakumar, C.; Rao, K.V.; Kaur, R.; Rao, A.R.; Murthy, J.N.; Ravikumar, M. Molecular Modeling Evaluation of Non-Steroidal Aromatase Inhibitors†. *Chem. Boil. Drug Des.* **2012**, *79*, 674–682. [CrossRef] [PubMed]
23. Fischer, J.; Ganellin, C.R. *Analogue-Based Drug Discovery*; John Wiley & Sons: Hoboken, NJ, USA, 2006; p. 516.
24. Maltais, R.; Ayan, D.; Trottier, A.; Barbeau, X.; Lagüe, P.; Bouchard, J.-E.; Poirier, D. Discovery of a Non-Estrogenic Irreversible Inhibitor of 17β-Hydroxysteroid Dehydrogenase Type 1 from 3-Substituted-16β-(m-carbamoylbenzyl)-estradiol Derivatives. *J. Med. Chem.* **2013**, *57*, 204–222. [CrossRef]
25. Lespérance, M.; Barbeau, X.; Roy, J.; Maltais, R.; Lagüe, P.; Poirier, D. Chemical synthesis of C3-oxiranyl/oxiranylmethyl-estrane derivatives targeted by molecular modeling and tested as potential inhibitors of 17β-hydroxysteroid dehydrogenase type 1. *Steroids* **2018**, *140*, 104–113. [CrossRef]
26. Maltais, R.; Djiemeny, A.N.; Roy, J.; Barbeau, X.; Lambert, J.-P.; Poirier, D. Design and synthesis of dansyl-labeled inhibitors of steroid sulfatase for optical imaging. *Bioorg. Med. Chem.* **2020**, *28*, 115368. [CrossRef]
27. Rakers, C.; Schumacher, F.; Meinl, W.; Glatt, H.; Kleuser, B.; Wolber, G. In SilicoPrediction of Human Sulfotransferase 1E1 Activity Guided by Pharmacophores from Molecular Dynamics Simulations. *J. Boil. Chem.* **2015**, *291*, 58–71. [CrossRef]
28. Kisselev, P.; Schunck, W.-H.; Roots, I.; Schwarz, D. Association of CYP1A1 Polymorphisms with Differential Metabolic Activation of 17 -Estradiol and Estrone. *Cancer Res.* **2005**, *65*, 2972–2978. [CrossRef]
29. Hong, C.-C.; Tang, B.-K.; Hammond, G.L.; Tritchler, D.; Yaffe, M.; Boyd, N.F. Cytochrome P450 1A2 (CYP1A2) activity and risk factors for breast cancer: A cross-sectional study. *Breast Cancer Res.* **2004**, *6*, R352–R365. [CrossRef] [PubMed]

30. Nishida, C.R.; Everett, S.; De Montellano, P.R.O. Specificity Determinants of CYP1B1 Estradiol Hydroxylation. *Mol. Pharmacol.* **2013**, *84*, 451–458. [CrossRef] [PubMed]
31. Selby, C. Sex Hormone Binding Globulin: Origin, Function and Clinical Significance. *Ann. Clin. Biochem. Int. J. Lab. Med.* **1990**, *27*, 532–541. [CrossRef]
32. Dechering, K.; Boersma, C.; Mosselman, S. Estrogen receptors alpha and beta: Two receptors of a kind? *Curr. Med. Chem.* **2000**, *7*, 561–576. [CrossRef] [PubMed]
33. Paterni, I.; Granchi, C.; Katzenellenbogen, J.A.; Minutolo, F. Estrogen receptors alpha (ERα) and beta (ERβ): Subtype-selective ligands and clinical potential. *Steroids* **2014**, *90*, 13–29. [CrossRef]
34. Ruff, M.; Gangloff, M.; Wurtz, J.; Moras, D. Estrogen receptor transcription and transactivation Structure-function relationship in DNA- and ligand-binding domains of estrogen receptors. *Breast Cancer Res.* **2000**, *2*, 353–359. [CrossRef]
35. Kumar, R.; Zakharov, M.N.; Khan, S.H.; Miki, R.; Jang, H.; Toraldo, G.; Singh, R.; Bhasin, S.; Jasuja, R. The Dynamic Structure of the Estrogen Receptor. *J. Amino Acids* **2011**, *2011*, 1–7. [CrossRef]
36. Billon-Galés, A.; Krust, A.; Fontaine, C.; Abot, A.; Flouriot, G.; Toutain, C.; Berges, H.; Gadeau, A.-P.; Lenfant, F.; Gourdy, P.; et al. Activation function 2 (AF2) of estrogen receptor-α is required for the atheroprotective action of estradiol but not to accelerate endothelial healing. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 13311–13316. [CrossRef]
37. Gao, L.; Tu, Y.; Ågren, H.; Eriksson, L.A. Characterization of Agonist Binding to His524 in the Estrogen Receptor α Ligand Binding Domain. *J. Phys. Chem. B* **2012**, *116*, 4823–4830. [CrossRef]
38. Hu, G.; Wang, J. Ligand selectivity of estrogen receptors by a molecular dynamics study. *Eur. J. Med. Chem.* **2014**, *74*, 726–735. [CrossRef]
39. Gu, X. Helix 12 in the human estrogen receptor (hER) is essential for the hER function by overcoming nucleosome repression in yeast. *J. Cell. Biochem.* **2002**, *86*, 224–238. [CrossRef] [PubMed]
40. Shiau, A.K.; Barstad, D.; Loria, P.M.; Cheng, L.; Kushner, P.J.; Agard, D.A.; Greene, G.L. The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen. *Cell* **1998**, *95*, 927–937. [CrossRef]
41. Fuentes, N.; Silveyra, P. Estrogen receptor signaling mechanisms. *Adv. Protein Chem. Struct. Biol.* **2019**, *116*, 135–170. [CrossRef] [PubMed]
42. Stefkovich, M.L.; Arao, Y.; Hamilton, K.J.; Korach, K.S. Experimental models for evaluating non-genomic estrogen signaling. *Steroids* **2018**, *133*, 34–37. [CrossRef] [PubMed]
43. Vrtacnik, P.; Ostanek, B.; Mencej-Bedrac, S.; Marc, J. The many faces of estrogen signaling. *Biochem. Medica* **2014**, *24*, 329–342. [CrossRef]
44. Rosano, C.; Ponassi, M.; Santolla, M.F.; Pisano, A.; Felli, L.; Vivacqua, A.; Maggiolini, M.; Lappano, R. Macromolecular Modelling and Docking Simulations for the Discovery of Selective GPER Ligands. *AAPS J.* **2015**, *18*, 41–46. [CrossRef]
45. Bruno, A.; Aiello, F.; Costantino, G.; Radi, M. Homology Modeling, Validation and Dynamics of the G Protein-coupled Estrogen Receptor 1 (GPER-1). *Mol. Inform.* **2016**, *35*, 333–339. [CrossRef]
46. Méndez-Luna, D.; Martínez-Archundia, M.; Maroun, R.C.; Ceballos, G.; Fragoso-Vázquez, M.; González-Juárez, D.; Correa-Basurto, J. Deciphering the GPER/GPR30-agonist and antagonists interactions using molecular modeling studies, molecular dynamics, and docking simulations. *J. Biomol. Struct. Dyn.* **2015**, *33*, 1–12. [CrossRef]
47. Taylor, R.; Jewsbury, P.; Essex, J.W. A review of protein-small molecule docking methods. *J. Comput. Mol. Des.* **2002**, *16*, 151–166. [CrossRef]
48. RCSB PDB. Available online: https://www.rcsb.org/ (accessed on 25 August 2020).
49. Landeros-Martinez, L.-L.; Glossman-Mitnik, D.; Orrantia-Borunda, E.; Flores-Holguin, N. A Combined Molecular Docking and Electronic Structure Study for a Breast Cancer Drug Design. In *Molecular Docking*; IntechOpen: London, UK, 2018. [CrossRef]
50. Muchtaridi, M.; Megantara, S.; Dermawan, D.; Yusuf, M. Antagonistic mechanism of α-mangostin derivatives against human estrogen receptor α of breast cancer using molecular dynamics simulation. *Rasayan J. Chem.* **2019**, *12*, 1927–1934. [CrossRef]
51. Cavasotto, C.; Aucar, M.G. High-Throughput Docking Using Quantum Mechanical Scoring. *Front. Chem.* **2020**, *8*, 246. [CrossRef] [PubMed]

52. Spiriti, J.; Subramanian, S.R.; Palli, R.; Wu, M.; Zuckerman, D.M. Middle-way flexible docking: Pose prediction using mixed-resolution Monte Carlo in estrogen receptor α. *PLoS ONE* **2019**, *14*, e0215694. [CrossRef] [PubMed]
53. Wolohan, P.; Reichert, D.E. CoMSIA and docking study of rhenium based estrogen receptor ligand analogs. *Steroids* **2007**, *72*, 247–260. [CrossRef] [PubMed]
54. Wierbowski, S.D.; Wingert, B.M.; Zheng, J.; Camacho, C.J. Cross-docking benchmark for automated pose and ranking prediction of ligand binding. *Protein Sci.* **2019**, *29*, 298–305. [CrossRef]
55. Shtaiwi, A.; Adnan, R.; Khairuddean, M.; Khan, S.U. Computational investigations of the binding mechanism of novel benzophenone imine inhibitors for the treatment of breast cancer. *RSC Adv.* **2019**, *9*, 35401–35416. [CrossRef]
56. Ling, J.; Oh, D.L.; Chia, A.Y.Y. Molecular Docking Studies of Glycyrrhizic Acid (GA), Glycyrrhetic Acid (GE) and Glabridin (GLA) with Estrogen Receptors (ERs). *Biosci. Biotechnol. Res. Asia* **2017**, *14*, 1–11. [CrossRef]
57. Pang, X.; Fu, W.; Wang, J.; Kang, D.; Xu, L.; Zhao, Y.; Liu, A.; Du, G.-H. Identification of Estrogen Receptor α Antagonists from Natural Products via In Vitro and In Silico Approaches. *Oxidative Med. Cell. Longev.* **2018**, *2018*, 6040149. [CrossRef]
58. Muchtaridi, M.; Dermawan, D.; Yusuf, M. Molecular Docking, 3D Structure-Based Pharmacophore Modeling, and ADME Prediction of Alpha Mangostin and Its Derivatives against Estrogen Receptor Alpha. *J. Young Pharm.* **2018**, *10*, 252–259. [CrossRef]
59. Yu, E.; Xu, Y.; Shi, Y.; Yu, Q.; Liu, J.; Xu, L. Discovery of novel natural compound inhibitors targeting estrogen receptor α by an integrated virtual screening strategy. *J. Mol. Model.* **2019**, *25*, 278. [CrossRef]
60. Abdelsamie, A.S.; Salah, M.; Siebenbürger, L.; Hamed, M.M.; Börger, C.; Van Koppen, C.J.; Frotscher, M.; Hartmann, R.W. Development of potential preclinical candidates with promising in vitro ADME profile for the inhibition of type 1 and type 2 17β-Hydroxysteroid dehydrogenases: Design, synthesis, and biological evaluation. *Eur. J. Med. Chem.* **2019**, *178*, 93–107. [CrossRef]
61. Hasan, T.N.; B, L.G.; A Masoodi, T.; Shafi, G.; Alshatwi, A.A.; Sivashanmugham, P. Affinity of estrogens for human progesterone receptor A and B monomers and risk of breast cancer: A comparative molecular modeling study. *Adv. Appl. Bioinform. Chem.* **2011**, *4*, 29–36. [CrossRef] [PubMed]
62. Fu, X.; Wang, P.; Zhu, B.T. Characterization of the Estradiol-Binding Site Structure of Human Protein Disulfide Isomerase (PDI). *PLoS ONE* **2011**, *6*, e27185. [CrossRef] [PubMed]
63. Da Silva, A.J.; Dos Santos, E.S. Aqueous solution interactions with sex hormone-binding globulin and estradiol: A theoretical investigation. *J. Boil. Phys.* **2018**, *44*, 539–556. [CrossRef] [PubMed]
64. Thomas, M.P.; Potter, B.V. The structural biology of oestrogen metabolism. *J. Steroid Biochem. Mol. Boil.* **2013**, *137*, 27–49. [CrossRef]
65. Maltais, R.; Fournier, D.; Poirier, D.; Maltais, R. Quantitative Structure-Activity Relationship (QSAR) Study with a Series of 17Î±-Derivatives of Estradiol: Model for the Development of Reversible Steroid Sulfatase Inhibitors. *QSAR Comb. Sci.* **2009**, *28*, 1284–1299. [CrossRef]
66. Granados, S.T.; Castillo, K.; Bravo-Moraga, F.; Sepúlveda, R.V.; Carrasquel-Ursulaez, W.; Rojas, M.; Carmona, E.; Lorenzo-Ceballos, Y.; González-Nilo, F.D.; González, C.; et al. The molecular nature of the 17β-Estradiol binding site in the voltage- and Ca2+-activated K+ (BK) channel β1 subunit. *Sci. Rep.* **2019**, *9*, 9965. [CrossRef]
67. McCullough, C.; Neumann, T.S.; Gone, J.R.; He, Z.; Herrild, C.; (Nee Lukesh) Wondergem, J.; Pandey, R.K.; Donaldson, W.A.; Sem, D.S. Probing the human estrogen receptor-α binding requirements for phenolic mono- and di-hydroxyl compounds: A combined synthesis, binding and docking study. *Bioorg. Med. Chem.* **2013**, *22*, 303–310. [CrossRef]
68. Kerdivel, G.; Le Guével, R.; Habauzit, D.; Brion, F.; Ait-Aissa, S.; Pakdel, F. Estrogenic Potency of Benzophenone UV Filters in Breast Cancer Cells: Proliferative and Transcriptional Activity Substantiated by Docking Analysis. *PLoS ONE* **2013**, *8*, e60567. [CrossRef]
69. Sauvée, C.; Schäfer, A.; Sundén, H.; Ma, J.-N.; Gustavsson, A.-L.; Burstein, E.S.; Olsson, R. The A-CD analogue of 16β,17α-estriol is a potent and highly selective estrogen receptor β agonist. *MedChemComm* **2013**, *4*, 1439. [CrossRef]
70. Kucinska, M.; Giron, M.D.; Piotrowska, H.; Lisiak, N.; Granig, W.H.; Lopez-Jaramillo, F.-J.; Salto, R.; Murias, M.; Erker, T. Novel Promising Estrogenic Receptor Modulators: Cytotoxic and Estrogenic Activity of Benzanilides and Dithiobenzanilides. *PLoS ONE* **2016**, *11*, e0145615. [CrossRef]

71. Gonzalez, T.L.; Rae, J.M.; Colacino, J.A.; Richardson, R.J. Homology models of mouse and rat estrogen receptor-α ligand-binding domain created by in silico mutagenesis of a human template: Molecular docking with 17β-estradiol, diethylstilbestrol, and paraben analogs. *Comput. Toxicol.* **2019**, *10*, 1–16. [CrossRef] [PubMed]
72. Yarger, J.G.; E Babine, R.; Bittner, M.; Shanle, E.; Xu, W.; Hershberger, P.; Nye, S.H. Structurally similar estradiol analogs uniquely alter the regulation of intracellular signaling pathways. *J. Mol. Endocrinol.* **2012**, *50*, 43–57. [CrossRef] [PubMed]
73. Maltais, R.; Trottier, A.; Barbeau, X.; Lagüe, P.; Perreault, M.; Thériault, J.-F.; Lin, S.X.; Poirier, D. Impact of structural modifications at positions 13, 16 and 17 of 16β-(m-carbamoylbenzyl)-estradiol on 17β-hydroxysteroid dehydrogenase type 1 inhibition and estrogenic activity. *J. Steroid Biochem. Mol. Boil.* **2016**, *161*, 24–35. [CrossRef] [PubMed]
74. E Starkey, N.J.; Li, Y.; Drenkhahn-Weinaug, S.K.; Liu, J.; Lubahn, D.B. 27-Hydroxycholesterol Is an Estrogen Receptor β-Selective Negative Allosteric Modifier of 17β-Estradiol Binding. *Endocrinology* **2018**, *159*, 1972–1981. [CrossRef] [PubMed]
75. Singh, K.; Munuganti, R.S.N.; Leblanc, E.; Lin, Y.L.; Leung, E.; Lallous, N.; Butler, M.S.; Cherkasov, A.; Rennie, P.S. In silico discovery and validation of potent small-molecule inhibitors targeting the activation function 2 site of human oestrogen receptor α. *Breast Cancer Res.* **2015**, *17*, 27. [CrossRef]
76. Durrant, J.D.; Carlson, K.E.; Martin, T.A.; Offutt, T.L.; Mayne, C.G.; Katzenellenbogen, J.A.; Amaro, R.E. Neural-Network Scoring Functions Identify Structurally Novel Estrogen-Receptor Ligands. *J. Chem. Inf. Model.* **2015**, *55*, 1953–1961. [CrossRef]
77. Ng, H.W.; Zhang, W.; Shu, M.; Luo, H.; Ge, W.; Perkins, R.; Tong, W.; Hong, H. Competitive molecular docking approach for predicting estrogen receptor subtype α agonists and antagonists. *BMC Bioinform.* **2014**, *15*, S4. [CrossRef]
78. Amr, A.E.-G.E.; Elsayed, E.; Al-Omar, M.; Eldin, H.O.B.; Nossier, E.S.; Abdalla, M.M. Design, Synthesis, Anticancer Evaluation and Molecular Modeling of Novel Estrogen Derivatives. *Molecules* **2019**, *24*, 416. [CrossRef]
79. Stjernschantz, E.; Reinen, J.; Meinl, W.; George, B.J.; Glatt, H.; Vermeulen, N.P.; Oostenbrink, C. Comparison of murine and human estrogen sulfotransferase inhibition in vitro and in silico—Implications for differences in activity, subunit dimerization and substrate inhibition. *Mol. Cell. Endocrinol.* **2010**, *317*, 127–140. [CrossRef]
80. Hanson, R.N.; McCaskill, E.; Tongcharoensirikul, P.; Dilis, R.; Labaree, D.; Hochberg, R.B. Synthesis and evaluation of 17α-(dimethylphenyl)vinyl estradiols as probes of the estrogen receptor-α ligand binding domain. *Steroids* **2012**, *77*, 471–476. [CrossRef]
81. Grande, F.; Rizzuti, B.; Occhiuzzi, M.A.; Ioele, G.; Casacchia, T.; Gelmini, F.; Guzzi, R.; Garofalo, A.; Statti, G. Identification by Molecular Docking ofHomoisoflavones from Leopoldia comosa as Ligands of Estrogen Receptors. *Molecules* **2018**, *23*, 894. [CrossRef] [PubMed]
82. Powers, C.N.; Setzer, W.N. A molecular docking study of phytochemical estrogen mimics from dietary herbal supplements. *Silico Pharmacol.* **2015**, *3*, 4. [CrossRef] [PubMed]
83. Yuan, P.; Liang, K.; Ma, B.; Zheng, N.; Nussinov, R.; Huang, J. Multiple-Targeting and Conformational Selection in the Estrogen Receptor: Computation and Experiment. *Chem. Boil. Drug Des.* **2011**, *78*, 137–149. [CrossRef] [PubMed]
84. Chen, Y.; Wang, J.; Hong, D.-Y.; Chen, L.; Zhang, Y.-Y.; Xu, Y.-N.; Pan, D.; Fu, L.-Y.; Tao, L.; Luo, H.; et al. Baicalein has protective effects on the 17β-estradiol-induced transformation of breast epithelial cells. *Oncotarget* **2017**, *8*, 10470–10484. [CrossRef] [PubMed]
85. Fokialakis, N.; Alexi, X.; Aligiannis, N.; Boulaka, A.; Meligova, A.K.; Lambrinidis, G.; Kalpoutzakis, E.; Pratsinis, H.; Cheilari, A.; Mitsiou, D.J.; et al. Biological evaluation of isoflavonoids from Genista halacsyi using estrogen-target cells: Activities of glucosides compared to aglycones. *PLoS ONE* **2019**, *14*, e0210247. [CrossRef]
86. Ayoub, N.M.; Siddique, A.B.; Ebrahim, H.Y.; Mohyeldin, M.M.; El Sayed, K.A. The olive oil phenolic (-)-oleocanthal modulates estrogen receptor expression in luminal breast cancer in vitro and in vivo and synergizes with tamoxifen treatment. *Eur. J. Pharmacol.* **2017**, *810*, 100–111. [CrossRef]
87. Yuseran, H.; Hartoyo, E.; Nurseta, T.; Kalim, H. Molecular docking of genistein on estrogen receptors, promoter region of BCLX, caspase-3, Ki-67, cyclin D1, and telomere activity. *J. Taibah Univ. Med. Sci.* **2018**, *14*, 79–87. [CrossRef]

88. Nanashima, N.; Horie, K.; Maeda, H. Phytoestrogenic Activity of Blackcurrant Anthocyanins Is Partially Mediated through Estrogen Receptor Beta. *Molecules* **2017**, *23*, 74. [CrossRef]
89. Alam, S.; Khan, F. Virtual screening, Docking, ADMET and System Pharmacology studies on Garcinia caged Xanthone derivatives for Anticancer activity. *Sci. Rep.* **2018**, *8*, 5524. [CrossRef]
90. Puranik, N.V.; Srivastava, P.; Bhatt, G.; Mary, D.J.S.J.; Limaye, A.M.; Sivaraman, J. Determination and analysis of agonist and antagonist potential of naturally occurring flavonoids for estrogen receptor (ERα) by various parameters and molecular modelling approach. *Sci. Rep.* **2019**, *9*, 7450. [CrossRef]
91. Wang, T.; Wang, Y.; Zhuang, X.; Luan, F.; Zhao, C.; Cordeiro, M.N.D.S. Interaction of Coumarin Phytoestrogens with ERα and ERβ: A Molecular Dynamics Simulation Study. *Molecules* **2020**, *25*, 1165. [CrossRef]
92. Wang, X.; Wang, G.-C.; Rong, J.; Wang, S.W.; Ng, T.B.; Zhang, Y.B.; Lee, K.F.; Zheng, L.; Wong, H.-K.; Yung, K.K.-L.; et al. Identification of Steroidogenic Components Derived From Gardenia jasminoides Ellis Potentially Useful for Treating Postmenopausal Syndrome. *Front. Pharmacol.* **2018**, *9*, 9. [CrossRef]
93. Jeong, J.; Kim, H.; Choi, J. In Silico Molecular Docking and In Vivo Validation with *Caenorhabditis elegans* to Discover Molecular Initiating Events in Adverse Outcome Pathway Framework: Case Study on Endocrine-Disrupting Chemicals with Estrogen and Androgen Receptors. *Int. J. Mol. Sci.* **2019**, *20*, 1209. [CrossRef]
94. Neves, B.J.; Braga, R.C.; Melo-Filho, C.C.; Moreira-Filho, J.T.; Muratov, E.N.; Andrade, C.H. QSAR-Based Virtual Screening: Advances and Applications in Drug Discovery. *Front. Pharmacol.* **2018**, *9*, 1275. [CrossRef]
95. Dror, O.; Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H.J. Novel Approach for Efficient Pharmacophore-Based Virtual Screening: Method and Applications. *J. Chem. Inf. Model.* **2009**, *49*, 2333–2343. [CrossRef]
96. Fiser, A. Template-based protein structure modeling. *Breast Cancer* **2010**, *673*, 73–94. [CrossRef]
97. Zhao, X.; Chen, M.; Huang, B.; Ji, H.; Yuan, M. Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA) Studies on α1A-Adrenergic Receptor Antagonists Based on Pharmacophore Molecular Alignment. *Int. J. Mol. Sci.* **2011**, *12*, 7022–7037. [CrossRef]
98. Sharma, R.; Dhingra, N.; Patil, S. CoMFA, CoMSIA, HQSAR and Molecular Docking Analysis of Ionone-based Chalcone Derivatives as Antiprostate Cancer Activity. *Indian J. Pharm. Sci.* **2016**, *78*, 54–64. [CrossRef]
99. Vuorinen, A.; Engeli, R.T.; Leugger, S.; Kreutz, C.R.; Schuster, D.; Odermatt, A.; Matuszczak, B. Phenylbenzenesulfonates and -sulfonamides as 17β-hydroxysteroid dehydrogenase type 2 inhibitors: Synthesis and SAR-analysis. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 2982–2985. [CrossRef]
100. Vuorinen, A.; Engeli, R.; Meyer, A.; Bachmann, F.; Griesser, U.J.; Schuster, D.; Odermatt, A. Ligand-Based Pharmacophore Modeling and Virtual Screening for the Discovery of Novel 17β-Hydroxysteroid Dehydrogenase 2 Inhibitors. *J. Med. Chem.* **2014**, *57*, 5995–6007. [CrossRef]
101. Chang, Y.-H.; Chen, J.-Y.; Hor, C.-Y.; Chuang, Y.-C.; Yang, C.-B.; Yang, C.-N. Computational Study of Estrogen Receptor-Alpha Antagonist with Three-Dimensional Quantitative Structure-Activity Relationship, Support Vector Regression, and Linear Regression Methods. *Int. J. Med. Chem.* **2013**, *2013*, 1–13. [CrossRef]
102. Sodero, A.C.R.; Romeiro, N.C.; Da Cunha, E.F.F.; Magalhães, U.D.O.; De Alencastro, R.B.; Rodrigues, C.R.; Cabral, L.M.; Castro, H.C.; Albuquerque, M.G. Application of 4D-QSAR Studies to a Series of Raloxifene Analogs and Design of Potential Selective Estrogen Receptor Modulators. *Molecules* **2012**, *17*, 7415–7439. [CrossRef]
103. Wang, P.; McInnes, C.; Zhu, B.T. Structural Characterization of the Binding Interactions of Various Endogenous Estrogen Metabolites with Human Estrogen Receptor α and β Subtypes: A Molecular Modeling Study. *PLoS ONE* **2013**, *8*, e74615. [CrossRef]
104. Bhhatarai, B.; Wilson, D.M.; Price, P.S.; Marty, S.; Parks, A.K.; Carney, E. Evaluation of OASIS QSAR Models Using ToxCast™ in Vitro Estrogen and Androgen Receptor Binding Data and Application in an Integrated Endocrine Screening Approach. *Environ. Health Perspect.* **2016**, *124*, 1453–1461. [CrossRef]
105. Bohari, M.; Srivastava, H.K.; Sastry, G.N. Analogue-based approaches in anti-cancer compound modelling: The relevance of QSAR models. *Org. Med. Chem. Lett.* **2011**, *1*, 3. [CrossRef]
106. Dems, M.A.E.; Laib, S.; Latelli, N.; Ouddai, N. A DFT-based Quantitative structure activity relationship Study of organometallic estradiol derivatives. *JCPS* **2017**, *10*, 483–487.
107. Zhang, T.; Wei, D.-Q.; Chou, K.-C. A pharmacophore model specific to active site of CYP1A2 with a novel molecular modeling explorer and CoMFA. *Med. Chem.* **2012**, *8*, 198–207.

108. Poirier, D.; Roy, J.; Cortes-Benitez, F.; Dutour, R. Targeting cytochrome P450 (CYP) 1B1 with steroid derivatives. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 5272–5276. [CrossRef]
109. Kar, S.; Roy, K.; Leszczynski, J. Impact of Pharmaceuticals on the Environment: Risk Assessment Using QSAR Modeling Approach. *Methods Mol. Biol.* **2018**, *1800*, 395–443. [CrossRef]
110. Colosi, L.M.; Huang, Q.; Weber, W.J. QSAR-assisted design of an environmental catalyst for enhanced estrogen remediation. *Chemosphere* **2010**, *81*, 897–903. [CrossRef]
111. Rokhina, E.V.; Suri, R. Application of density functional theory (DFT) to study the properties and degradation of natural estrogen hormones with chemical oxidizers. *Sci. Total Environ.* **2012**, *417*, 280–290. [CrossRef] [PubMed]
112. Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Boil.* **1976**, *103*, 227–249. [CrossRef]
113. Jorgensen, W. The Many Roles of Computation in Drug Discovery. *Science* **2004**, *303*, 1813–1818. [CrossRef] [PubMed]
114. Simonson, T.; Archontis, G.; Karplus, M. Free Energy Simulations Come of Age: Protein−Ligand Recognition. *Acc. Chem. Res.* **2002**, *35*, 430–437. [CrossRef]
115. Fratev, F.; Steinbrecher, T.; Jónsdóttir, S.O. Prediction of Accurate Binding Modes Using Combination of Classical and Accelerated Molecular Dynamics and Free-Energy Perturbation Calculations: An Application to Toxicity Studies. *ACS Omega* **2018**, *3*, 4357–4371. [CrossRef]
116. Schindler, C.; Rippmann, F.; Kuhn, D. Relative binding affinity prediction of farnesoid X receptor in the D3R Grand Challenge 2 using FEP+. *J. Comput. Mol. Des.* **2017**, *32*, 265–272. [CrossRef]
117. Olsson, M.A.; García-Sosa, A.T.; Ryde, U. Binding affinities of the farnesoid X receptor in the D3R Grand Challenge 2 estimated by free-energy perturbation and docking. *J. Comput. Mol. Des.* **2017**, *32*, 211–224. [CrossRef]
118. Åqvist, J.; Luzhkov, V.B.; Brandsdal, B.O. Ligand Binding Affinities from MD Simulations. *Acc. Chem. Res.* **2002**, *35*, 358–365. [CrossRef]
119. Van Lipzig, M.M.H.; Ter Laak, A.M.; Jongejan, A.; Vermeulen, N.P.; Wamelink, M.; Geerke, D.P.; Meerman, J.H.N. Prediction of Ligand Binding Affinity and Orientation of Xenoestrogens to the Estrogen Receptor by Molecular Dynamics Simulations and the Linear Interaction Energy Method. *J. Med. Chem.* **2004**, *47*, 1018–1030. [CrossRef]
120. Rifai, E.A.; Van Dijk, M.; Vermeulen, N.P.; Geerke, D.P. Binding free energy predictions of FXR agonists using LIE with reliability estimation: Application to the D3R Grand Challenge 2. *J. Comput. Mol. Des.* **2017**, *32*, 239–249. [CrossRef]
121. Simonson, T. Protein: Ligand recognition: Simple models for electrostatic effects. *Curr. Pharm. Des.* **2013**, *19*, 4241–4256. [CrossRef] [PubMed]
122. Gelpí, J.L.; Hospital, A.; Goñi, J.R.; Orozco, M. Molecular dynamics simulations: Advances and applications. *Adv. Appl. Bioinform. Chem.* **2015**, *8*, 37–47. [CrossRef] [PubMed]
123. Freindorf, M.; Furlani, T.R.; Kong, J.; Cody, V.; Davis, F.B.; Davis, P.J. Combined QM/MM Study of Thyroid and Steroid Hormone Analogue Interactions with αvβ3 Integrin. *J. Biomed. Biotechnol.* **2012**, *2012*, 1–12. [CrossRef] [PubMed]
124. Kalaiarasi, C.; Manjula, S.; Kumaradhas, P. Combined quantum mechanics/molecular mechanics (QM/MM) methods to understand the charge density distribution of estrogens in the active site of estrogen receptors. *RSC Adv.* **2019**, *9*, 40758–40771. [CrossRef]
125. Costa, A.H.L.; Clemente, W.S.; Bezerra, K.S.; Neto, J.X.L.; De Albuquerque, É.L.; Fulco, U.; Clemente, W.S., Jr. Computational biochemical investigation of the binding energy interactions between an estrogen receptor and its agonists. *New J. Chem.* **2018**, *42*, 19801–19810. [CrossRef]
126. Hilder, T.A.; Hodgkiss, J.M. Molecular Mechanism of Binding between 17β-Estradiol and DNA. *Comput. Struct. Biotechnol. J.* **2016**, *15*, 91–97. [CrossRef] [PubMed]
127. Eisold, A.; LaBudde, D. Detailed Analysis of 17β-Estradiol-Aptamer Interactions: A Molecular Dynamics Simulation Study. *Molecules* **2018**, *23*, 1690. [CrossRef] [PubMed]
128. Lakhin, A.V.; Tarantul, V.Z.; Gening, L.V. Aptamers: Problems, Solutions and Prospects. *Acta Naturae* **2013**, *5*, 34–43. [CrossRef] [PubMed]
129. Dutkiewicz, Z.; Mikstacka, R. Structure-Based Drug Design for Cytochrome P450 Family 1 Inhibitors. *Bioinorg. Chem. Appl.* **2018**, *2018*, 3924608. [CrossRef] [PubMed]

130. Moldogazieva, N.T.; Ostroverkhova, D.S.; Kuzmich, N.N.; Kadochnikov, V.V.; Terentiev, A.A.; Porozov, Y. Elucidating Binding Sites and Affinities of ERα Agonists and Antagonists to Human Alpha-Fetoprotein by In Silico Modeling and Point Mutagenesis. *Int. J. Mol. Sci.* **2020**, *21*, 893. [CrossRef] [PubMed]
131. Ma, K.; Saha, P.K.; Chan, L.; Moore, D.D. Farnesoid X receptor is essential for normal glucose homeostasis. *J. Clin. Investig.* **2006**, *116*, 1102–1109. [CrossRef] [PubMed]
132. Yang, F.; Huang, X.; Yi, T.; Yen, Y.; Moore, D.D.; Huang, W. Spontaneous Development of Liver Tumors in the Absence of the Bile Acid Receptor Farnesoid X Receptor. *Cancer Res.* **2007**, *67*, 863–867. [CrossRef] [PubMed]
133. Levin, E.R. Cellular Functions of the Plasma Membrane Estrogen Receptor. *Trends Endocrinol. Metab.* **1999**, *10*, 374–377. [CrossRef]
134. Soltysik, K.; Czekaj, P. Membrane estrogen receptors - is it an alternative way of estrogen action? *J. Physiol. Pharmacol. Off. J. Pol. Physiol. Soc.* **2013**, *64*, 129–142.
135. Guo, J.J.; Yang, D.-P.; Tian, X.; Vemuri, V.K.; Yin, D.; Li, C.; Duclos, R.I.; Shen, L.; Ma, X.; Janero, D.R.; et al. 17β-estradiol (E2) in membranes: Orientation and dynamic properties. *Biochim. Biophys. Acta (BBA) Biomembr.* **2016**, *1858*, 344–353. [CrossRef]
136. Vogel, A.; Scheidt, H.A.; Feller, S.E.; Metso, J.; Badeau, R.M.; Tikkanen, M.J.; Wähälä, K.; Jauhiainen, M.; Huster, D. The Orientation and Dynamics of Estradiol and Estradiol Oleate in Lipid Membranes and HDL Disc Models. *Biophys. J.* **2014**, *107*, 114–125. [CrossRef]
137. Goh, J.Y.; Goh, K.S.; Yip, Y.M.; Ng, C.K. High salinity enhances adsorption of 17α-ethinyl estradiol by polyethersulfone membrane: Isotherm modelling and molecular simulation. *engrXiv Preprints* **2019**. [CrossRef]
138. Lakshmanan, S.; Kanwal, A.; Liu, S.; Patlolla, A.; Iqbal, Z.; Mitra, S.; Thomas, G.A.; Fagan, J.A.; Farrow, R.C. Improved Electrophoretic Deposition of Vertical Single Wall Carbon Nanotubes with Nanoscopic Electrostatic Lenses. *Micromachines* **2020**, *11*, 324. [CrossRef]
139. Ulissi, Z.W.; Zhang, J.; Sresht, V.; Blankschtein, D.; Strano, M.S. 2D Equation-of-State Model for Corona Phase Molecular Recognition on Single-Walled Carbon Nanotube and Graphene Surfaces. *Langmuir* **2014**, *31*, 628–636. [CrossRef]
140. Sun, W.; Li, M.; Zhang, W.; Wei, J.; Chen, B.; Wang, C. Sediments inhibit adsorption of 17β-estradiol and 17α-ethinylestradiol to carbon nanotubes and graphene oxide. *Environ. Sci. Nano* **2017**, *4*, 1900–1910. [CrossRef]
141. Jiang, L.; Liu, Y.; Liu, S.; Zeng, G.; Hu, X.-J.; Hu, X.; Guo, Z.; Tan, X.; Wang, L.; Wu, Z. Adsorption of Estrogen Contaminants by Graphene Nanomaterials under Natural Organic Matter Preloading: Comparison to Carbon Nanotube, Biochar, and Activated Carbon. *Environ. Sci. Technol.* **2017**, *51*, 6352–6359. [CrossRef] [PubMed]
142. Boateng, L.K.; Heo, J.; Flora, J.R.V.; Park, Y.-G.; Yoon, Y. Molecular level simulation of the adsorption of bisphenol A and 17α-ethinyl estradiol onto carbon nanomaterials. *Sep. Purif. Technol.* **2013**, *116*, 471–478. [CrossRef]
143. Zaib, Q.; Khan, I.A.; Saleh, N.B.; Flora, J.R.V.; Park, Y.-G.; Yoon, Y. Removal of Bisphenol A and 17β-Estradiol by Single-Walled Carbon Nanotubes in Aqueous Solution: Adsorption and Molecular Modeling. *Water Air Soil Pollut.* **2012**, *223*, 3281–3293. [CrossRef]
144. Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864. [CrossRef]
145. Kohn, W.; Sham, L.J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138. [CrossRef]
146. Grimme, S. Density functional theory with London dispersion corrections. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 211–228. [CrossRef]
147. Clark, S.J.; Segall, M.D.; Pickard, C.J.; Hasnip, P.J.; Probert, M.I.J.; Refson, K.; Payne, M.C. First principles methods using CASTEP. *Z. Krist. Cryst. Mater.* **2005**, *220*, 567–570. [CrossRef]
148. Amusia, M.Y.; Msezane, A.Z.; Shaginyan, V.R. Density Functional Theory versus the Hartree–Fock Method: Comparative Assessment. *Phys. Scr.* **2003**, *68*, C133–C140. [CrossRef]
149. Mota, K.; Neto, J.X.L.; Costa, A.L.; Oliveira, J.; Bezerra, K.; De Albuquerque, É.L.; Caetano, E.W.S.; Freire, V.; Fulco, U. A quantum biochemistry model of the interaction between the estrogen receptor and the two antagonists used in breast cancer treatment. *Comput. Theor. Chem.* **2016**, *1089*, 21–27. [CrossRef]

150. Mazurek, A.H.; Szeleszczuk, Ł.; Pisklak, D.M. Periodic DFT Calculations—Review of Applications in the Pharmaceutical Sciences. *Pharmaceutics* **2020**, *12*, 415. [CrossRef]
151. Stevenson, E.L.; Lancaster, R.W.; Buanz, A.B.M.; Price, L.S.; Tocher, D.A.; Price, S.L. The solid state forms of the sex hormone 17-β-estradiol. *CrystEngComm* **2019**. [CrossRef]
152. Du, R.; Xu, J.; Zhang, L.; Ning, L.; Li, S. Ethinyl estradiol cocrystals assembled by chain structures: Improvement in stability and solubility. *New J. Chem.* **2019**, *43*, 16889–16897. [CrossRef]
153. Wang, J.-R.; Yang, Y.; Chen, X.; Mei, X. Solid-state characterization of 17β-estradiol co-crystals presenting improved dissolution and bioavailability. *CrystEngComm* **2016**, *18*, 3498–3505. [CrossRef]
154. Watanabe, H.; Okiyama, Y.; Nakano, T.; Tanaka, S. Incorporation of solvation effects into the fragment molecular orbital calculations with the Poisson–Boltzmann equation. *Chem. Phys. Lett.* **2010**, *500*, 116–119. [CrossRef]
155. Szeleszczuk, Ł.; Pisklak, D.M.; Zielińska-Pisklak, M.; Wawer, I. Effects of structural differences on the NMR chemical shifts in cinnamic acid derivatives: Comparison of GIAO and GIPAW calculations. *Chem. Phys. Lett.* **2016**, *653*, 35–41. [CrossRef]
156. Charpentier, T. The PAW/GIPAW approach for computing NMR parameters: A new dimension added to NMR study of solids. *Solid State Nucl. Magn. Reson.* **2011**, *40*, 1–20. [CrossRef] [PubMed]
157. Elyashberg, M.E. Identification and structure elucidation by NMR spectroscopy. *TrAC Trends Anal. Chem.* **2015**, *69*, 88–97. [CrossRef]
158. Szeleszczuk, Ł.; Pisklak, D.M.; Zielińska-Pisklak, M.; Jurczak, E. A new polymorph of 17-β-estradiol and the application of different analytical techniques (ssNMR, PXRD, DSC, and FTIR) for its study. *J. Mol. Struct.* **2019**, *1183*, 274–280. [CrossRef]
159. Singh, H.; Singh, S.; Srivastava, A.; Tandon, P.; Bharti, P.; Kumar, S.; Maurya, R. Conformational analysis and vibrational study of daidzein by using FT-IR and FT-Raman spectroscopies and DFT calculations. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2014**, *120*, 405–415. [CrossRef] [PubMed]
160. Machado, N.; De Carvalho, L.A.E.B.; Otero, J.C.; Marques, M.P.M. A conformational study of hydroxyflavones by vibrational spectroscopy coupled to DFT calculations. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2013**, *109*, 116–124. [CrossRef]
161. A Minaeva, V.; Minaev, B.F.; Hovorun, D.M. Vibrational spectra of the steroid hormones, estradiol and estriol, calculated by density functional theory. The role of low-frequency vibrations. *Ukr. Biokhimichnyi Zhurnal* **2009**, *80*, 82–95.
162. Scheidt, H.A.; Badeau, R.M.; Huster, D. Investigating the membrane orientation and transversal distribution of 17β-estradiol in lipid membranes by solid-state NMR. *Chem. Phys. Lipids* **2010**, *163*, 356–361. [CrossRef] [PubMed]
163. Vedad, J.; Mojica, E.-R.E.; Desamero, R.Z. Raman spectroscopic discrimination of estrogens. *Vib. Spectrosc.* **2018**, *96*, 93–100. [CrossRef] [PubMed]
164. Oren, I.; Fleishman, S.J.; Kessel, A.; Tal, N.B. Free Diffusion of Steroid Hormones Across Biomembranes: A Simplex Search with Implicit Solvent Model Calculations. *Biophys. J.* **2004**, *87*, 768–779. [CrossRef] [PubMed]
165. Ellena, J.; De Paula, K.; De Melo, C.C.; Da Silva, C.C.P.; Bezerra, B.P.; Venâncio, T.; Ayala, A.P. Temperature-Driven Isosymmetric Reversible Phase Transition of the Hormone Estradiol 17β Valerate. *Cryst. Growth Des.* **2014**, *14*, 5700–5709. [CrossRef]
166. Morishima, F.; Inokuchi, Y.; Ebata, T. Laser Spectroscopic Study of β-Estradiol and Its Monohydrated Clusters in a Supersonic Jet. *J. Phys. Chem. A* **2012**, *116*, 8201–8208. [CrossRef]
167. Borah, M.M.; Devi, T.G. The vibrational spectroscopic studies and molecular property analysis of Estradiol, Tamoxifen and their interaction by density functional theory. *J. Mol. Struct.* **2018**, *1163*, 205–220. [CrossRef]
168. Cherkasova, O.; Nazarov, M.; Mankova, A.; Fedulova, E.; Volodin, V.; Minaeva, V.A.; Minaev, B.F.; Baryshnikov, G.V. Terahertz time-domain spectroscopy of testosterone, estradiol and estriol. In Proceedings of the 2010 International Kharkov Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves, Kharkiv, Ukraine, 21–26 June 2010; pp. 1–3.
169. Hafizi, R.; Taheri, R.A.; Moghimi, H. Liquid phase extraction of nanosized biologically active estrogenic pollutants by using an efficient adsorbent. *J. Mol. Liq.* **2018**, *266*, 535–539. [CrossRef]
170. Donini, C.A.; Silva, M.K.L.; Simões, R.; Cesarino, I. Reduced graphene oxide modified with silver nanoparticles for the electrochemical detection of estriol. *J. Electroanal. Chem.* **2018**, *809*, 67–73. [CrossRef]

171. Klamt, A.; Schüürmann, G. COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans. 2* **1993**, *2*, 799–805. [CrossRef]
172. Mennucci, B.; Tomasi, J.; Cammi, R.; Cheeseman, J.R.; Frisch, M.J.; Devlin, F.J.; Gabriel, S.; Stephens, P.J. Polarizable Continuum Model (PCM) Calculations of Solvent Effects on Optical Rotations of Chiral Molecules. *J. Phys. Chem. A* **2002**, *106*, 6102–6113. [CrossRef]
173. Chinnasamy, K.; Kumaradhas, P. Intermolecular interactions and charge density distribution of endocrine-disrupting molecules (xenoestrogens) with ERα: QM/MM perspective. *Struct. Chem.* **2020**, *31*, 1013–1028. [CrossRef]
174. Ruiz, P.; Ingale, K.; Wheeler, J.S.; Mumtaz, M. 3D QSAR studies of hydroxylated polychlorinated biphenyls as potential xenoestrogens. *Chemosphere* **2016**, *144*, 2238–2246. [CrossRef] [PubMed]
175. Kim, M.; Li, L.Y.; Grace, J.R. Predictability of physicochemical properties of polychlorinated dibenzo-p-dioxins (PCDDs) based on single-molecular descriptor models. *Environ. Pollut.* **2016**, *213*, 99–111. [CrossRef]
176. Eguchi, A.; Hanazato, M.; Suzuki, N.; Matsuno, Y.; Todaka, E.; Mori, C. Maternal–fetal transfer rates of PCBs, OCPs, PBDEs, and dioxin-like compounds predicted through quantitative structure–activity relationship modeling. *Environ. Sci. Pollut. Res.* **2015**, *25*, 7212–7222. [CrossRef]
177. Delfosse, V.; Grimaldi, M.; Cavaillès, V.; Balaguer, P.; Bourguet, W. Structural and Functional Profiling of Environmental Ligands for Estrogen Receptors. *Environ. Health Perspect.* **2014**, *122*, 1306–1313. [CrossRef]
178. Nwachukwu, J.C.; Srinivasan, S.; Bruno, N.E.; Nowak, J.; Wright, N.J.; Minutolo, F.; Rangarajan, E.S.; Izard, T.; Yao, X.-Q.; Grant, B.J.; et al. Systems Structural Biology Analysis of Ligand Effects on ERα Predicts Cellular Response to Environmental Estrogens and Anti-hormone Therapies. *Cell Chem. Boil.* **2017**, *24*, 35–45. [CrossRef]
179. Cozzini, P.; Dellafiora, L. In silico approach to evaluate molecular interaction between mycotoxins and the estrogen receptors ligand binding domain: A case study on zearalenone and its metabolites. *Toxicol. Lett.* **2012**, *214*, 81–85. [CrossRef]
180. Zhang, J.; Wu, W.; Song, Y.; Hou, L.; Li, T.; Guan, T.; Zhang, T.; Wang, Y. Homogeneous assay for zearalenone analogues and their docking studies with apo-/holo-estrogen receptors. *Anal. Methods* **2019**, *11*, 192–199. [CrossRef]
181. Dellafiora, L.; Oswald, I.P.; Dorne, J.-L.; Galaverna, G.; Battilani, P.; Dall'Asta, C. An in silico structural approach to characterize human and rainbow trout estrogenicity of mycotoxins: Proof of concept study using zearalenone and alternariol. *Food Chem.* **2020**, *312*, 126088. [CrossRef]
182. Yang, J.; Hu, C.T.; Zhu, X.; Zhu, Q.; Ward, M.D.; Kahr, B. DDT Polymorphism and the Lethality of Crystal Forms. *Angew. Chem.* **2017**, *129*, 10299–10303. [CrossRef]
183. Zhang, H.; He, W.; Luo, X.; Lin, X.; Lu, X. Adsorption of 2,3,7,8-tetrochlorodibenzo-p-dioxins on intrinsic, defected, and Ti (N, Ag) doped graphene: A DFT study. *J. Mol. Model.* **2014**, *20*, 1–7. [CrossRef] [PubMed]
184. Liu, C.; Li, H.; Johnston, C.T.; Boyd, S.A.; Teppen, B.J. Relating Clay Structural Factors to Dioxin Adsorption by Smectites: Molecular Dynamics Simulations. *Soil Sci. Soc. Am. J.* **2012**, *76*, 110–120. [CrossRef]
185. Zhang, J.; Wu, W.; Wang, Y.; Xing, X.; Zhong, S.; Guan, T.; Zhang, T.; Hou, L.; Li, T. Estrogen receptor-based fluorescence polarization assay for bisphenol analogues and molecular modeling study of their complexation mechanism. *Anal. Chim. Acta* **2018**, *1032*, 107–113. [CrossRef] [PubMed]
186. Sengupta, S.; Obiorah, I.; Maximov, P.; Curpan, R.; Jordan, V.C. Molecular mechanism of action of bisphenol and bisphenol A mediated by oestrogen receptor alpha in growth and apoptosis of breast cancer cells. *Br. J. Pharmacol.* **2013**, *169*, 167–178. [CrossRef] [PubMed]
187. Liu, Y.; Qu, K.; Hai, Y.; Zhao, C. Bisphenol A (BPA) binding on full-length architectures of estrogen receptor. *J. Cell. Biochem.* **2018**, *119*, 6784–6794. [CrossRef] [PubMed]
188. Zhang, J.; Li, T.; Wang, T.; Yuan, C.; Zhong, S.; Guan, T.; Li, Z.; Wang, Y.; Yu, H.; Luo, Q.; et al. Estrogenicity of halogenated bisphenol A: In vitro and in silico investigations. *Arch. Toxicol.* **2017**, *92*, 1215–1223. [CrossRef] [PubMed]
189. Cao, H.; Wang, F.; Liang, Y.; Wang, H.; Zhang, A.; Song, M. Experimental and computational insights on the recognition mechanism between the estrogen receptor α with bisphenol compounds. *Arch. Toxicol.* **2017**, *91*, 3897–3912. [CrossRef]
190. Zhuang, S.; Zhang, C.; Liu, W. Atomic Insights into Distinct Hormonal Activities of Bisphenol A Analogues toward PPARγ and ERα Receptors. *Chem. Res. Toxicol.* **2014**, *27*, 1769–1779. [CrossRef]

191. Li, L.; Wang, Q.; Zhang, Y.; Niu, Y.; Yao, X.; Liu, H. The Molecular Mechanism of Bisphenol A (BPA) as an Endocrine Disruptor by Interacting with Nuclear Receptors: Insights from Molecular Dynamics (MD) Simulations. *PLoS ONE* **2015**, *10*, e0120330. [CrossRef]
192. Delfosse, V.; Grimaldi, M.; Pons, J.-L.; Boulahtouf, A.; Le Maire, A.; Cavaillès, V.; Labesse, G.; Bourguet, W.; Balaguer, P. Structural and mechanistic insights into bisphenols action provide guidelines for risk assessment and discovery of bisphenol A substitutes. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 14930–14935. [CrossRef] [PubMed]
193. Wei, D.; Li, J.; Chen, Z.; Liang, L.; Ma, J.; Wei, M.; Ai, Y.; Wang, X. Understanding bisphenol-A adsorption in magnetic modified covalent organic frameworks: Experiments coupled with DFT calculations. *J. Mol. Liq.* **2020**, *301*, 112431. [CrossRef]
194. Bao, S.; Wu, S.; Huang, L.; Xu, X.; Xu, R.; Li, Y.; Liang, Y.; Yang, M.; Yoon, D.K.; Lee, M.; et al. Supramolecular Nanopumps with Chiral Recognition for Moving Organic Pollutants from Water. *ACS Appl. Mater. Interfaces* **2019**, *11*, 31220–31226. [CrossRef] [PubMed]
195. Díaz, I.; Díez, E.; Camacho, J.; León, S.; Ovejero, G.; Cabanillas, S.L. Comparison between three predictive methods for the calculation of polymer solubility parameters. *Fluid Phase Equilibria* **2013**, *337*, 6–10. [CrossRef]
196. Wang, C.-Y.; Zhang, Y.-J.; Wang, W.-K.; Pei, D.-N.; Huang, G.-X.; Chen, J.-J.; Zhang, X.; Yu, H.-Q. Enhanced photocatalytic degradation of bisphenol A by Co-doped BiOCl nanosheets under visible light irradiation. *Appl. Catal. B Environ.* **2018**, *221*, 320–328. [CrossRef]
197. Motta, A.; La Mantia, F.P.; Ascione, L.; Mistretta, M. Theoretical study on the decomposition mechanism of bisphenol A polycarbonate induced by the combined effect of humidity and UV irradiation. *J. Mol. Graph. Model.* **2020**, *99*, 107622. [CrossRef]
198. Liu, Y.; Liu, Y.; Liu, Z.; Du, F.; Qin, G.; Li, G.; Hu, X.; Xu, Z.; Cai, Z. Supramolecularly imprinted polymeric solid phase microextraction coatings for synergetic recognition nitrophenols and bisphenol A. *J. Hazard. Mater.* **2019**, *368*, 358–364. [CrossRef]
199. Dvorakova, M.; Kejlová, K.; Rucki, M.; Jírová, D. Selected bisphenols and phthalates screened for estrogen and androgen disruption by in silico and in vitro methods. *Neuro Endocrinol. Lett.* **2018**, *39*, 409–416.
200. Zhu, Q.; Liu, L.; Zhou, X.; Ma, M. In silico study of molecular mechanisms of action: Estrogenic disruptors among phthalate esters. *Environ. Pollut.* **2019**, *255*, 113193. [CrossRef]
201. Josh, M.S.; Pradeep, S.; Adarsh, V.; Amma, K.V.; Devi, R.S.; Balachandran, S.; Sreejith, M.; Jaleel, U.A.; Benjamin, S. In silicoevidences for the binding of phthalates onto human estrogen receptor α, β subtypes and human estrogen-related receptor γ. *Mol. Simul.* **2013**, *40*, 408–417. [CrossRef]
202. Sheikh, I.A.; Turki, R.F.; Abuzenadah, A.M.; Damanhouri, G.A.; A Beg, M. Endocrine Disruption: Computational Perspectives on Human Sex Hormone-Binding Globulin and Phthalate Plasticizers. *PLoS ONE* **2016**, *11*, e0151444. [CrossRef] [PubMed]
203. Wang, S.; Wang, S.; Chen, M.; Xu, D.; Tang, L.; Wang, S. Elucidating Adsorption Mechanisms of Phthalate Esters upon Carbon Nanotubes/Graphene and Natural Organic Acid Competitive Effects in Water by DFT and MD Calculations. *Bull. Korean Chem. Soc.* **2015**, *36*, 1631–1636. [CrossRef]
204. Liu, Y.; Zhang, R.; Wang, X.; Sun, P.; Chen, W.; Shen, J.; Xue, G. Hydrogenation induced deviation of temperature and concentration dependences of polymer-solvent interactions in poly(vinyl chloride) and a new eco-friendly plasticizer. *Eur. Phys. J. Plus* **2015**, *130*, 11. [CrossRef]
205. Jacob, R.B.; Andersen, T.; McDougal, O.M. Accessible High-Throughput Virtual Screening Molecular Docking Software for Students and Educators. *PLoS Comput. Boil.* **2012**, *8*, e1002499. [CrossRef]
206. Chaput, L.; Mouawad, L. Efficient conformational sampling and weak scoring in docking programs? Strategy of the wisdom of crowds. *J. Chem.* **2017**, *9*, 37. [CrossRef]
207. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **2015**, *10*, 449–461. [CrossRef]
208. Software Website. GOLD—Protein Ligand Docking Software. Available online: www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/ (accessed on 25 August 2020).
209. Clark, M.; Cramer, R.D.; Van Opdenbosch, N. Validation of the general purpose tripos 5.2 force field. *J. Comput. Chem.* **1989**, *10*, 982–1012. [CrossRef]
210. Software Website. Available online: https://www.charmm.org/ (accessed on 25 August 2020).

211. Brooks, B.R.; Brooks, C.L.; Mackerell, A.D., Jr.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545–1614. [CrossRef]
212. Software Website. Available online: http://www.swissdock.ch/ (accessed on 25 August 2020).
213. Grosdidier, A.; Zoete, V.; Michielin, O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.* **2011**, *39*, W270–W277. [CrossRef]
214. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; Beer, T.A.P.D.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. [CrossRef]
215. Dominguez, C.; Boelens, R.; Bonvin, A.M.J.J. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **2003**, *125*, 1731–1737. [CrossRef] [PubMed]
216. Jian, Y.; He, Y.; Yang, J.; Han, W.; Zhai, X.; Zhao, Y.; Li, Y. Molecular Modeling Study for the Design of Novel Peroxisome Proliferator-Activated Receptor Gamma Agonists Using 3D-QSAR and Molecular Docking. *Int. J. Mol. Sci.* **2018**, *19*, 630. [CrossRef]
217. Software Website. Available online: https://www.schrodinger.com/maestro (accessed on 25 August 2020).
218. McAliley, J.H.; Bruce, D.A. Development of Force Field Parameters for Molecular Simulation of Polylactide. *J. Chem. Theory Comput.* **2011**, *7*, 3756–3767. [CrossRef] [PubMed]
219. Halgren, T.A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Compt. Chem.* **1996**, *17*, 490–519. [CrossRef]
220. Software Website. Available online: https://gaussian.com/ (accessed on 25 August 2020).
221. Forli, S.; Huey, R.; Pique, M.E.; Sanner, M.F.; Goodsell, D.S.; Olson, A.J. Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* **2016**, *11*, 905–919. [CrossRef]
222. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2009**, *31*, 455–461. [CrossRef]
223. Software Website. Available online: http://autodock.scripps.edu/ (accessed on 25 August 2020).
224. Lipparini, F.; Mennucci, B. Perspective: Polarizable continuum models for quantum-mechanical descriptions. *J. Chem. Phys.* **2016**, *144*, 160901. [CrossRef]
225. Li, D.-D.; Meng, X.-F.; Wang, Q.; Yu, P.; Zhao, L.-G.; Zhang, Z.-P.; Wang, Z.; Xiao, W. Consensus scoring model for the molecular docking study of mTOR kinase inhibitor. *J. Mol. Graph. Model.* **2018**, *79*, 81–87. [CrossRef]
226. Software Website. Available online: https://ambermd.org/ (accessed on 25 August 2020).
227. Case, D.A.; Iii, T.E.C.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688. [CrossRef]
228. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C.L. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins Struct. Funct. Bioinform.* **2006**, *65*, 712–725. [CrossRef]
229. Available online: http://www.cosmo-model.org/ (accessed on 25 August 2020).
230. Grubmüller, H.; Heller, H.; Windemuth, A.; Schulten, K. Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-range Interactions. *Mol. Simul.* **1991**, *6*, 121–142. [CrossRef]
231. Softwae Website. Available online: www.gromacs.org (accessed on 25 August 2020).
232. Softwae Website. Available online: https://www.vasp.at/ (accessed on 25 August 2020).
233. Softwae Website. Available online: https://www.ks.uiuc.edu/Research/namd/ (accessed on 25 August 2020).
234. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [CrossRef] [PubMed]
235. Karamertzanis, P.G.; Pantelides, C.C. Ab initio crystal structure prediction. I. Rigid molecules. *J. Comput. Chem.* **2005**, *26*, 304–324. [CrossRef] [PubMed]
236. Karamertzanis, P.G.; Pantelides, C.C. Ab initio crystal structure prediction. II. Flexible molecules. *Mol. Phys.* **2007**, *105*, 273–291. [CrossRef]
237. Vasileiadis, M.; Pantelides, C.C.; Adjiman, C.S. Prediction of the crystal structures of axitinib, a polymorphic pharmaceutical molecule. *Chem. Eng. Sci.* **2015**, *121*, 60–76. [CrossRef]
238. Delley, B. DFT studies: From molecules and molecular environments to surfaces and solids. *Comput. Mater. Sci.* **2000**, *17*, 122–126. [CrossRef]

239. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868. [CrossRef]
240. Softwae Website. Available online: http://www.castep.org/ (accessed on 25 August 2020).
241. Becke, A.D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652. [CrossRef]
242. Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954–9960. [CrossRef]

molecules

MDPI

Review

Current Status of Quantum Chemical Studies of Cyclodextrin Host–Guest Complexes

Anna Helena Mazurek [1] and Łukasz Szeleszczuk [2,*]

[1] Department of Physical Chemistry, Chair of Physical Pharmacy and Bioanalysis, Faculty of Pharmacy, Doctoral School, Medical University of Warsaw, Banacha 1 Str., 02-093 Warsaw, Poland; anna.mazurek@wum.edu.pl

[2] Department of Physical Chemistry, Chair of Physical Pharmacy and Bioanalysis, Faculty of Pharmacy, Medical University of Warsaw, Banacha 1 Str., 02-093 Warsaw, Poland

* Correspondence: lukasz.szeleszczuk@wum.edu.pl; Tel.: +48-501-255-121

Abstract: This article aims to review the application of various quantum chemical methods (semi-empirical, density functional theory (DFT), second order Møller–Plesset perturbation theory (MP2)) in the studies of cyclodextrin host–guest complexes. The details of applied approaches such as functionals, basis sets, dispersion corrections or solvent treatment methods are analyzed, pointing to the best possible options for such theoretical studies. Apart from reviewing the ways that the computations are usually performed, the reasons for such studies are presented and discussed. The successful applications of theoretical calculations are not limited to the determination of stable conformations but also include the prediction of thermodynamic properties as well as UV–Vis, IR, and NMR spectra. It has been shown that quantum chemical calculations, when applied to the studies of CD complexes, can provide results unobtainable by any other methods, both experimental and computational.

Keywords: cyclodextrin; host–guest complexes; DFT; QC; quantum chemistry; density functional theory; CD complexes

Citation: Mazurek, A.H.; Szeleszczuk, Ł. Current Status of Quantum Chemical Studies of Cyclodextrin Host–Guest Complexes. *Molecules* **2022**, *27*, 3874. https://doi.org/10.3390/molecules27123874

Academic Editors: Rosa Iacovino, Marina Isidori and Margherita Lavorgna

Received: 30 May 2022
Accepted: 13 June 2022
Published: 16 June 2022

1. Introduction

Due to their unique structural, physical, and chemical properties, cyclodextrins (CDs) and their derivatives have been of great interest for more than a century [1]. The biodegradability, biocompatibility, and versatility of CDs and CDs-based materials extend their applications to new areas every year; however, the main property that makes CDs so popular is their ability to form host–guest complexes with a variety of compounds [2].

CDs are commonly used in pharmaceutical formulations as they increase the solubility of poorly soluble drugs and protect substances against external factors, such as light, humidity, and heat [3]. CDs can mask unpleasant smells or flavors of drugs, which is especially important in formulations dedicated to children [4]. More than 100 original drugs are currently being manufactured with CDs as excipients [5–7].

Interactions between CDs (host) and guest molecules may yield a stable complex with a high equilibrium constant; for example, β-CD forms highly stable inclusion complexes with adamantyl derivatives with a binding constant of ~10^4–10^5 M^{-1} [8,9]. It is not surprising then that the number of newly obtained cyclodextrin host–guest complexes is constantly increasing. However, only a small amount of those complexes is being reported with their crystal structures. This is caused by the fact that most of those complexes are either amorphous or polycrystalline, and even for the crystalline complexes it is usually very hard to obtain a crystal of a size suitable for single crystal X-ray measurements [10,11]. This is probably one of the reasons why a lot of experimental works describing the structure and properties of CDs complexes are supported by theoretical calculations.

To fully understand the behavior and physicochemical properties of a complex, knowledge of its structure is essential. However, it is not just the desire to reveal how the complex looks that makes the application of molecular modeling methods so popular in the studies of CDs complexes. By choosing an appropriate computational approach, it is possible to determine (or explain) the molar stoichiometry of the complex, the differences observed in the spectra (UV–VIS, IR, NMR) of host–guest physical mixtures and their complexes, and also to predict the stability of such structures under various conditions such as different solvents, temperature or pressure.

CDs host–guest complexes are surely very flexible structures, which is the common cause of their polycrystallinity. This is why a lot of the molecular modelling studies devoted to those complexes utilize molecular dynamics simulation at the molecular mechanics level. Those works have been recently reviewed by us [12]. However, the types of intermolecular forces that stabilize such complexes such as hydrogen bonding, van der Waals, and hydrophobic and dipole–dipole interactions, usually cannot be modeled with the required accuracy using the molecular mechanics methods. This is why the number of works in which the calculations of CD complexes at the higher level of theory, namely, quantum chemical (QC), has constantly increased since 2005 (Figure 1). Now, after 20 years of studies, the number of such articles is large enough to draw some general conclusions and trends as well as the advantages and disadvantages of such an approach. Therefore, the aim of this review was to gather and analyze the works in which the CD host–guest complexes have been modeled using QC methods.

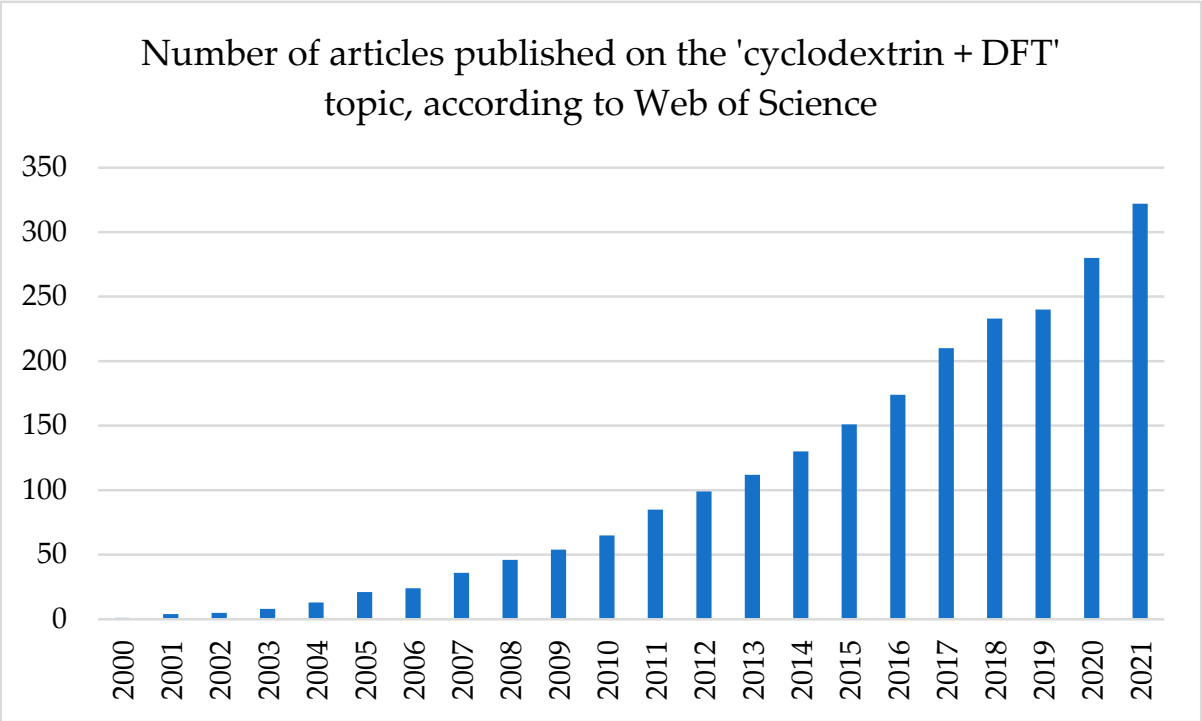


Figure 1. Number of the search results for the 'cyclodextrin AND DFT' phrase in the Web of Science. Each column shows the number of articles in the given year and all years before. For example, the column entitled '2010' depicts the number of articles published in the period 2000–2010, including 2010.

There were at least a few reasons behind writing this review. First, it was worth analyzing whether some conclusions can be made on the choice of the most accurate method, including the applied DFT functional, basis set, dispersion correction or solvent model. Those aspects are discussed at the very beginning. Then, it was interesting to check the main reasons behind the QC method chosen by the authors of the reviewed works. Was

it solely to predict or suggest the possible structure of the complex, or were the calculations used for something more such as the simulation of the spectra or explanation of the reaction mechanism to support the experimental findings? Further, we wanted to find out what type of CDs and what possible guests were studied in those computational works. Therefore, an informative table presenting the most essential information such as the composition of the complex, applied functional, basis set, and solvent treatment has been prepared to serve as an informative and easy-to-follow guide for future studies. Finally, the chosen examples are described in a more detailed way, suggesting the possible solutions and future indications.

As the authors of this review have been using the QC methods to model the structure and properties of CD complexes and found this approach very useful, it was our hope and desire to convince other researchers to try such solutions in their works.

2. Applied Calculational Methods and Parameters

2.1. Choice of QC Method

The computational methods that were used in the reviewed works nicely correspond with the general increase in the computational power available to researchers worldwide. The earliest (before 2005) QC works studying CD complexes were done using the least demanding semi-empirical methods such as AM1, PM3 and later PM6 or PM7. Subsequently, those methods have been gradually superseded by density functional theory (DFT) calculations, while recently a few works have been published in which the Møller–Plesset perturbation theory (MP) was applied. It is worth noticing that CD complexes are not small objects, in terms of QC calculations, especially when the γ-CD or substituted CDs are the hosts with the large ligand as a guest. Application of QC can also be a problem when studying complexes with a host:guest ratio higher than 1:1. This is why even in the 2020s in some works, semiempirical methods have been applied; however, the ratio of DFT to the semi-empirical ones is constantly increasing (Figure 2).

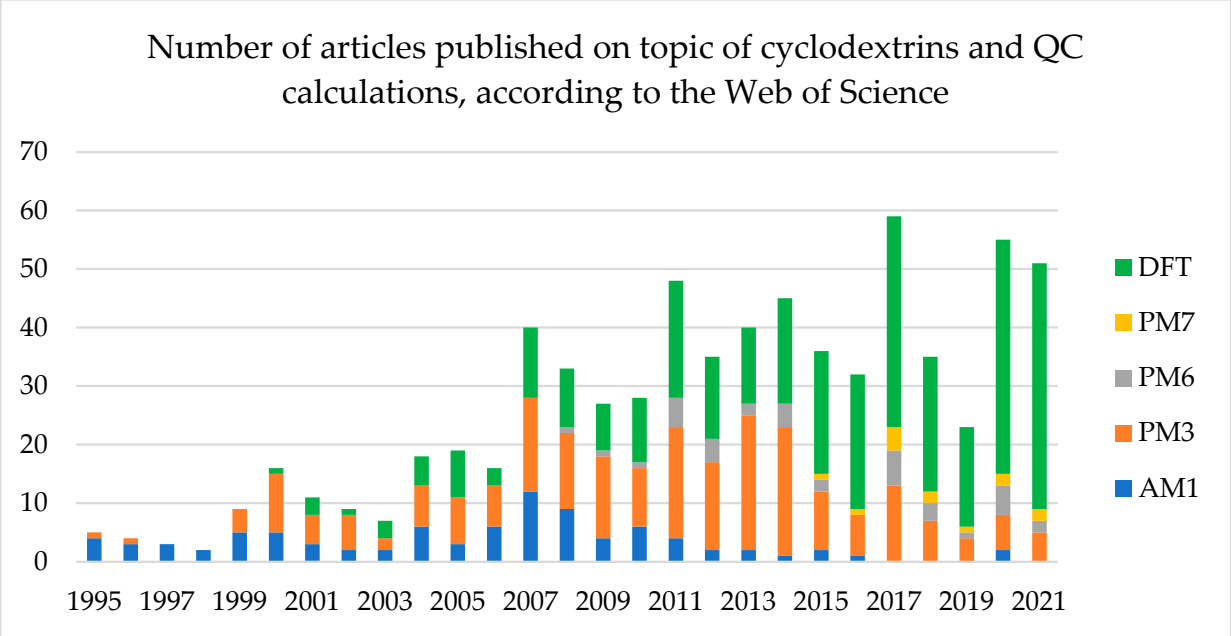


Figure 2. Changes in the number of articles published on the topic of cyclodextrin and either different semi-empirical or DFT methods over the years.

2.2. General Remarks

- Calculations of CD other than in the form of typical complexes

As stated in the title of this manuscript, this review focuses on the application of QC methods in the analysis of CD complexes. However, computational studies are also

performed for the systems where CDs play different roles, i.e., as nanocarriers, corrosion inhibitors or building blocks for even more complex structures, often with some additives such as gold particles. Those studies have not been listed in Table 1. The selected newest articles on the topic are [13–18].

It should be noticed that to increase the accuracy of DFT calculations in solid state, periodic boundary conditions are often applied, using crystal unit cells as simulation boxes. During the computations, only the properties of the original unit cell need to be calculated and then propagated in the chosen dimensions. However, to perform such computations, the crystal structure of the studied object is mandatory. More information on such calculations can be found in a recent review [19] with an example of such calculations for CDs presented in [20].

- Software

In almost all of the reviewed works presenting the results of the DFT calculations, Gaussian software was used. There are only a few cases when ORCA [21–23], VASP [24] or ADF [25] were applied instead. Only for the solid state calculations or those involving adhesion, nanocarriers, etc., is DMol3 also commonly used.

- ONIOM

In several works that included CD where DFT was applied, the ONIOM method was presented (see Table 1). ONIOM stands for Our own N-layered Integrated molecular Orbital and Molecular mechanics and is a hybrid method which combines QC (either ab inito or semi-empirical) and molecular mechanics methods in order to reduce the computational cost [26]. On the basis of the ONIOM results, other properties such as thermodynamic ones are calculated. This approach was for years popular in the computation of the CD systems, as CDs are relatively big structures, and for a long time it was not possible to calculate both the CD and a guest using DFT. Therefore, CD was considered an outer layer and there a lower level of theory was applied, and the guest molecule was computed using a higher level of theory. However, with the general increase in computational power available, this approach is currently rarely used in the studies of CD complexes due to its lower accuracy when compared with pure QC calculations.

- Molecular dynamics

An MD simulation is a well-established technique used for the study of various molecules complexes and mixtures in any state of matter and at almost any temperature and pressure condition. It can be used to determine structural, energetic, and thermodynamic properties as well as a means to scan the potential energy surface of a studied system.

For MD simulations of large molecular complexes, such as ligand–protein, molecular mechanics (MM) methods are commonly used. On the contrary, when MD simulations are performed on relatively small molecules, it is usually at the quantum mechanics (QM) level of theory, which significantly increases the accuracy of calculations, but also their computational costs. In terms of the sizes of the modeled objects, CD complexes are somewhere in between. While geometry optimization calculations on the static structures of CD complexes are, nowadays, performed mostly at the QM level, usually by the means of DFT, the MD simulations are still being performed at the MM level [12].

2.3. Semi-Empirical Methods

Semi-empirical methods are based on the Hartree–Fock equation but simplified by the application of the empirical corrections [27]. Semi-empirical calculations are much faster than their ab initio counterparts, mostly due to the use of the zero differential overlap approximation. Their results, however, can be very wrong if the molecule being computed is not similar enough to the molecules in the database used to parameterize the method. Here, we will concentrate on the most popular semi-empirical methods, that is AM1, PM3, PM6, and PM7.

According to Figure 2, the most popular semi-empirical method applied for the cyclodextrin complexes is PM3. However, it may be argued whether PM3 delivers better results than the newer generations, PM6 and PM7. The reason for wider application of PM3 instead of PM6 and PM7 may be the fact that some researchers got used to PM3 and some software does not support the newer parametrizations (PM6 and PM7). Since 2019, the DFT methods have strongly overtaken the role of a leading quantum chemistry-based calculation method in the cyclodextrin systems. Naturally, DFT is more precise than any semi-empirical approach, which leaves no space for further investigation, and among the semi-empirical methods, seems to be the most reliable in the complexes in question. Below, the application of the semi-empirical methods in the cyclodextrin-including complexes is described based on the examples from 2015–2022. However, it should also be noticed that the PM3, PM6, and PM7 approaches often co-exist with the DFT ones. In those cases, the systems are first optimized with the semiempirical method to obtain the initial structure for the DFT calculations.

The oldest of all here presented semi-empirical methods is the AM1 approach. There are just a few examples of the cyclodextrin complexes calculated using this approach. In some of them, AM1 was applied to perform the geometrical optimization of a whole guest–CD complex [28]; in others, both substrates were calculated by AM1 whereas the complex underwent the DFT treatment [29], and in others, AM1 was applied solely for the CD.

The next generation of semi-empirical methods is PM3. There are studies showing that PM3 predicts the presence and energy of hydrogen bonds better than AM1 [30]. Moreover, it has been reported that after α- and β-CD optimization with AM1 and PM3, the former resulted in badly distorted geometries, whereas the latter reproduced the crystalline structure rather well [31].

Semi-empirical methods are repetitively reported to deliver good insight into the complex formation process as well as reliable order of the configuration stabilities [32–35]. These methods help to determine global or local minima. However, it is stressed that to obtain reliable complexation energy values, DFT calculations should be performed [36]. Often, the PM3 approach is undertaken along with the ONIOM DFT/PM3 approach [34,37–44].

The most often and standard PM3 application in the cyclodextrin complexes is to move the guest along the selected axis going through the CD cavity. The guest is stopped every 1 Å, usually between −8 Å or −10 Å and respectively +8 Å or +10 Å. Additionally, the guest is rotated from 0°to 360° usually every 20° or 45°. At each such stopping point, the complexation energy is measured with PM3 [37,44–47].

In the way of their application to the cyclodextrin systems, the PM6 and PM7 approaches follow the same pattern as PM3. Namely, they are used to gain insight into the complex structure [48] and thermodynamic properties [48–52], which allows to determine the most stable complex [52–55], and, as in case of ofloxacin enantiomers, rank the eluted substances in the order in which they will be eluted [49]. Similarly, as in the case of PM3, FT-IR spectra can be simulated [50,56]. Further, PM6 and PM7 are often combined with the DFT methods in form of the ONIOM approach [42,52,57,58].

In contrast to PM6, in PM7, the description of dispersion interactions and hydrogen bonding has been improved, and consequently the errors associated with modelling large molecules and complexes have been reduced [49]. The description of properties such as heat of formation or height of the reaction barrier has been improved [59]. In turn, it has been reported that PM6, compared with PM3, can yield better agreement with the experimental values [59]. In another study, when compared to the experiment, PM3 provided wrong and at the same time opposite results to those obtained by PM7 [60]. To depict another example, in a study where β-CD, dimethyl-β-CD, and hydroxypropyl-β-CD were analyzed by both PM6 and PM7, in all cases PM7 delivered complexation energies of significantly lower values [58].

A separate topic is ADMP, the Atom Centered Density Matrix Propagation Molecular Dynamics approach, which can be performed with semi-empirical, Hartree–Fock or DFT

methods. It provides equivalent functionality to Born–Oppenheimer molecular dynamics at a considerably reduced computational cost. The ADMP method has a number of attractive features. Systems can be simulated by accurately treating all electrons or by using pseudopotentials. Through the use of a tensorial fictitious mass and smaller values of the mass, reasonably large time steps can be employed, and lighter atoms such as hydrogens need not be replaced with heavier isotopes. A wide variety of exchange-correlation functionals can be utilized, including hybrid density functionals [61]. In the last decade, only two cases have been published: β-CD-olsalazine with PM3-ADMP [62] and β-CD-propranolol with PM6-ADMP, ONIOM(DFT/PM3)-ADMP, and DFT-ADMP approaches [63]. ADMP results confirm the importance of the non-bonded interactions in the complex stabilization.

To sum up this part of the review, a relatively new and complex study should be cited. In the article entitled 'Prediction of correct intermolecular interactions in host–guest systems involving cyclodextrins' [63] published in 2020, the following approaches were tested: AM1, PM3, PM6, and DFT with the most standard B3LYP/6-31G(d,p), as well as PM6 and DFT with and without dispersion correction. The study involved 15 α-CD and 28 β-CD inclusion complexes in terms of both geometrical parameters and complexation energy values. The results showed that the most accurate was the B3LYP/6-31G(d,p)-D3 approach, followed by PM6-D3. Nevertheless, taking into account the high computational requirements of the DFT methods, the authors suggest that PM6-D3 is the most accurate and cost-effective approach. However, it must be stated that as the availability of the computational power is developing quickly, the DFT approach might shortly be, or already is, the best option for the computational analysis of the structure and energy of CD systems.

2.4. Density Functional Tight Binding (DFTB)

An approach that can be positioned between semi-empirical methods and DFT is the Density Functional Tight Binding Self Consistent Charge method (DFTB-SCC, referred here to as DFTB), which is often described as DFT approximation [64]. According to the Web of Science, just a few articles referring to DFTB and CD complexes have been published. However, those works show a relatively wide spectrum of possible DFTB applications. In the oldest works [64,65], DFTB was applied to verify the experimental NMR results and deliver some additional structural information. In the first case [64], the method's application confirmed which conformation of spironolactone was preferred in the complex. In the second case [65], DFTB confirmed that in the analyzed peptide, tyrosine was a favored residue to access the CD's cavity. In this second study, DFTB was applied within the QM/MM approach. In both of the described cases, the authors claimed good agreement of the obtained computation data with the experimental data.

A more complex situation is described in [66], where the topic is self-inclusion (in the own cavity) of the CD's substituents. Here, DFTB was compared with the DFT approach. DFTB is said to be method of choice as it predicts the stability order of the analyzed complexes properly, delivers the energy data that are close to the DFT data, and is faster than DFT. Both options including dispersion correction and the absence of this correction have been tested in DFTB and in DFT. The results plainly show that in both cases, application of the correction is necessary. In the case of DFTB, the dispersion-corrected version of calculations result in lower complexation energies and the order is maintained. Nevertheless, the authors emphasize that the application of an empirical dispersion correction may significantly overestimate dispersion interactions, and therefore a comparison with a rigorous DFT method should be done.

Some drawbacks of DFTB have been pointed out in the work that targeted the largest CD for which the crystallographic structure is known [67]. The heavy atoms' RMSD between the optimized structure and the crystallographic one were between 0.89 Å and 1.35 Å for DFT, depending on the parameters applied, with the best results for the B3LYP functional and 0.95 Å for DFTB. However, even if the overall RMSD looks good, large

discrepancies in the angles sizes when compared with the experiment have been reported in the case of the DFTB approach as opposed to all DFT methods.

In turn, the article from 2018 utilizes the DFTB approach to analyze the tautomerization process during encapsulation of genistein in CD [68]. The results are clear and deliver important information on the complexation, namely, 'DFTB-based MD simulations reveal that spontaneous keto-enol tautomerization occurs even within a hundred picoseconds, which suggests that the encapsulated genistein is complexed in the ordinary enol form of the drug molecule'.

2.5. Density Functional Theory (DFT)

- Functionals

While performing the DFT calculations, the main two parameters that must be decided on are type of functional and a basis set. Functionals mathematically define the electronic energy, which when added to the kinetic and electrostatic energy of the system, sums up to the total system's energy [68,69]. According to the literature (see Table 1) for the systems that included CD, the hybrid B3LYP [70], semi-empirical GGA (generalized gradient approximation): wB97XD, B97D3 [71] or meta-GGA kinetic energy density incorporating Minnesota [72] functionals have been applied so far. Among the last category, M06-2X, M05-2X, and M06-L are used, with M06-2X being the most common in the analysis of the non-covalent interactions, whereas M05-2X includes 0% Hartree–Fock (HF) exchange, and M06-2X has 54% HF exchange [73]. In one of the studies including CDs, it was concluded that M06-L delivered poor results [74]. This was expected as the analyzed complex was β-CD-alprazolam, whereas M06L has been designed for calculations of the systems including transition metals, inorganic or organometallics [21]. wB97X and 97D3 are comprised of 22% Hartree–Fock exchange in the short range and 100% Hartree–Fock in the long range [75].

- Dispersion correction

Noncovalent forces, such as hydrogen bonding and van der Waals interactions, are crucial for the formation, stability, and function of most CD complexes. At present, ubiquitous van der Waals interactions can only be accounted for properly by high-level quantum-chemical wavefunctions or by the Quantum Monte Carlo method. In contrast, the correct long-range interaction tail is absent from all popular local-density or gradient corrected exchange-correlation functionals of DFT, as well as from the Hartree–Fock (HF) approximation. A long-range electron correlation effect, known as the London part of the dispersion energy term, is not included in the Kohn–Sham DFT equation [76]. For years this was an issue affecting the accuracy of the DFT calculations. Nowadays, several dispersion correction methods are available. Nevertheless, their inclusion not always improves the calculation effect, hence this should be tested separately for each system in question. The most widely used dispersion corrections are TS (Tkatchenko-Scheffler) [77], GD (Grimme Dispersion, written also as D) [78], and MBD (Many-Body Dispersion) [79]. However, to perform calculations on the systems that included CD, almost solely the semi-empirical Grimme dispersion correction was applied (see the Table 1). It occurs in the D2, D3, D4, and D3(BJ) versions, where BJ indicates Becke Johnson damping. This last one is rarely used in CD-complex calculations. It is claimed that 'the damping function in DFT-D methods has only a minor impact on the quality of the results' [80] and even the comparison between D3 and D3 (BJ) published by Stefan Grimme clearly indicates that 'the differences between the two methods are much smaller than the overall dispersion effect' [80]. D3 includes less empirical input than D2, can be called a newer D2-version, and is the dispersion correction that is currently the most widely used.

No dispersion correction is applied to the Minnesota functionals that are parametrized for dispersion. The same applies to wB97X (written also as ωB97X), which is the dispersion-incorporating version of B97X. B3LYP is sometimes used as B3LYP-CAM (Cambridge extension) [81], which includes the long-range correction; however, this is not common,

and there is no strong evidence that this type of dispersion correction gives better results than application of the Grimme correction.

Sometimes it seems reasonable to perform the same calculations with two chosen functionals as has been done for the β-CD-2,2′-bipyridine complex [82]. The authors' conclusion is that wB97XD showed reliability in elucidating weak interactions, whereas B3LYP allowed one to achieve a good time–precision compromise.

Another interesting example is comparison of three types of functionals performed for the β-CD-procaine HCl system [83]. According to the authors, B3LYP showed the highest efficiency and quality of results, wB97XD was specifically used to analyze the long-range interactions, and M06-2X was applied to predict the presence of hydrogen bonds. On the other hand, there are works such as [84] (β-CD-8-Anilinonaphthalene-1-sulfonate) where it is claimed that among tested functionals, B3LYP, wB97XD and M06-2X, the overall best results were delivered by wB97XD.

To take one more example, for the β-CD-benzocaine system [53], where B3LYP, CAM-B3LYP, M05-2X, and M06-2X were tested, M06-2X was claimed to deliver excellent results when used to obtain the NMR spectra. In another study, for UV–Vis spectrum simulation, the B3LYP-D3-based results showed the best agreement with the experimental data. The tested functionals were BLYP-D3, B3LYP-D3, and M06-2X-D3 [22]. In the case of β-CD-5-fluorouracil, inclusion of the dispersion correction changed the interaction energy by 15% and by 20% in water and ethanol solvent, respectively [85].

This only allows us to draw three important conclusions. Firstly, a couple of functionals, preferably representing all three groups (hybrid, GGA, Minnesota) should be tested for each system. Secondly, the choice of a functional depends on the goal of the study: geometry optimization, thermodynamic parameters, NMR spectra, etc. Thirdly, this literature review sends a clear message that the three most used and effective functionals are B3LYP-(D3), wB97XD, and M06-2X.

- Basis set

It should not be forgotten that in the study preparation, the functional and dispersion correction and the basis set choice play a significant role. The basis sets applied for the CD complexes so far are the Pople, correlation-consistent (cc-pVDZ), and Karlsruhe (def-TZVP, def2-SVP) basis sets (see Table 1). However, the Pople ones are definitely the most common and the variety among them is large, for instance, 6-31G(d), 6-31G(d,p), 6-31+G(d,p), 6-311++G(d,p), 6-31G**, where '1′ means basis set enlargement, '+' means an additional diffuse function, and '*' means a polarization function. The choice among different Pople basis sets depends partly on the available computational possibilities and partly on the type of studied objects. Inclusion of diffuse functions is needed to properly calculate the long-range interactions, such as hydrogen bonds. In turn, by extending the size of a basis set, the addition of polarization functions is meaningless. This is why, in the works presented in Table 1, the most presented basis sets are 6-31G(d,p) and 6-311G(d,p).

In Table 1, it is noticeable that the '6-31+G* for H, N, O and 4-31G for C' basis set combination is present. However, first of all, several of these works have been published with one affiliation, in other words, there is one laboratory that uses such an approach, and secondly, such a combination of basis sets for one system has been used mainly in the past when insufficient computation possibilities were at hand. In order to perform computation using the limited available tools, it is common practice to perform the geometry optimization in a lower basis set and later, for instance, for single point calculations, a larger basis set is used. An analogical approach applies to the more and less computationally demanding functionals.

2.6. Solvent

The last parameter to decide on refers to the environment of the system. Calculations can be performed in gas or in solvent. To simulate a solvent in DFT, implicit solvent models are typically used. For the CD systems, the most popular is the family of the Polarizable Continuum Models (PCM) [86]. The other possibility is the Solvation Model Based on

Density (SMD). IEFPCM [87] is a reformulation of the dielectric PCM, and the C-PCM is a conductor-like PCM, closer to the COSMO model. SMD defines the free energy of solvation via two components: the one is electrostatic contribution arising from the self-consistent reaction field, the other comes from the short-range interactions between the solute and solvent molecules [88].

Both PCM and SMD treat the solvent as a continuum because using Quantum Mechanics, it would not be possible to calculate a system with a large number of explicit solvent molecules. However, this issue can be partly approached as has been done in the study [64]. There, for general calculations, IEFPCM has been applied, but additionally a separate set of calculations has been done on the limited number of water molecules placed inside the CD cavity. Such an approach is useful if there is the probability that the CD–guest interaction is influenced significantly by the solvent's presence. For example, when it is assumed that the hydrogen bonds between the guest and host molecules are water mediated. Sometimes, inclusion of the solvent effect decreases the complexation energy significantly, as in the case of the dexamethasone and SMD model [89]. However, it must be pointed out that inclusion of a solvent in the calculated system not always results in better (closer to the experimental data) complexation energies.

Taking into account all what has been written above, in order to make a fully justified selection of the parameters for the DFT calculations, a cross-study including different but most commonly used functionals, dispersion corrections (presence or absence), basis sets, and the environment (gas or solvent, type of solvation) should be performed. To the authors' best knowledge, so far no such study has been undertaken on any system that includes CD. The already published benchmark studies are quite uncommon and usually focus on modification of one of the parameters, i.e., functionals, dispersion correction or the solvation method [67,90–92].

Only afterwards, with the parameters chosen carefully for the analyzed system (e.g., CDs+steroidal hormones, CDs+flavonoids etc.), should further calculations be performed.

2.7. Møller–Plesset Perturbation Theory 2 (MP2)

Only a few (five in the period from 2014–2021) articles in which the MP2 method has been applied for CD analysis have been published. The reason is the fact that this technique is computationally more demanding than DFT and since CD complexes are relatively large, as the objects for QC studies, this method is currently not affordable for most of the computational researchers.

The most recent work in this topic was published in 2017 and concerns β-CDs with one large substituent that can either be located in or outside of the CD's cavity [68]. The geometry optimization has been performed with the B3LYP-D3 functional but single point calculations already with both B3LYP-D3 and MP2 using various basis sets, for MP2: 6-31G* and 6-311G*.

In another study (β-CD-sertraline) [93], MP2/6-31G(d,p) was applied for the single point calculations, although even the authors of the work state that such a small basis set does not allow one to obtain results with the required accuracy.

3. Preparation of Structures, Post-Processing Methods, and Some Examples

3.1. Preparation of the CD Complexes for the QC Calculations

To obtain the structure of the CD complex that can be used for DFT calculations, both the structure of the chosen CD as well as the structure of the guest molecule must be prepared beforehand. The method of in silico complex preparation is also important, as it may have a major influence on the results.

Since the crystal structures of all of the native and also some of the modified CDs can be found in the CCDC [94], they are usually used as the starting points for calculations. The structure of the guest is either simply drawn using one of the multiple available software packages or taken from the CCDC, assuming that its crystal structure has been deposited previously.

Very important, for the accuracy of the results, is the method of preparation of the complex from its components. This must be done unless a crystal structure of the complex has been obtained or deposited previously in the CCDC, which is unfortunately quite uncommon. Usually, one of the two approaches is used. In the first one, molecular docking is applied, treating CD as a macromolecule and the guest as a ligand. In the reviewed works, the authors usually use the popular Auto Dock [95] software for that purpose. However, some other programs are also used such as Schrodinger Maestro [96] or BIOVIA Discovery Studio [97]. Surprisingly, in the reviewed studies, not much attention was being paid to the description of this part, which was justified by the fact that the initial (docked) structures would be optimized at the higher theory level. This may, however, lead to some inaccurate or even wrong results as the energetically lowest conformation obtained from the docking part may not necessarily be close to the global DFT minimum. This is nicely reflected in Figure 3, where the energetically lowest orientation obtained from molecular docking is substantially different from the experimental one, even after optimization using DFT. However, when the other pose from molecular docking was optimized using QC, much better agreement between the experimental and theoretical results was obtained. Therefore, in some studies, the authors decided to optimize not one but a few different complexes from molecular docking. While this approach is reasonable as it increases the likelihood of finding the deep minimum, it should be noticed that the time of calculations increases linearly with the number of initial structures.

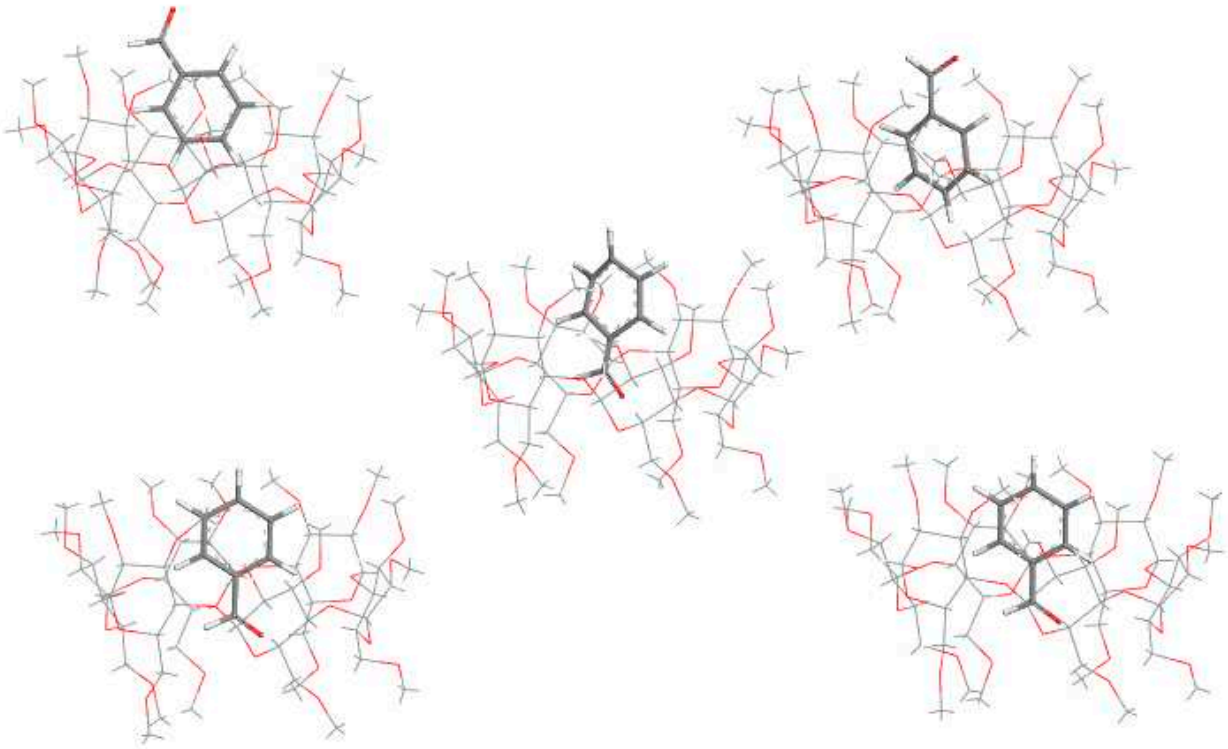


Figure 3. Comparison between the structures of the αCD complex with benzaldehyde. Top left: the best pose from molecular docking; top right: the best pose from molecular docking after optimization using DFT; bottom left: one of the poses obtained from molecular docking; bottom right: "bottom left" structure after optimization using DFT; middle one: experimental structure (CSDC ref. code: BOHWUQ). It should be noted that while the top left structure has energy lower than the bottom left by 3.4 kcal/mol, the top right structure has energy higher than the bottom right by 4.2 kcal/mol. Source: author's archive.

Instead of molecular docking based on the molecular mechanics calculations, in some cases, the authors decided to manually dock the guest into the CD. In order to find the best pose within the cavity, the guest molecule is put in different positions along the selected

axis, so that the guest has different levels of immersion into the CD's cavity. For example, in [98], the guest was moved along the *Z*-axis from +7.5 Å to −7.5 Å with an interval of 0.3 Å. Additionally, in this particular study, the guest was rotated around the *Z*-axis by 3° from 0° to 360°. In each step, the generated systems underwent geometry optimization calculations. This type of systematic search seems to be the most accurate approach, especially for the ligands with limited conformational space, with the only drawback being the increase in the calculation time.

The other important factor that is usually neglected in the QC studies of CD complexes is the conformational flexibility of the guest molecule. The optimal conformation of the guest found *in vacuo* is not necessarily the one that it takes in the complex. To increase the likelihood of finding the deep energetical minimum, the conformational space search of the guest molecule should be performed.

Another aspect that must be taken into consideration is the host–guest molar ratio of the complex. When there are some experimental indications for a specific value, the assumption can be tested using QC calculations. Otherwise, it seems reasonable to prepare the complexes of various stoichiometry and confirm their stability via geometry optimization.

3.2. Description of QC Results

After the host, guest, and complex are optimized at the chosen QC level, the interaction and stabilization energies are obtained. Stabilization energy is defined as the difference between the energy of the fully optimized geometry complex and complex components: CD and guest (Equation (1)) [99]. Interaction energy is defined as the corresponding single point energy.

$$E_{stb} = E_{cplx\_opt} - (E_{CD\_opt} - E_{guest\_opt}) \quad (1)$$

Sometimes the solvation energy is taken into account as well. It is calculated as the difference between the complex energy in water and in gas. Thermodynamic parameters (TD) are often calculated, as they give more insight into the stability of the analyzed systems.

The collected data allow us to draw conclusions, which forces determination of the complex creation: dispersion [100], van der Waals [63] interactions or hydrogen bonds [61]. TD results allow us to determine whether the complexation process is enthalpy driven [51,84,100], which is said to relate to the number and strength of the intermolecular interactions within the system, or entropy driven, which is quite rare for those complexes. Inclusion of the temperature effects allows us to observe if and how the temperature affects the complex stoichiometry, as in the case of β-CD-pentoxifilline [101].

Additionally, often IR or UV–Vis spectra or NMR chemical shifts are calculated (see Table 1) and in the majority of cases, the results are claimed to have very good agreement with the experimental data.

Another common practice is the application of the QTAIM method (Quantum Theory of Atoms In Molecules) for the DFT-optimized complexes in order to analyze weak interactions and therefore obtain a better understanding of the complex's structure at the molecular level. Several articles about CD complexes including this approach have been published (see Table 1).

3.3. Analyzed CD Complexes

Among the large variety of CDs, in the DFT studies, mainly only β-CD has been applied so far, and a variety of guests in CD complexes has been analyzed, as presented in Table 1. The guests are mainly drugs among which the antidepressants seem to be especially targeted (Table 1A.1) as well as plant derivatives with (potential) medical use (Table 1B). A separate group consists of substances that could be defined as functionalized food (Table 1C). In those cases, CDs serve to protect or even increase the antioxidative capacity of the substances in question, for example (−)-gallocatechin, (−)-catechin gallate and (−)-gallocatechin present in tea [102–104], or to reduce the bitter taste of coffee [105,106]. Moreover, CDs can be used as chiral selectors, and this has its reflection in the DFT articles (Table 1D).

Some of the analyzed systems in Table 1 have been already described in the previous paragraphs as examples regarding the applied parameters and computation. Other selected examples showing a particular usefulness and applicability of the DFT methods as well as the obtained results are described below.

The examples that directly show a vast area of DFT applicability for CD systems analysis are studies of the β-CD complexes, with 8-anilinophthalene-sulfonate [84], benzyl isothiocyanthe [23], metheonine [98] or vanillina [107]. In these works, the DFT approach was used to perform geometry optimization, obtain interaction and stabilization energies, calculate thermodynamic properties, use QTAIM and NBO analysis approaches, and simulate NMR and absorption spectra. The obtained data allow the screening of possible conformations, to define the interactions (van der Waals, hydrogen bonds, etc.) determining the CD–guest interaction, rank the complexes according to their stability, complement experimental spectra, and support the signal assignment.

Further, such chemical information happens to be a crucial part of new theses. For instance, DFT calculations revealed that the nicotine forms have considerably stronger binding with β-CD rather than with Mβ-CD in the same orientation with lower complexation energy. This explains why after 21 days the remaining nicotine increased from 65.56% in pure nicotine to 89.32% and 76.22% in β-CD-nicotine and Mβ-CD-nicotine complexes, respectively [34].

Another example is imipramine and desipramine β-CD complexes [108]. DFT calculations revealed an alternative inclusion scenario: via a guest's side chain and not via the aromatic moiety. Thus, the controversy in the experiments has been explained because such bimodal complexation increased the therapeutic effect of the substances.

For another antidepressant, paroxetine, the DFT calculations helped to confirm the existence of a new CD inclusion polymorph: a new 2:1 stoichiometry complex has been described [109]. It is characterized by a stronger presence of the dispersion interactions and is more energetically favorable than the 1:1 complex, which improves the drug's bioavailability.

Again, when it comes to structural information, the DFT approach showed that in 1:2 Cu-flavonoid and 1:3 Fe-flavonoid β-CD complexes, in morin and quercetin, the 3-OH site, and in primuletin, the 5-OH site, were utilized as preferable chelation sites [25]. These data are helpful for scientists trying to obtain an effective CD-flavonoid antidiabetic formulation.

Table 1. Selected articles published in the years 2015–2022 on the application of DFT methods for systems that included CD. The functional and basis set information concerns CD complexes, not guests. Abbreviations used in table: DM-CD (2,6-dimethylo-CD), TM-CD (trimetylo-CD), per-M-CD (permethylated-CD), geo. opt. (geometry optimization), SP (single point calculations), NMR (1H NMR spectra simulation), NBO (Natural Bond Orbitals), BJ (Becke Johnson damping function),TD (thermodynamics calculations), n.i.p. (no information provided). In the case where the DFT application in the published research occurs only as an ONIOM component, the article has not been included in the table. The ONIOM approach along with the examples has been described in Section 2.2.

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
A (potential) drugs							
1	β	(s)-2-Isopropyl-1-(o-nitrophenyl) Sulfonyl) Aziridine	B3LYP, WB97X-D, B97D3	6-31G(d)	gas, water		[110]
2	β	boron-based aromatic systems	BLYP-D3(BJ)	def2-SVP	vacuum, CPCM	geo. opt., natural bond orbital calculations (NBO), complexation energy	[100]

Table 1. *Cont.*

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
3	α, β, γ	alprazolam	B3LYP, M06L	def-TZVP	vacuum	geo. opt. in gas, NMR spectra	[21]
4	β	lenalidomide	B3LYP, M06-2X	6-31G(d,p)	PCM		[111]
5	β	dexamethasone	BLYP-D4	def2-TZVP	gas, water	geo. opt., complexation energy	[89]
6	β	2,2′-Bipyridine	B3LYP, wB97XD	6-31G(d)	PCM (eight solvents)	geo. opt., UV–Vis spectrum, HOMO-LUMO	[82]
7	β	2,2′-Dipyridylamine	B3LYP	6-311++G(d,p)	PCM		[112]
8		vardenafil hydrochloride	B3LYP	6-311G(2d,2p)	vacuum	geo. opt., FT-IR	[113]
9	amino-CD	doxorubicin	B3LYP	6-31G	vacuum	geo. opt., complexation energy, HOMO-LUMO, dipole moment, chemical potential, electrophilicity	[114]
10	β	5-fluorouracil	B3LYP-D3	6-31+G(d,p)	vacuum, PCM	geo. opt., complexation energy, harmonic frequency calculations	[85]
11	HP-β	2-methyl mercapto phenothiazine	B97-D3, BP86-D3	6-31G(d,p)	gas, CPCM	geo. opt., vibrational spectra, NBO, QTAIM, HOMO-LUMO	[115]
12	β	vemurafenib	ωB97XD	6-31+G(d)	vacuum, PCM	Geo. opt., vibrational spectra, MD, NBO, TD, HOMO-LUMO	[116]
13	β	procaine hydrochloride	B3LYP, M06-2X, WB97XD	6-31G(d,p)	gas, PCM	Geo. opt., NBO	[83]
14	β, SBE-β	fluorometholone, cholesterol	M06-2X	6-31G**	PCM	Geo opt., interaction energy	[117]
15	α, β, γ	chlordecone	M06-2X-D3	6-31G(d,p)	SMD	Geo. opt., QTAIM	[118]
16	β, methyl-β	nicotine	M06-2X	6-31G(d,p)	n.i.p.	Geo., opt., complexation enrgy	[34]
17	β	8-Anilinonaphthalene-1-sulfonate	B3LYP, M06-2X, WB97X-D	6-31G(d)	gas, water	Geo. opt., interaction energy, NMR, TD, NBO	[84]
18	β	benzocaine	B3LYP, CAM-B3LYP, M05-2X, M06-2X	6-31G(d,p)	PCM	Geo. opt., QTAIM, NBO, NMR, HOMO-LUMO, TD	[53]
19	β	aryl pentazole	M06-2X	6-31+G(d,p)	PCM	Geo. opt.	[119]
20	β	2,4D, dicamba pesticides	PBE1PBE (PBE0), B97-D, M06-2X	6-31G(d,p)	gas, SMD	Geo. opt.	[120]

Table 1. *Cont.*

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
21	Monochlorotriazinyl-β	permethrin, cyppermethrin	BLYP (geo. opt.); BLYP-D3, B3LYP-D3, M06-2X-D3 (UV–Vis)	def2-SV(P) (geo. opt.); TZVP (UV–Vis)	COSMO	Geo. opt.	[22]
22	β	dopamine	B3LYP, MPW1PW91, M05-2X, M06-2X, ωB97X-D	3-21G*	CPCM	Geo. opt., complexation energy, QTAIM, NBO	[61]
23	α	benzoate derivatives	M06L (geo. opt.); M06-2X//M06-L (SP)	6-31+G(d,p)	gas	Geo. opt.	[121]
24	α, β, γ	cholic, deoxycholic acid	B97-D, M06-2X, B3LYP	6-31G(d)	PCM	Geo. opt., interaction energy	[122]
25	α	benzoate derivatives	M06-2X//M06-L, M06-2X//BLYP, BLYP, M06-2X	6-31+G(d,p)	gas	Geo. opt., interaction energy	[123]
26	γ	cetirizine	B3LYP	def-TZVP	n.i.p.	Geo.opt., interaction energy, HOMO-LUMO, DOS, NMR	[124]
27	succinyl-β	uranium	M06-2X	6-31G(d,p)	SMD	Geo. opt.	[125]
28	β-CD, DM -β	thymidine-carbonate	B3LYP-GD2	6-31G(d,p)	PCM	Geo. opt., complexation energy, TD, HOMO-LUMO, NMR	[126]
29	β	glycyl-L-phenylalanine	B3LYP	3-21G(d)	PCM	Geo. opt., interaction energy, HOMO-LUMO	[127]
30	β	sodium salicylate	B3LYP	6-31G(d)	gas, PCM	Geo. opt., solvation energy, relative stabilization energy, complexation energy, change of volume	[128]
31	β	benzyl isothiocyanthe	B97-D3	def2-SVP	vacuum	Geo. opt., complexation energy, HOMO-LUMO, NBO, NMR	[23]
32	α	iodine solution	CAM-B3LYP	6-31*G	PCM	Geo. opt., absorption spectra, HOMO-LUMO	[129]
33	β	meta-aminophenol	M06-2X	6-31G(d,p)	IEFPCM	Geo. opt., complexation energy, HOMO-LUMO, TD, NBO	[130]

Table 1. *Cont.*

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
35	β	L-glutamine	B97-D3	6-31G(d)	n.i.p.	Geo. opt., complexation energy, TD, NBO, QTAIM	[131]
36	β	R and S ibuprofen	M062X	6-31G(d,p) (geo. opt.); 6-311++G(d,p) (SP)	gas, SMD	Geo. opt., solvation energy	[132]
37	α, β	thioureides	B97-D3	6–31G(d,p)		Geo. opt., interaction energy	[133]
38	β	mepivacaine	B97-D3	6-31G(d,p)	gas, SMD	Geo. opt., interaction energy, TD	[134]
39	β	L-metheonine	WB97-D3	6-31G(d)	PCM	geo. opt., interaction energy, QTAIM, TD, NMR	[98]
40	β	prazosin, losartan	B3LYP	6–311+G(d,p)	gas	Geo. opt.	[135]
41	β	olsalazine	B3LYP, WB97-D3, CAM-B3LYP (UV-vis)	6-31+G(d)	PCM	Geo. opt., ADMP	[62]
42	β	aspirin	B3LYP-D3	cc-pVDZ	gas	Geo. opt., qTAIM, NBO	[136]
43	β	quinine	B3PW91	6-311++G(d,p)	PCM	Geo. opt.	[137]
44	β	erlotinib	B3LYP	6-31+G*	n.i.p.	Geo. opt., harmonic frequencies, HOMO-LUMO	[138]
45	γ	rocuronium, vecuronium	B3LYP	6–31+G(d,p)	n.i.p.	Geo. opt., NBO, HOMO-LUMO	[139]
46	α, β, γ	cathinone	M05-2X	6-31G(d)	gas, CPCM (water, chloroform, methanol)	Geo. opt., QTAIM, NBO, IR spectra, TD	[140]
47	α	CO2	B3LYP	G-31G*	PCM	NMR	[141]
48	β	flutafemic acid	B3LYP, M05-2X	6-31G(d)	vacuum, water	Geo. opt., complexation energy, TD, NMR	[142]
49	2-HP-β	Cu (II) and Fe (III) complexes of quercetin, morin, primuletin	B3LYP	6-311++G**	n.i.p.	Geo. opt., complexation energy, HOMO-LUMO	[25]
50	β	6-thioguanine, 6-mercaptopurine	B3LYP	6-31+g(d,p)	IEFPCM (DMSO)	Geo. opt., interaction energy, TD	[37]
51	β	N-(2-chloroethyl),N -nitroso,N′,N′-dicyclohexylsulfamid	B3LYP	6-31G(d)	PCM (DMSO)	Geo. opt., NBO, QTAIM	[143]
52	β	benzaldehyde	B97-D	6-31G(d,p) (geo. opt.); 6-311++G(2d,p) (SP)	gas, SMD	Geo. opt., interaction energy, TD	[144]

Table 1. *Cont.*

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
53	α	chitibiose	M06-2X	6-311++G**	n.i.p.	Geo. opt., NBO, QTAIM	[145]
54	α	hydrated and nonhydrated IIA/IIB group metal cations	M06-2X	6-31G(d,p)	gas, PCM	Geo. opt., interaction energy, TD	[146]
55	β	nabumetone	WB97X-D, B97-D, B3LYP, M05-2X, M06-2X	6-31G(d)	IEFPCM	Geo. opt., NBO, QTAIM	[40]
56	β	propranolol	B3LYP, ωB97XB (ONIOM)	6-31+G(d)	gas, IEFPCM, explicit solvent effect: explicit water molecules inside of the complex	Geo. opt., interaction energy, ADMP, TD	[59]
57	functionalized CDs	8-hydroxyquinoline ligands	B3LYP	6-31G**	n.i.p.	Geo. opt.	[147]
58	β	pentoxifilline	M06-2X	6-31g(d,p)	gas	Geo. opt., NBO, HOMO-LUMO	[101]
59	β	p-nitropenthyl acetate	B3LYP	6-31G(d,p)	n.i.p.	Geo. opt., interaction energy, NBO, HOMO-LUMO	[148]
60	β	norfloxacin	B97D (geo. opt.), B3LYP (SP, NMR)	6-31G(d,p)	IEFPCM	Geo. opt., interaction and stabilization energy, NMR, TD	[149]
			A1. Antidepressants				
61	β	paroxetine	B3LYP (geo. opt.); B97D (SP)	6-31+G* for H, N, O and 4-31G for C	vacuum	Geo. opt., interaction energy, TD	[109]
62	2,6-DM-β	mianserin	B3LYP-GD2 (geo. opt.); M05-GD3, M06-GD3, M062X-GD3, ωB97XD, mPW1PW91, M11 (SP)	6-31G(d,p)	PCM, vacuum	Geo. opt., interaction energy, NMR	[150]
63	β	sertraline HCl, fluoxetine HCl	B3LYP	6-31+G* for H, N, O and 4-31G for C	gas	Geo. opt., interaction energy	[151]
64	β	protriptyline, maprotiline	B3LYP	6-31+G* for H, N, O and 4-31G for C	vacuum	Geo. opt., interaction and stabilization energy	[152]
65	β	clomipramine, doxepin	B3LYP	31+G(d) for H, N, O, Cl, and 4-31G for C	gas	Geo. opt., interaction energy	[153]

Table 1. *Cont.*

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
66	β	desipramine, imipramine	B3LYP	6-31þG(d) for H, N, O and 4- 31G for C	gas, implicit solvent (water)	Geo. opt., interaction and stabilization energy	[108]
67	β	amitryptyline, nortryptiline	n.i.p.	6-31+G* for H, N, O, Cl and 4-31G for C	vacuum, SMD	Geo. opt., interaction and stabilization energy	[154]
			B. Plant derivatives				
68	HP-β	thymoquinone	B3LYP-D2, B3LYP-D3	6-31G(d,p)	PCM	Geo. opt., NBO, QTAIM, HOMO-LUMO, NMR	[155]
69	α	β-carotene	B3LYP	cc-pVDZ	vacuum	Geo. opt., interaction energy, Raman spectra	[156]
70	γ	3-hydroxyflavone	PBE0	def2-SV	PCM	Geo. opt., HOMO-LUMO, IT spectra	[157]
71	β	vanillina	B3LYP, ωB97xD, M06- 2X	6-311G(d,p)	vacuum, CPCM	Geo. opt., interaction energy, NMR, HOMO-LUMO, NBO, QTAIM, UV–Vis	[107]
72	β	alfa-terpineol	B3LYP (for UV–Vis), B3LYP/CAM, M062X, WB97-D3	6-311G(d,p)	vacuum, CPCM	Geo. opt., complexation energy, NBO, QTAIM, TD, UV-vis	[158]
73	TM-β, β	naringenin	B3LYP, M06-2X, wB97X-D	6-31G(d)	vacuum	Geo. opt., interaction energy, NBO, QTAIM, NMR, HOMO-LUMO	[159]
74	2,6-DMβ, 2HP-β, 2,6-DH-β, β	eucalyptol	M06-2X	6-31G(d,p)		Geo. opt., interaction energy	[160]
75	β	fisetin	M06-2X	6-31G(d,p)	gas, PCM	Geo. opt., interaction energy	[161]
76	β	gallic acid	B97-D3	6-31G*, for GIAO: 6-311++g**	gas, solvent	Geo. opt., HOMO-LUMO, NBO, NMR	[162]
77	β	gabapentin	B3LYP-D3	6- 31G(d)	vacuum, PCM	Geo. opt., interaction energy, NBO, HOMO-LUMO	[163]
78	B, γ	tropane alkaloids	B3LYP	6-31+G(d,p)	PCM	Geo. opt., interaction energy, NMR	[164]
79	β	coumarins	EDF2	6-311G(d,p)	PCM	Geo. opt.	[165]
80	2-HP-β	quercetin	B3LYP	6-31G*		Geo. opt.	[166]
81	β	carvacrol, thymol	B3LYP	6-31G, 6-31+G(d)	SMD	Geo. opt., interaction energy, NBO, HOMO-LUMO	[167]
82	β	thymol	B3LYP, PBEPBE, CAM-B3LYP	6-31G(d,p)	PCM	Geo. opt., interaction energy, UV–Vis	[168]
83	β	carvacrol	B3LYP, M05-2X	6-31G(d)	PCM	Geo. opt., HOMO-LUMO, NBO	[42]

Table 1. *Cont.*

No.	CD	Guest	Functional	Basis Set	Environment	DFT Application	Ref.
C. Functionalized food							
84	β	(−)-gallocatechin, (−)-catechin gallate, (−)-gallocatechin gallate	B3LYP	6-31+G* for H, O and 4-31G for C	gas	Geo. opt., interaction and stabilization energy	[102,103]
85	β	(−)-epigallocatechin, (−)-epigallocatechin gallate	B3PW91	cc-pVDZ	gas	Geo. opt., interaction energy	[104]
86	β	catechol derivatives: protocatechuic aldehyde, protocatechuic acid	B3LYP	6-31+G* for H, O and 4-31G for C	gas (geo. opt.), implicit solvent (TD)	Geo. opt., interaction energy, TD	[169]
87	β	oleuropein, hydroxytyrosol, tyrosol	n.i.p.	6-31+G* for H, O and 4-31G for C	gas	Geo. opt., interaction energy, TD	[170]
88	β	chlorogenic, caffeic, quinic acids	B3LYP	6-31+G* for H, O and 4-31G for C	gas	Geo. opt., interaction energy, TD	[105,106]
D. CD as a chiral selector							
89	β	D- and L-penicillamine	B3LYP-D3 (geo. opt.); M062X-D3, xB97X-D, B3LYP-D3 (interaction energy)	6-31G(d, p) (geo. opt.); G-311+G(d,p) (interaction energy)	water	Geo. opt., interaction energy	[171]
90	metal-ion coupled β	D- and L-penicillamine	DFT, M062X	6-31G(d,p)	vacuum	Geo. opt.	[172]
91	β	R- and S-propranolol	B3LYP	6-311+G(d,p)	vacuum	Geo. opt., vibrational spectra	[173]
92	per-M β	D- and L-isoleucine	B3LYP (geo. opt.), wB97X-D (IR)	6-31G*, 6-311G**	gas	Geo. opt., interaction energy, IR spectra, TD	[174]
93	per-M β	D- and L-alanine	B3LYP, wB97X-D, M06-2X	6-31G**, 6-311G**		Geo. opt., IR spectra	[175]
94	2,3,6-TM-β	cis-(2S,4R) and -(2R,4S) ketoconazole	B3LYP	6-311G(d,p)	gas (geo. opt.), PCM (SP)	Geo. opt., interaction energy	[36]
95	2-HP-β	abacavir enentiomers	PBE	6-31G*	PCM	Geo. opt., interaction energy	[176]

4. Conclusions

The number of studies concerning CDs complexes in which the theoretical calculations at the QC level have been used has constantly increased since the beginning of the 21st century. Solely for the DFT-based works in this topic, the number of published articles has already exceeded 300. While this number is still relatively low, when compared to the amount of reported molecular dynamics simulations at the molecular mechanics level [12], the reviewed works reveal that the application of QC calculations in the studies of CD complexes can be essential, providing the results unobtainable by any other method, both experimental and computational.

Initially, for those kind of studies, less computationally demanding semi-empirical methods have been applied (mostly PM3, PM6, and PM7). However, since 2015, each year there have been more papers in which DFT has been chosen instead of the semi-empirical approach. Nevertheless, even in some works published after 2020, the authors have found that the PM6 or PM7, when used with appropriate dispersion correction, can provide results with similar accuracy to those obtained using DFT.

Regarding calculations of geometries and interaction energies with DFT methods, in most of the reviewed works, the inclusion of dispersion correction was found to be crucial to obtain accurate energies, irrespective of the basis set and functional used. As for the functionals, there is no surprise that B3LYP, which is the most commonly applied one in the field of organic molecule calculations, is also the one that is most extensively used for the studies of CD complexes. However, in some works where the authors have used wB97X or M06-2X instead, different results have been obtained both in terms of predicted geometry as well as stability ranking.

When preparing the complexes, some authors prefer to manually dock the molecules, systematically moving the guest towards the CD cavity and rotating it; however, in most of the works, the molecular mechanics docking procedure has been applied, usually employing the popular and freeware Auto Dock. It should be noticed that the best pose from docking was not always the one with the lowest DFT energy; therefore, at least a few different poses should be optimized in order to achieve credible results.

Though in most of the reviewed works the authors have limited their calculations solely to geometry optimization of one or a few conformations, in some of the articles, the complex properties have been computed. Successful application of DFT methods include prediction of UV–Vis, IR, and NMR spectra as well as HOMO-LUMO and NBO calculations.

The major problem with the DFT calculations seems to be the solvent treatment. In the vast majority of the reviewed works, the authors decided to apply an implicit solvation model, usually PCM or SMD. However, other studies have shown that the role of water in the complex formation can be crucial as the water-mediated hydrogen bonds between the host and guest have been observed many times. Further, the role of water release from the cyclodextrin cavity during the complexation significantly affects the thermodynamics of such a process, which can be modeled accurately only using explicit water models.

Finally, it should be emphasized that even when carefully choosing the appropriate DFT method (applied functional, basis set, solvation scheme, dispersion correction, credible initial conformation), the obtained results can still be different from the corresponding experimental ones. This is due to the high flexibility and dynamics of most of CD complexes. Therefore, it seems reasonable to explore the application of ab initio molecular dynamics simulations. While, at this moment, it may be computationally not affordable for many researchers, combining the benefits of molecular dynamics simulations with the accuracy of DFT calculations seems to be the solution to obtain even more accurate results.

Author Contributions: Conceptualization, A.H.M. and Ł.S. writing—original draft preparation, A.H.M., Ł.S.; writing—review and editing, A.H.M., Ł.S.; supervision, Ł.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Aiassa, V.; Garnero, C.; Longhi, M.R.; Zoppi, A. Cyclodextrin Multicomponent Complexes: Pharmaceutical Applications. *Pharmaceutics* **2021**, *13*, 1099. [CrossRef] [PubMed]
2. Poulson, B.G.; Alsulami, Q.A.; Sharfalddin, A.; El Agammy, E.F.; Mouffouk, F.; Emwas, A.-H.; Jaremko, L.; Jaremko, M. Cyclodextrins: Structural, Chemical, and Physical Properties, and Applications. *Polysaccharides* **2022**, *3*, 1. [CrossRef]
3. Hădărugă, N.G.; Bandur, G.N.; David, I.; Hădărugă, D.I. A review on thermal analyses of cyclodextrins and cyclodextrin complexes. *Environ. Chem. Lett.* **2019**, *17*, 349–373. [CrossRef]
4. Szejtli, J.; Szente, L. Elimination of bitter, disgusting tastes of drugs and foods by cyclodextrins. *Eur. J. Pharm. Biopharm.* **2005**, *61*, 115–125. [CrossRef]
5. European Medicines Agency. Available online: https://www.ema.europa.eu/en (accessed on 29 May 2022).
6. U.S. Food & Drug Administration (FDA). Available online: http://www.fda.gov/cder/guidance/index.htm (accessed on 29 May 2022).
7. Pharmaceutical and Medical Devices Agencty. Available online: https://www.pmda.go.jp/english/index.html (accessed on 29 May 2022).
8. Tran, D.N.; Colesnic, D.; de Beaumais, S.A.; Pembouong, G.; Portier, F.; Queijo, A.; Tato, J.V.; Zhang, Y.; Ménand, M.; Bouteiller, L.; et al. Cyclodextrin-adamantane conjugates, self-inclusion and aggregation versus supramolecular polymer formation. *Org. Chem. Front.* **2014**, *1*, 703–706. [CrossRef]
9. Carrazana, J.; Jover, A.; Meijide, F.; Soto, A.V.H.; Tato, J.V. Complexation of Adamantyl Compounds by β-Cyclodextrin and Monoaminoderivatives. *J. Phys. Chem. B* **2005**, *109*, 9719–9726. [CrossRef]
10. Jug, M.; Mura, P.A. Grinding as Solvent-Free Green Chemistry Approach for Cyclodextrin Inclusion Complex Preparation in the Solid State. *Pharmaceutics* **2018**, *10*, 189. [CrossRef]
11. Wdowiak, K.; Rosiak, N.; Tykarska, E.; Żarowski, M.; Płazińska, A.; Płaziński, W.; Cielecka-Piontek, J. Amorphous Inclusion Complexes: Molecular Interactions of Hesperidin and Hesperetin with HP-B-CD and Their Biological Effects. *Int. J. Mol. Sci.* **2022**, *23*, 4000. [CrossRef]
12. Mazurek, A.H.; Szeleszczuk, Ł.; Gubica, T. Application of Molecular Dynamics Simulations in the Analysis of Cyclodextrin Complexes. *Int. J. Mol. Sci.* **2021**, *22*, 9422. [CrossRef]
13. Dehghani, A.; Bahlakeh, G.; Ramezanzadeh, B.; Mofidabadi, A.H.J.; Mostafatabar, A.H. Benzimidazole loaded β-cyclodextrin as a novel anti-corrosion system; Coupled experimental/computational assessments. *J. Colloid Interface Sci.* **2021**, *603*, 716–727. [CrossRef]
14. Roschi, E.; Gellini, C.; Ricci, M.; Sanchez-Cortes, S.; Focardi, C.; Neri, B.; Otero, J.C.; López-Tocón, I.; Smulevich, G.; Becucci, M. Surface-Enhanced Raman Spectroscopy for Bisphenols Detection: Toward a Better Understanding of the Analyte–Nanosystem Interactions. *Nanomaterials* **2021**, *11*, 881. [CrossRef] [PubMed]
15. Paulino, P.H.S.; Silva, C.F.; De Almeida, B.D.; Guimarães, L.; Nascimento, C.S. A theoretical study of poly(p-phenylenes) and their cyclodextrin-based insulated molecular wires. *Comput. Theor. Chem.* **2021**, *1197*, 113157. [CrossRef]
16. Dehghani, A.; Bahlakeh, G.; Ramezanzadeh, B. Synthesis of a non-hazardous/smart anti-corrosion nano-carrier based on beta-cyclodextrin-zinc acetylacetonate inclusion complex decorated graphene oxide (β-CD-ZnA-MGO). *J. Hazard. Mater.* **2020**, *398*, 122962. [CrossRef] [PubMed]
17. Tomer, A.; Kusema, B.T.; Paul, J.-F.; Przybylski, C.; Monflier, E.; Pera-Titus, M.; Ponchel, A. Cyclodextrin-assisted low-metal Ni-Pd/Al2O3 bimetallic catalysts for the direct amination of aliphatic alcohols. *J. Catal.* **2018**, *368*, 172–189. [CrossRef]
18. Quan, X.; Yi, S.; Wang, X. Theoretical study of an anti-Markovnikov addition reaction catalyzed by β-cyclodextrin. *J. Mol. Model.* **2018**, *24*, 77. [CrossRef]
19. Mazurek, A.H.; Szeleszczuk, Ł.; Pisklak, D.M. Periodic DFT Calculations—Review of Applications in the Pharmaceutical Sciences. *Pharmaceutics* **2020**, *12*, 415. [CrossRef]
20. Van De Streek, J.; Neumann, A.M. Validation of molecular crystal structures from powder diffraction data with dispersion-corrected density functional theory (DFT-D). *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2014**, *B70*, 1020–1032. [CrossRef]
21. Imtiaz, S.; Ali, S.M. Atom accurate structure determination of alprazolam/cyclodextrin inclusion complexes by ROESY and computational approaches. *J. Indian Chem. Soc.* **2022**, *99*, 100299. [CrossRef]
22. Lu, H.; Wang, Y.; Xie, X.; Chen, F.; Li, W. Molecular dynamics simulation and TDDFT study of the structures and UV–vis absorption spectra of MCT-β-CD and its inclusion complexes. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2015**, *149*, 564–570. [CrossRef]
23. Bouhadiba, A.; Rahali, S.; Belhocine, Y.; Allal, H.; Nouar, L.; Rahim, M. Structural and energetic investigation on the host/guest inclusion process of benzyl isothiocyanate into β-cyclodextrin using dispersion-corrected DFT calculations. *Carbohydr. Res.* **2020**, *491*, 107980. [CrossRef]

24. Wang, J.; Zhang, R.; Lu, Z.; Ai, Y. Experimental and theoretical studies of spherical β-cyclodextrin modified titanium dioxide composites for uranium removal. *Ecol. Eng.* **2020**, *149*, 105835. [CrossRef]
25. Jabeen, E.; Janjua, N.K.; Ahmed, S.; Murtaza, I.; Ali, T.; Masood, N.; Rizvi, A.S.; Murtaza, G. DFT predictions, synthesis, stoichiometric structures and anti-diabetic activity of Cu (II) and Fe (III) complexes of quercetin, morin, and primuletin. *J. Mol. Struct.* **2017**, *1150*, 459–468. [CrossRef]
26. Yuan, S.; Shi, W.; Li, B.; Wang, J.; Jiao, H.; Li, Y.-W. Theoretical ONIOM2 Study on Pyridine Adsorption in the Channels and Intersection of ZSM-5. *J. Phys. Chem. A* **2005**, *109*, 2594–2601. [CrossRef]
27. Christensen, A.; Kubař, T.; Cui, Q.; Elstner, M. Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications. *Chem. Rev.* **2016**, *116*, 5301–5337. [CrossRef]
28. Ferrero, R.; Pantaleone, S.; Delle Piane, M.; Caldera, F.; Corno, M.; Trotta, F.; Brunella, V. On the Interactions of Melatonin/β-Cyclodextrin Inclusion Complex: A Novel Approach Combining Efficient Semiempirical Extended Tight-Binding (xTB) Results with Ab Initio Methods. *Molecules* **2021**, *26*, 5881. [CrossRef] [PubMed]
29. Vidal, R.B.P.; Ibañez, G.A.; Escandar, G.M. Spectrofluorimetric study of phenolic endocrine disruptors in cyclodextrin media. *RSC Adv.* **2015**, *5*, 20914–20923. [CrossRef]
30. Li, X.-S.; Liu, L.; Mu, T.-W.; Guo, Q.-X. A Systematic Quantum Chemistry Study on Cyclodextrins. *Mon. Chem.* **2000**, *131*, 849–855. [CrossRef]
31. Fatiha, M.; Khatmi, D.E.; Largate, L. Theoretical approach in the study of the inclusion processes of sulconazole with β-cyclodextrin. *J. Mol. Liq.* **2010**, *154*, 1–5. [CrossRef]
32. Setiadji, S.; Sundari, C.D.D.; Ramdhani, M.A.; Umam, A.B.K.; Ivansyah, A.L. Theoretical Investigation of Inclusion Complex between Omeprazole Enantiomers and Carboxymethyl-β-Cyclodextrin. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *288*, 012138. [CrossRef]
33. Saroj, M.K.; Payal, R.; Jain, S.; Sharma, N.; Rastogi, R.C. Investigation of indole chalcones encapsulation in β-cyclodextrin: Determination of stoichiometry, binding constants and thermodynamic parameters. *J. Incl. Phenom. Macrocycl. Chem.* **2018**, *90*, 306–320. [CrossRef]
34. Nurhidayah, E.S.; Martoprawiro, M.; Zulfikar, M.A. PM3 and ONIOM2 modelling of inclusion complex of ibuprofen enantiomers with dimethyl-β-Cyclodextrin. *J. Chem. Tech. Metall.* **2019**, *54*, 673–678.
35. Srihakulung, O.; Triamchaisri, N.; Toochinda, P.; Lawtrakul, L. Theoretical study on ferrocenyl hydrazones inclusion complexes with β-cyclodextrin and its three methylated derivatives. *J. Incl. Phenom. Macrocycl. Chem.* **2020**, *98*, 79–91. [CrossRef]
36. Arsad, S.R.; Maarof, H.; Ibrahim, W.A.W.; Aboul-Enein, H.Y. Theoretical and Molecular Docking Study of Ketoconazole on Heptakis(2,3,6-tri-*O* -methyl)-β-cyclodextrin as Chiral Selector. *Chirality* **2016**, *28*, 209–214. [CrossRef] [PubMed]
37. Bensouilah, N.; Fisli, H.; Bensouilah, H.; Zaater, G.; Abdaoui, M.; Boutemeur-Kheddis, B. Host-guest complex of N-(2-chloroethyl), N-nitroso,N′,N′-dicyclohexylsulfamid with β-cyclodextrin: Fluorescence, QTAIM analysis and structure-chemical reactivity. *J. Mol. Struct.* **2017**, *1146*, 179–190. [CrossRef]
38. Ivansyah, A.L.; Martoprawiro, M. Buchari Computational modeling of inclusion complex of r/s-omeprazole with β-cyclodextrin using oniom2 method. *J. Phys. Conf. Ser.* **2017**, *812*, 012070. [CrossRef]
39. Yang, Z.; Huang, L.; Yao, X.; Ji, H. Host-guest complexes of estragole with β-cyclodextrin: An experimental and theoretical investigation. *Flavour Fragr. J.* **2017**, *32*, 102–111. [CrossRef]
40. Bensouilah, N.; Boutemeur-Kheddis, B.; Meddour, I.; Abdaoui, M. Host-guest complex of nabumetone: β-cyclodextrin: Quantum chemical study and QTAIM analysis. *J. Incl. Phenom. Macrocycl. Chem.* **2017**, *87*, 191–206. [CrossRef]
41. Reis, V.S.; Santos, E.S.; Bonsolhos, D.N.F.; Guimarães, L.; De Almeida, W.G.; Nascimento, C.S. Theoretical study on the formation process of Cross-Linked β-Cyclodextrin molecular tubes. *Chem. Phys. Lett.* **2017**, *677*, 13–18. [CrossRef]
42. Abdelmalek, L.; Fatiha, M.; Leila, N.; Mouna, C.; Nora, M.; Djameleddine, K. Computational study of inclusion complex formation between carvacrol and β-cyclodextrin in vacuum and in water: Charge transfer, electronic transitions and NBO analysis. *J. Mol. Liq.* **2016**, *224*, 62–71. [CrossRef]
43. Al Azzam, K.M.; Muhammad, E. Host-guest Inclusion Complexes between Mitiglinide and the Naturally Occurring Cyclodextrins α, β, and γ: A Theoretical Approach. *Adv. Pharm. Bull.* **2015**, *5*, 289–291. [CrossRef]
44. Yang, Z.; Yao, X.; Xiao, Z.; Chen, H.; Ji, H. Preparation and release behaviour of the inclusion complexes of phenylethanol withβ-cyclodextrin. *Flavour Fragr. J.* **2016**, *31*, 206–216. [CrossRef]
45. Nazarov, V.B.; Avakyan, V.G.; Bagrii, E.I.; Vershinnikova, T.G.; Alfimov, M.V. Long-lived phosphorescence of arenes in complexes with cyclodextrins 2. Room-temperature phosphorescence of ternary complexes of naphthalene and phenanthrene with β-cyclodextrin and adamantane derivatives in the presence of oxygen. *Russ. Chem. Bull.* **2005**, *54*, 2752–2756. [CrossRef]
46. Bouzit, H.; Stiti, M.; Abdaoui, M. Spectroscopic and molecular modelling investigations of supramolecular complex of β-cyclodextrin with N-[(4-sulfonamidophenyl)ethyl]-5-(1,2-dithiolan-3-yl)pentanamide. *J. Incl. Phenom. Macrocycl. Chem.* **2016**, *86*, 121–134. [CrossRef]
47. Laspidou, C.S.; Archimandritis, A.S.; Papadimitriou, T.; Kormas, K.A.; Yannakopoulou, K.; Lazarou, Y.G. Theoretical investigation of microcystin-LR, microcystin-RR and nodularin-R complexation with α-, β-, and γ-cyclodextrin as a starting point for the targeted design of efficient cyanotoxin traps. *Sustain. Chem. Pharm.* **2016**, *3*, 25–32. [CrossRef]

48. Suliman, F.O.; Elbashir, A.A.; Schmitz, O.J. Study on the separation of ofloxacin enantiomers by hydroxyl-propyl-β-cyclodextrin as a chiral selector in capillary electrophoresis: A computational approach. *J. Incl. Phenom. Macrocycl. Chem.* **2015**, *83*, 119–129. [CrossRef]
49. Mizera, M.; Lewandowska, K.; Miklaszewski, A.; Cielecka-Piontek, J. Machine Learning Approach for Determining the Formation of β-Lactam Antibiotic Complexes with Cyclodextrins Using Multispectral Analysis. *Molecules* **2019**, *24*, 743. [CrossRef] [PubMed]
50. Djilani, I.; Madi, F.; Nouar, L.; Haiahem, S.; Rahim, M.; Khatmi, D.E.; Bouhadiba, A. Theoretical investigation to characterize the inclusion complex of α-lipoic acid and β-cyclodextrin. *Comptes Rendus Chim.* **2015**, *18*, 170–177. [CrossRef]
51. Bouhadiba, A.; Belhocine, Y.; Rahim, M.; Djilani, I.; Nouar, L.; Khatmi, D.E. Host-guest interaction between tyrosine and β-cyclodextrin: Molecular modeling and nuclear studies. *J. Mol. Liq.* **2017**, *233*, 358–363. [CrossRef]
52. Sambrook, M.R.; Vincent, J.C.; Ede, J.A.; Gass, I.A.; Cragg, P.I. Experimental and computational study of the inclusion complexes of β-cyclodextrin with the chemical warfare agent soman (GD) and commonly used simulants. *RSC Adv.* **2017**, *7*, 38069–38076. [CrossRef]
53. Yahia, H.A.; Yahia, O.A.; Khatmi, D.; Belghiche, R.; Bouzitouna, A. Quantum chemical investigations on hydrogen bonding interactions established in the inclusion complex β-cyclodextrin/benzocaine through the DFT, AIM and NBO approaches. *J. Incl. Phenom. Macrocycl. Chem.* **2017**, *89*, 353–365. [CrossRef]
54. Ceborska, M.; Kędra-Królik, K.; Kowalska, A.A.; Koźbiał, M. Comparative study of molecular recognition of folic acid subunits with cyclodextrins. *Carbohydr. Polym.* **2018**, *184*, 47–56. [CrossRef] [PubMed]
55. Appell, M.; Evans, K.O.; Jackson, M.A.; Compton, D.L. Determination of ochratoxin A in grape juice and wine using nanosponge solid phase extraction clean-up and liquid chromatography with fluorescence detection. *J. Liq. Chromatogr. Relat. Technol.* **2018**, *41*, 949–954. [CrossRef]
56. López-Méndez, L.J.; Rojas-Aguirre, Y.; Vázquez-Lima, H.; Cassani, J.; Enríquez, R.G.; Rojo-Domínguez, A.; Guadarrama, P. On the conformational search of a βCD dendritic derivative: NMR and theoretical calculations working together reveal a donut-like amphiphilic structure. *J. Mol. Struct.* **2020**, *1204*, 127535. [CrossRef]
57. Sahra, K.; Dinar, K.; Seridi, A.; Kadri, M. Investigation on the inclusion of diclofenac with β-cyclodextrin: A molecular modeling approach. *Struct. Chem.* **2015**, *26*, 61–69. [CrossRef]
58. Srihakulung, O.; Maezono, R.; Toochinda, P.; Kongprawechnon, W.; Intarapanich, A.; Lawtrakul, A.L. Host-Guest Interactions of Plumbagin with β-Cyclodextrin, Dimethyl-β-Cyclodextrin and Hydroxypropyl-β-Cyclodextrin: Semi-Empirical Quantum Mechanical PM6 and PM7 Methods. *Sci. Pharm.* **2018**, *86*, 20. [CrossRef]
59. Bani-Yaseen, A.W. Computational molecular perspectives on the interaction of propranolol with β-cyclodextrin in solution: Towards the drug-receptor mechanism of interaction. *J. Mol. Liq.* **2017**, *227*, 280–290. [CrossRef]
60. Prabhu, A.A.M.; Fatiha, M.; Leila, N.; Raj, T.A.; Navarro-González, I.; Periago, M.J.; Yáñez-Gascón, M.J.; Pérez-Sánchez, H. Investigation of 3D Contour Map and Intermolecular Interaction of Dopamine with β-Cyclodextrin and 2-Hydroxypropyl-β-cyclodextrin. *J. Solut. Chem.* **2018**, *47*, 409–429. [CrossRef]
61. Iyengar, S.S.; Schlegel, B.H.; Voth, G.A. Atom-Centered Density Matrix Propagation (ADMP): Generalizations Using Bohmian Mechanics. *J. Phys. Chem. A* **2003**, *107*, 7269–7277. [CrossRef]
62. Al-Jaber, A.S.; Bani-Yaseen, A.D. On the encapsulation of Olsalazine by β-cyclodextrin: A DFT-based computational and spectroscopic investigations. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2019**, *214*, 531–536. [CrossRef]
63. Silva, D.A.; Xavier, M.J.; Dutra, J.D.L.; Gimenez, I.F.; Freire, R.O.; da Costa, N.B. Prediction of correct intermolecular interactions in host-guest systems involving cyclodextrins. *J. Mol. Struct.* **2020**, *1205*, 127517. [CrossRef]
64. Lula, I.; Gomes, M.F.; Piló-Veloso, D.; Noronha, A.L.O.; Duarte, H.A.; Santos, R.A.S.; Sinisterra, R.D. Spironolactone and its Complexes with β-cyclodextrin: Modern NMR Characterization and Structural DFTB-SCC Calculations. *J. Incl. Phenom. Macrocycl. Chem.* **2006**, *56*, 293–302. [CrossRef]
65. Lula, I.; Denadai, L.; Resende, J.M.; de Sousa, F.B.; de Lima, G.F.; Pilo-Veloso, D.; Heine, T.; Duarte, H.A.; Santos, R.A.; Sinisterra, R.D. Study of angiotensin-(1–7) vasoactive peptide and its β-cyclodextrin inclusion complexes: Complete sequence-specific NMR assignments and structural studies. *Peptides* **2007**, *28*, 2199–2210. [CrossRef] [PubMed]
66. Lukin, O.; Dolgonos, G.; Leszczynski, J. A comprehensive test of computational approaches for evaluation of cyclodextrin complexes. Self-inclusion in monosubstituted β-cyclodextrins—A case study. *Tetrahedron* **2017**, *73*, 5302–5306. [CrossRef]
67. Schnupf, U.; Momany, F.A. DFT Energy Optimization of a Large Carbohydrate: Cyclomaltohexaicosaose (CA-26). *J. Phys. Chem. B* **2011**, *116*, 6618–6627. [CrossRef] [PubMed]
68. Hanpaibool, C.; Chakcharoensap, T.; Hijikata, Y.; Irle, S.; Wolschann, P.; Kungwan, N.; Pongsawasdi, P.; Ounjai, P.; Rungrotmongkol, T. Theoretical analysis of orientations and tautomerization of genistein in β-cyclodextrin. *J. Mol. Liq.* **2018**, *265*, 16–23. [CrossRef]
69. Yunta, M. Using Molecular Modelling to Study Interactions between Molecules with Biological Activity. *IntechOpen* **2012**. [CrossRef]
70. Lu, L. Can B3LYP be improved by optimization of the proportions of exchange and correlation functionals? *Int. J. Quantum Chem.* **2015**, *115*, 502–509. [CrossRef]
71. Ivanov, P. Performance of some DFT functionals with dispersion on modeling of the translational isomers of a solvent-switchable [2]rotaxane. *J. Mol. Struct.* **2016**, *1107*, 31–38. [CrossRef]

72. Mardirossian, N.; Head-Gordon, M. How Accurate Are the Minnesota Density Functionals for Noncovalent Interactions, Isomerization Energies, Thermochemistry, and Barrier Heights Involving Molecules Composed of Main-Group Elements? *J. Chem. Theory Comput.* **2016**, *12*, 4303–4325. [CrossRef]
73. Valero, R.; Costa, R.; Moreira, I.D.P.R.; Truhlar, D.; Illas, F. Performance of the M06 family of exchange-correlation functionals for predicting magnetic coupling in organic and inorganic molecules. *J. Chem. Phys.* **2008**, *128*, 114103. [CrossRef]
74. Wang, Y.; Jin, X.; Yu, H.S.; Truhlar, D.G.; He, X. Revised M06-L functional for improved accuracy on chemical reaction barrier heights, noncovalent interactions, and solid-state physics. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 8487–8492. [CrossRef] [PubMed]
75. Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924. [CrossRef] [PubMed]
76. Kohn, W.; Sham, L.J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133–A1138. [CrossRef]
77. Bučko, T.; Lebègue, S.; Hafner, J.; Ángyán, J.G. Tkatchenko-Scheffler van der Waals correction method with and without self-consistent screening applied to solids. *Phys. Rev. B* **2013**, *87*, 064110. [CrossRef]
78. Grimme, S.; Steinmetz, M. Effects of London dispersion correction in density functional theory on the structures of organic molecules in the gas phase. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16031–16042. [CrossRef] [PubMed]
79. Xu, P.; Alkan, M.; Gordon, M.S. Many-Body Dispersion. *Chem. Rev.* **2020**, *120*, 12343–12356. [CrossRef]
80. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465. [CrossRef]
81. Yanai, T.; Tew, D.P.; Handy, N.C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57. [CrossRef]
82. Salma, A.; Fatiha, M.; Leila, N. Effect of solvent on absorption and emission spectra of 2,2′-Bipyridine and its inclusion complexinto β-cyclodextrin: DFT and TD-DFT study. *Comput. Theor. Chem.* **2021**, *1206*, 113481. [CrossRef]
83. Azayez, M.; Fergoug, T.; Meddah-Araibi, N.; Zelmat, C.; Bouhadda, Y. Theoretical Investigation of the Complexation Reaction of Procaine-hydrochloride by β-cyclodextrin. *Phys. Chem. Res.* **2020**, *8*, 155–165. [CrossRef]
84. Safia, H.; Ismahan, L.; Abdelkrim, G.; Mouna, C.; Leila, N.; Fatiha, M. Density functional theories study of the interactions between host β-Cyclodextrin and guest 8-Anilinonaphthalene-1-sulfonate: Molecular structure, HOMO, LUMO, NBO, QTAIM and NMR analyses. *J. Mol. Liq.* **2019**, *280*, 218–229. [CrossRef]
85. Buczek, A.; Staś, A.; Hebenstreit, C.; Maller, C.; Broda, M.A.; Kupka, T.; Kelterer, A. Interaction of 5-fluorouracil with β-cyclodextrin: A density functional theory study with dispersion correction. *Int. J. Quantum Chem.* **2021**, *121*, e26487. [CrossRef]
86. Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* **2003**, *24*, 669–681. [CrossRef] [PubMed]
87. Tomasi, J.; Mennucci, B.; Cancès, E. The IEF version of the PCM solvation method: An overview of a new method addressed to study molecular solutes at the QM ab initio level. *J. Mol. Struct.* **1999**, *464*, 211–226. [CrossRef]
88. Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396. [CrossRef]
89. Belhocine, Y.; Rahali, S.; Allal, H.; Assaba, I.M.; Ghoniem, M.G.; Ali, F.A.M. A Dispersion Corrected DFT Investigation of the Inclusion Complexation of Dexamethasone with β-Cyclodextrin and Molecular Docking Study of Its Potential Activity against COVID-19. *Molecules* **2021**, *26*, 7622. [CrossRef]
90. Oqmhula, K.; Hongo, K.; Maezono, R.; Ichibha, T. Ab Initio Evaluation of Complexation Energies for Cyclodextrin-Drug Inclusion Complexes. *ACS Omega* **2020**, *5*, 19371–19376. [CrossRef]
91. Sambrook, M.R.; Gass, I.A.; Cragg, P.J. Spectroscopic and inclusion properties of G-series chemical warfare agents and their simulants: A DFT study. *Supramol. Chem.* **2017**, *30*, 206–217. [CrossRef]
92. Sure, R.; Grimme, S. Comprehensive Benchmark of Association (Free) Energies of Realistic Host–Guest Complexes. *J. Chem. Theory Comput.* **2015**, *11*, 3785–3801. [CrossRef]
93. Lopes, J.F.; Nascimento, C.S.; Anconi, C.P.A.; Santos, H.F.; Almeida, W.B. Inclusion complex thermodynamics: The β-cyclodextrin and sertraline complex example. *J. Mol. Graph. Model.* **2015**, *62*, 11–17. [CrossRef]
94. CCDC, Cambridge Crystallographic Data Centre. Available online: https://www.ccdc.cam.ac.uk/ (accessed on 29 May 2022).
95. Autodock. Available online: https://autodock.scripps.edu/ (accessed on 29 May 2022).
96. Available online: https://www.schrodinger.com/products/maestro (accessed on 29 May 2022).
97. Biovia. Available online: https://www.3ds.com/ (accessed on 29 May 2022).
98. Nora, M.; Ismahan, L.; Abdelkrim, G.; Mouna, C.; Leila, N.; Fatiha, M.; Nada, B.; Brahim, H. Interactions in inclusion complex of β-cyclodextrin/l-Metheonine: DFT computational studies. *J. Incl. Phenom. Macrocycl. Chem.* **2020**, *96*, 43–54. [CrossRef]
99. Pan, A.; Kar, T.; Rakshit, A.K.; Moulik, S.P. Enthalpy–Entropy Compensation (EEC) Effect: Decisive Role of Free Energy. *J. Phys. Chem. B* **2016**, *120*, 10531–10539. [CrossRef] [PubMed]
100. Rahali, S.; Belhocine, Y.; Allal, H.; Bouhadiba, A.; Assaba, I.M.; Seydou, M. A DFT Investigation of The Host-Guest Interactions Between Boron-Based Aromatic Systems and β-Cyclodextrin. *Res. Sq.* **2022**, 1–2. [CrossRef]

101. Morais, C.A.S.; Silva, B.L.; Denadai, A.M.L.; Lopes, J.F.; De Sousa, F.B. Structural and thermodynamic investigation of pentoxifylline-cyclodextrin inclusion complex. *Chem. Phys. Lett.* **2017**, *682*, 43–48. [CrossRef]
102. Aree, T.; Jongrungruangchok, S. β-Cyclodextrin encapsulation elevates antioxidant capacity of tea: A closing chapter on non-epicatechins, atomistic insights from X-ray analysis, DFT calculation and DPPH assay. *Carbohydr. Polym.* **2018**, *194*, 24–33. [CrossRef] [PubMed]
103. Aree, T.; Jongrungruangchok, S. Enhancement of antioxidant activity of green tea epicatechins in β-cyclodextrin cavity: Single-crystal X-ray analysis, DFT calculation and DPPH assay. *Carbohydr. Polym.* **2016**, *20*, 1139–1151. [CrossRef] [PubMed]
104. Ikeda, H.; Ohata, T.; Yukawa, M.; Tsutsumi, H.; Fujisawa, M.; Aki, H. Calculation study on complex formation of catechins with β-cyclodextrin using density function theory. *J. Incl. Phenom. Macrocycl. Chem.* **2021**, *100*, 99–107. [CrossRef]
105. Aree, T. Understanding structures and thermodynamics of β-cyclodextrin encapsulation of chlorogenic, caffeic and quinic acids: Implications for enriching antioxidant capacity and masking bitterness in coffee. *Food Chem.* **2019**, *293*, 550–560. [CrossRef]
106. Aree, T. Inclusion complex of β-cyclodextrin with coffee chlorogenic acid: New insights from a combined crystallographic and theoretical study. *Acta Crystallogr. Sect. C Struct. Chem.* **2019**, *75*, 15–21. [CrossRef]
107. Meryem, G.; Rabah, K.; Fatiha, M.; Leila, N.; Aziz, B.A.; Imane, D.; Rachid, M. Computational investigation of vanillin@βéta-cyclodextrin inclusion complex: Electronic and intermolecular analysis. *J. Mol. Liq.* **2021**, *321*, 114839. [CrossRef]
108. Aree, T. β-Cyclodextrin Inclusion Complexation With Tricyclic Antidepressants Desipramine and Imipramine: A Structural Chemistry Perspective. *J. Pharm. Sci.* **2020**, *109*, 3086–3094. [CrossRef] [PubMed]
109. Aree, T. Inclusion Scenarios and Conformational Flexibility of the SSRI Paroxetine as Perceived from Polymorphism of β-Cyclodextrin–Paroxetine Complex. *Pharmaceuticals* **2022**, *15*, 98. [CrossRef] [PubMed]
110. Keniche, A.; Slimani, M.Z.; Miranda, J.I.; Aizpurua, J.M.; Mulengi, J.K. NMR Investigation of the complexation of (S)-2-isopropyl-1-(o-nitrophenyl)sulfonyl)aziridine with β-cyclodextrin. *Mediterr. J. Chem.* **2013**, *2*, 620–631. [CrossRef]
111. Harati, H.; Morsali, A.; Bozorgmehr, M.R.; Beyramabadi, S.A. β-cyclodextrin-lenalidomide anticancer drug delivery nanosystem: A quantum chemical approach. *J. Mol. Liq.* **2021**, *344*, 117762. [CrossRef]
112. Rohman, M.A.; Phanrang, P.T.; Chamlagai, D.; Mitra, S. Deciphering Spectroscopic and Structural Insights into the Photophysical Behavior of 2,2′-Dipyridylamine: An Efficient Environment Sensitive Fluorescence Probe. *J. Phys. Chem. A* **2021**, *125*, 6964–6975. [CrossRef]
113. Wiergowska, G.; Ludowicz, D.; Wdowiak, K.; Miklaszewski, A.; Lewandowska, K.; Cielecka-Piontek, J. Combinations of Freeze-Dried Amorphous Vardenafil Hydrochloride with Saccharides as a Way to Enhance Dissolution Rate and Permeability. *Pharmaceuticals* **2021**, *14*, 453. [CrossRef]
114. Akhondi, M.; Jamalizadeh, E.; Mohebbi, A. MD and DFT calculations on the structural variations of amino-cyclodextrin as a pH-sensitive carrier for smart carriage and release of Doxorubicin. *J. Mol. Struct.* **2021**, *1230*, 129855. [CrossRef]
115. Mezari, Y.; Nouar, L.; Madi, F.; Guendouzi, A.; Djellala, I.; Lafifi, I.; Merdes, R.; Bouhadiba, A.; Houari, B. Theoretical investigation of inclusion complex of 2-methyl mercapto phenothiazine with hydroxy propyl β-cyclodextrin by DFT approaches. *Bulg. Chem. Comm.* **2021**, *53*, 196–210. [CrossRef]
116. Bani-Yaseen, A.D. The supramolecular host-guest complexation of Vemurafenib with β-cyclodextrin and cucurbit[7]uril as drug photoprotecting systems: A DFT/TD-DFT study. *Comput. Theor. Chem.* **2020**, *1191*, 113026. [CrossRef]
117. Jafari, G.; Raissi, H.; Hashemzadeh, H. Molecular insight into the interaction of fluorometholone and cholesterol molecules with β-cyclodextrin and sulfobutylether-β-cyclodextrin. *Comput. Theor. Chem.* **2022**, *1208*, 113554. [CrossRef]
118. Gamboa-Carballo, J.J.; Ferino-Pérez, A.; Rana, V.K.; Levalois-Grützmacher, J.; Gaspard, S.; Montero-Cabrera, L.A.; Haza, U.J.J. Theoretical Evaluation of the Molecular Inclusion Process between Chlordecone and Cyclodextrins: A New Method for Mitigating the Basis Set Superposition Error in the Case of an Implicit Solvation Model. *J. Chem. Inf. Model.* **2020**, *60*, 2115–2125. [CrossRef] [PubMed]
119. Yang, Y.-Z.; Liu, X.-F.; Zhang, R.-B.; Pang, S.-P. Joint experimental and theoretical studies of the surprising stability of the aryl pentazole upon noncovalent binding to β-cyclodextrin. *Phys. Chem. Chem. Phys.* **2017**, *19*, 31236–31244. [CrossRef] [PubMed]
120. Pereira, R.A.; Borges, W.M.D.S.; Peraro, C.R.; Anconi, C.P.A. Theoretical inclusion of deprotonated 2,4-D and dicamba pesticides in ß-cyclodextrin. *J. Incl. Phenom. Macrocycl. Chem.* **2016**, *86*, 343–349. [CrossRef]
121. Li, Z.; Couzijn, E.P.A.; Zhang, X. Intrinsic Properties of α-Cyclodextrin Complexes with Benzoate Derivatives in the Gas Phase: An Experimental and Theoretical Study. *J. Phys. Chem. B* **2012**, *116*, 943–950. [CrossRef] [PubMed]
122. Yao, L.; Mori, Y.; Takano, K. Theoretical Study on Intermolecular Interactions in Complexes of Cyclodextrins with Bile Acids: DFT and Ab Initio Fragment Molecular Orbital Calculations. *Bull. Chem. Soc. Jpn.* **2014**, *87*, 258–266. [CrossRef]
123. Li, Z.; Couzijn, E.P.A.; Zhang, X. A quantitative study of intrinsic non-covalent interactions within complexes of α-cyclodextrin and benzoate derivatives. *Chem. Commun.* **2012**, *48*, 9864–9866. [CrossRef]
124. Muzaffar, S.; Imtiaz, S.; Ali, S.M. Demonstrating accuracy of the proposed protocol for structure elucidation of cyclodextrin inclusion complexes by validation using DFT studies. *J. Mol. Struct.* **2020**, *1217*, 128419. [CrossRef]
125. Li, N.; Yang, L.; Ji, X.; Ren, J.; Gao, B.; Deng, W.-Q.; Wang, Z. Bioinspired succinyl-β-cyclodextrin membranes for enhanced uranium extraction and reclamation. *Environ. Sci. Nano* **2020**, *7*, 3124–3135. [CrossRef]
126. Ignaczak, A.; Orszański, Ł.; Adamiak, M.; Olejniczak, A.B. Comparative DFT study of inclusion complexes of thymidine-carborane conjugate with β-cyclodextrin and heptakis(2,6-O-dimethyl)-β-cyclodextrin in water. *J. Mol. Liq.* **2020**, *315*, 113767. [CrossRef]

127. Ramos, M.L.; Dias, D.C.; Justino, L.L.G.; Verissimo, L.M.P.; Valente, A.J.M.; Esteso, M.A.; Ribeiro, A.C.F.; Leaist, D.G.; Pina, J.; Cabral, A.; et al. Interactions between glycyl-L-phenylalanine and β-cyclodextrin from diffusion, spectroscopic and computational studies. *J. Mol. Liq.* **2020**, *315*, 113704. [CrossRef]
128. Deosarkar, S.D.; Sawale, R.T.; Pinjari, R.V.; Kalyankar, T.M. Interactions of sodium salicylate and β-cyclodextrin in water: A volumetric, ultraacoustic and optical study. *J. Mol. Liq.* **2020**, *130*, 113151. [CrossRef]
129. Okuda, M.; Hiramatsu, T.; Yasuda, M.; Ishigaki, M.; Ozaki, Y.; Hayashi, M.; Tominaga, K.; Chatani, E. Theoretical Modeling of Electronic Structures of Polyiodide Species Included in α-Cyclodextrin. *J. Phys. Chem. B* **2020**, *124*, 4089–4096. [CrossRef] [PubMed]
130. Majhi, K.; Bandyopadhyay, P.; Khatun, R.; Sinha, S. Prediction of the most preferable rotamer of meta-aminophenol in β-cyclodextrin cavity in aqueous medium by using spectroscopic and DFT computational studies. *J. Incl. Phenom. Macrocycl. Chem.* **2020**, *97*, 77–86. [CrossRef]
131. Ismahan, L.; Leila, N.; Fatiha, M.; Abdelkrim, G.; Mouna, C.; Nada, B.; Brahim, H. Computational study of inclusion complex of l-Glutamine/beta-Cycldextrin: Electronic and intermolecular interactions investigations. *J. Mol. Struct.* **2020**, *1206*, 127740. [CrossRef]
132. Pereva, S.; Nikolova, V.; Sarafska, T.; Angelova, S.; Spassov, T.; Dudev, T. Inclusion complexes of ibuprofen and β-cyclodextrin: Supramolecular structure and stability. *J. Mol. Struct.* **2020**, *1205*, 127575. [CrossRef]
133. Stoicescu, C.S.; Neacşu, A.D.; Bădiceanu, C.D.; Munteanu, G. Inclusion complexes of some thiourea derivatives in cyclodextrins. *J. Incl. Phenom. Macrocycl. Chem.* **2019**, *96*, 275–283. [CrossRef]
134. Paulino, P.H.S.; Sousaab, S.M.R.; Da Silva, H.C.; De Almeida, W.B.; Ferrari, J.L.; Guimarãesa, L.; Nascimento, C.S., Jr. A theoretical investigation on the encapsulation process of mepivacaine into β-cyclodextrin. *Chem. Phys. Lett.* **2020**, *740*, 137060. [CrossRef]
135. Wu, J.; Ma, H.; Bu, X.; Zhu, L.; Hao, B.; Zhao, B.; Tian, Y. SERS determination of the antihypertensive drugs prazosin and losartan by using silver nanoparticles coated with β-cyclodextrin. *Mikrochim. Acta* **2019**, *186*, 801. [CrossRef]
136. Bezzina, B.; Djemil, R.; Bensouilah, N. Quantitative and qualitative analyses of intermolecular interactions in neutral/deprotonated aspirin@β-CD inclusion complexes: QTAIM and NBO analyses. *Theor. Chim. Acta* **2019**, *138*, 43. [CrossRef]
137. Wójcik, J.; Ejchart, A.; Nowakowski, M. Shape adaptation of quinine in cyclodextrin cavities: NMR studies. *Phys. Chem. Chem. Phys.* **2019**, *21*, 6925–6934. [CrossRef]
138. Bakirhan, N.K.; Tok, T.T.; Ozkan, S.A. The redox mechanism investigation of non-small cell lung cancer drug: Erlotinib via theoretical and experimental techniques and its host–guest detection by β-Cyclodextrin nanoparticles modified glassy carbon electrode. *Sens. Actuators B Chem.* **2019**, *278*, 172–180. [CrossRef]
139. Li, L.; Zhou, Y.; Wang, Z.; Wu, C.; Li, Z.; Sun, C.; Sun, T. Theoretical studies on the mechanism of sugammadex for the reversal of aminosteroid-induced neuromuscular blockade. *J. Mol. Liq.* **2018**, *265*, 450–456. [CrossRef]
140. Khavani, M.; Kalantarinezhad, R.; Izadyar, M. A joint QM/MD study on α-, β- and γ-cyclodextrins in selective complexation with cathinone. *Supramol. Chem.* **2017**, *30*, 687–696. [CrossRef]
141. Eftaiha, A.F.; Qaroush, A.K.; Alsoubani, F.; Pehl, T.M.; Troll, C.; Rieger, B.; Al-Maythalony, B.A.; Assaf, K.I. A green sorbent for CO_2 capture: α-cyclodextrin-based carbonate in DMSO solution. *RSC Adv.* **2018**, *8*, 37757–37764. [CrossRef]
142. Belhocine, Y.; Bouhadiba, A.; Rahim, M.; Nouar, L.; Djilani, I.; Khatmi, D.I. Inclusion Complex Formation of β-Cyclodextrin with the Nonsteroidal Anti-inflammatory Drug Flufenamic Acid: Computational Study. *Macroheterocycles* **2018**, *11*, 203–209. [CrossRef]
143. Sierpe, R.; Noyong, M.; Simon, U.; Aguayo, D.; Huerta, J.; Kogan, M.J.; Yutronic, N. Construction of 6-thioguanine and 6-mercaptopurine carriers based on βcyclodextrins and gold nanoparticles. *Carbohydr. Polym.* **2017**, *177*, 22–31. [CrossRef] [PubMed]
144. Anconi, C.P.A.; Santos, T.M.R.; Souza, A.C.; Borges, W.M.S.; Sales, A.L.R. Host–guest intermolecular hydrogen bonds and stability in aqueous media: The benzaldehyde/β-CD case study. *J. Incl. Phenom. Macrocycl. Chem.* **2017**, *89*, 137–142. [CrossRef]
145. Cao, B.; Du, J.; Cao, Z.; Sun, X.; Sun, H.; Fu, H. DFT study on the dissolution mechanisms of α-cyclodextrin and chitobiose in ionic liquid. *Carbohydr. Polym.* **2017**, *169*, 227–235. [CrossRef]
146. Angelova, S.; Nikolova, V.; Molla, N.; Dudev, T. Factors Governing the Host–Guest Interactions between IIA/IIB Group Metal Cations and α-Cyclodextrin: A DFT/CDM Study. *Inorg. Chem.* **2017**, *56*, 1981–1987. [CrossRef]
147. Oliveri, V.; Pietropaolo, A.; Sgarlata, C.; Vecchio, G. Zinc Complexes of Cyclodextrin-bearing 8-Hydroxyquinoline Ligands: A Comparative Study. *Chem. Asian J.* **2016**, *12*, 110–115. [CrossRef]
148. Cheng, Y.; Wang, X.; Chang, D.; Li, W. DFT study on the effects of catalysis by β-cyclodextrin in the reaction of p-nitrophenyl acetate. *J. Mol. Model.* **2017**, *23*, 21. [CrossRef] [PubMed]
149. Maia, P.P.; de Sousa, S.M.R.; De Almeida, W.B.; Guimaraes, L.; Nascimento, C.S. Computational investigation on the host–guest inclusion process of norfloxacin into β-cyclodextrin. *J. Mol. Model.* **2016**, *22*, 220. [CrossRef] [PubMed]
150. Ignaczak, A.; Orszański, Ł. In Search of the Most Stable Molecular Configuration of Heptakis(2,6-O-dimethyl)-β-cyclodextrin and Its Complex with Mianserin: A Comparison of the B3LYP-GD2 and M062X-GD3 Results. *J. Phys. Chem. B* **2021**, *125*, 13077–13087. [CrossRef] [PubMed]
151. Aree, T. Advancing insights on β-cyclodextrin inclusion complexes with SSRIs through lens of X-ray diffraction and DFT calculation. *Int. J. Pharm.* **2021**, *609*, 121113. [CrossRef] [PubMed]

152. Aree, T. Distinctive Supramolecular Features of β-Cyclodextrin Inclusion Complexes with Antidepressants Protriptyline and Maprotiline: A Comprehensive Structural Investigation. *Pharmaceuticals* **2021**, *14*, 812. [CrossRef]
153. Aree, T. Supramolecular Complexes of β-Cyclodextrin with Clomipramine and Doxepin: Effect of the Ring Substituent and Component of Drugs on Their Inclusion Topologies and Structural Flexibilities. *Pharmaceuticals* **2020**, *13*, 278. [CrossRef]
154. Aree, T. β-Cyclodextrin encapsulation of nortriptyline HCl and amitriptyline HCl: Molecular insights from single-crystal X-ray diffraction and DFT calculation. *Int. J. Pharm.* **2020**, *575*, 118899. [CrossRef]
155. Rayene, K.; Imane, D.; Abdelaziz, B.; Leila, N.; Fatiha, M.; Abdelkrim, G.; Bouzid, G.; Ismahan, L.; Brahim, H.; Rabah, O. Molecular modeling study of structures, Hirschfield surface, NBO, AIM, RDG, IGM and 1HNMR of thymoquinone/hydroxypropyl-β-cyclodextrin inclusion complex from QM calculations. *J. Mol. Struct.* **2022**, *1249*, 131565. [CrossRef]
156. Macernis, M.; Bockuviene, A.; Gruskiene, R.; Krivorotova, T.; Sereikaite, J. Raman study for β-ring positioning in β-Carotene complexes with Cyclodextrins and Chitooligosaccharides. *J. Mol. Struct.* **2021**, *1226*, 129362. [CrossRef]
157. Kerdpol, K.; Daengngern, R.; Sattayanon, C.; Namuangruk, S.; Rungrotmongkol, T.; Wolschann, P.; Kungwan, N.; Hannongbua, S. Effect of Water Microsolvation on the Excited-State Proton Transfer of 3-Hydroxyflavone Enclosed in γ-Cyclodextrin. *Molecules* **2021**, *26*, 843. [CrossRef]
158. Bouchemela, H.; Madi, F.; Nouar, L. DFT investigation of host–guest interactions between α-Terpineol and β-cyclodextrin. *J. Incl. Phenom. Macrocycl. Chem.* **2019**, *95*, 247–258. [CrossRef]
159. Seridi, L.; Boufelfel, A. Naringenin encapsulation in β-CD and in heptakis(2,6-di-O-methyl)-β-CD:NMR, NBO and QTAIM analysis. *J. Incl. Phenom. Macrocycl. Chem.* **2018**, *90*, 287–304. [CrossRef]
160. Nutho, B.; Nunthaboot, N.; Wolschann, P.; Kungwan, N.; Rungrotmongkol, T. Metadynamics supports molecular dynamics simulation-based binding affinities of eucalyptol and beta-cyclodextrin inclusion complexes. *RSC Adv.* **2017**, *7*, 50899–50911. [CrossRef]
161. Nutho, B.; Khuntawee, W.; Rungnim, C.; Pongsawasdi, P.; Wolschann, P.; Karpfen, A.; Kungwan, N.; Rungrotmongkol, T. Binding mode and free energy prediction of fisetin/β-cyclodextrin inclusion complexes. *Beilstein J. Org. Chem.* **2014**, *10*, 2789–2799. [CrossRef] [PubMed]
162. Guendouzi, O.; Guendouzi, A.; Ouici, H.B.; Brahim, H.; Boumediene, M.; Elkeurti, M. A quantum chemical study of encapsulation and stabilization of gallic acid in β-cyclodextrin as a drug delivery system. *Can. J. Chem.* **2020**, *98*, 204–214. [CrossRef]
163. Yang, L.; Li, D.; Guo, B.; Wei, D. Theoretical Study on the Inclusion Interaction of β-Cyclodextrin with Gabapentin and Its Stability. *J. Struct. Chem.* **2019**, *60*, 564–574. [CrossRef]
164. Asztemborska, M.; Ceborska, M.; Pietrzak, M. Complexation of tropane alkaloids by cyclodextrins. *Carbohydr. Polym.* **2020**, *209*, 74–81. [CrossRef]
165. Abdel-Mottaleb, M.S.A.; Hamed, E.; Saif, M.; Hafez, H.S. Binding, and thermodynamics of β-cyclodextrin inclusion complexes with some coumarin laser dyes and coumarin-based enzyme substrates: A simulation study. *J. Incl. Phenom. Macrocycl. Chem.* **2018**, *92*, 319–327. [CrossRef]
166. Diamantis, D.A.; Ramesova, S.; Chatzigiannis, C.M.; Degano, I.; Gerogianni, P.S.; Karadima, K.E.; Perikleous, S.; Rekkas, D.; Gerothanassis, I.P.; Galaris, D.; et al. Exploring the oxidation and iron binding profile of a cyclodextrin encapsulated quercetin complex unveiled a controlled complex dissociation through a chemical stimulus. *Biochim. Biophys. Acta (BBA) Gen. Subj.* **2018**, *1862*, 1913–1924. [CrossRef]
167. Guendouzi, A.; Mekelleche, S.M.; Brahim, H.; Litim, K. Quantitative conformational stability host-guest complex of Carvacrol and Thymol with β-cyclodextrin: A theoretical investigation. *J. Incl. Phenom. Macrocycl. Chem.* **2017**, *89*, 143–155. [CrossRef]
168. Abdelaali, M.; Fatiha, M.; Leila, N.; Nora, M.; Mouna, C.; Sakina, H.; Eddine, K.D. Computational approach in the study of the inclusion processes of Thymol with β-cyclodextrin. *J. Mol. Liq.* **2017**, *242*, 714–721. [CrossRef]
169. Aree, T. β-Cyclodextrin Inclusion Complexes with Catechol-Containing Antioxidants Protocatechuic Aldehyde and Protocatechuic Acid—An Atomistic Perspective on Structural and Thermodynamic Stabilities. *Molecules* **2021**, *26*, 3574. [CrossRef] [PubMed]
170. Aree, T.; Jongrungruangchok, S. Structure–antioxidant activity relationship of β-cyclodextrin inclusion complexes with olive tyrosol, hydroxytyrosol and oleuropein: Deep insights from X-ray analysis, DFT calculation and DPPH assay. *Carbohydr. Polym.* **2018**, *199*, 661–669. [CrossRef] [PubMed]
171. Huang, F.; Zhuang, S.; Liu, W.; Lin, L.; Sun, L. Computational investigation on the chiral differentiation of D- and L-penicillamine by β-cyclodextrin. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2020**, *248*, 119277. [CrossRef]
172. Yang, S.; Wu, F.; Yu, F.; Gu, L.; Wang, H.; Liu, Y.; Chu, Y.; Wang, F.; Fang, X.; Ding, C.-F. Distinction of chiral penicillamine using metal-ion coupled cyclodextrin complex as chiral selector by trapped ion mobility-mass spectrometry and a structure investigation of the complexes. *Anal. Chim. Acta* **2021**, *1184*, 339017. [CrossRef]
173. Știufiuc, G.F.; Toma, V.; Onaciu, A.; Chiș, V.; Lucaciu, C.M.; Știufiuc, R.I. Proving Nanoscale Chiral Interactions of Cyclodextrins and Propranolol Enantiomers by Means of SERS Measurements Performed on a Solid Plasmonic Substrate. *Pharmaceutics* **2021**, *13*, 1594. [CrossRef]

174. Lee, S.-S.; Lee, J.-U.; Oh, J.H.; Park, S.; Hong, Y.; Min, B.K.; Lee, H.H.L.; Kim, H.I.; Kong, X.; Lee, S.; et al. Chiral differentiation of d- and l-isoleucine using permethylated β-cyclodextrin: Infrared multiple photon dissociation spectroscopy, ion-mobility mass spectrometry, and DFT calculations. *Phys. Chem. Chem. Phys.* **2018**, *20*, 30428–30436. [CrossRef]
175. Lee, S.; Park, S.; Hong, Y.; Lee, J.U.; Kim, J.H.; Yoon, D.; Kong, X.; Lee, S.; Bin Oh, H. Chiral differentiation of d- and l-alanine by permethylated β-cyclodextrin: IRMPD spectroscopy and DFT methods. *Phys. Chem. Chem. Phys.* **2017**, *19*, 14729–14737. [CrossRef]
176. Reyes-Reyes, M.L.; Roa-Morales, G.; Melgar-Fernández, R.; Reyes-Pérez, H.; Gómez-Oliván, L.M.; Gonzalez-Rivas, N.; Bautista-Renedo, J.; Balderas-Hernández, P. Chiral recognition of abacavir enantiomers by (2-hydroxy)propyl-β-cyclodextrin: UHPLC, NMR and DFT studies. *J. Incl. Phenom. Macrocycl. Chem.* **2015**, *82*, 373–382. [CrossRef]

ECOLE DOCTORALE

Titre: Une étude de certains perturbateurs endocriniens et de leur liaison aux molécules hôtes avec modélisation moléculaire

Mots clés: estradiol, cyclodextrine, champ de force AMOEBA, paramétrisation du champ de force, DFT, perturbateurs endocriniens

Résumé: Les perturbateurs endocriniens (Endocrine Disrupting Chemicals, EDC) sont des substances qui présentent des effets néfastes en raison d'un mode d'action endocrinien. Cela inclut souvent une interaction avec les récepteurs de la même manière que les ligands naturels des récepteurs. Parmi les EDC, il existe des ingrédients pharmaceutiques actifs (API) tels que les stéroïdes hormonaux. Les cyclodextrines (CD) sont des oligosaccharides cycliques utilisés comme systèmes d'administration de médicaments pour les API à faible solubilité dans l'eau et comme agents d'élimination des toxines. Le but de cette étude était de développer différentes techniques de modélisation moléculaire pour analyser les interactions entre les EDC choisis et les récepteurs d'œstrogènes ou CD. Les méthodes suivantes ont été appliquées : paramétrisation des EDC choisis (estradiol, progestérone, biephénol A) et CD dans le champ de force polarisable AMOEBA et simulation dynamique moléculaire réussie du système récepteur d'œstrogène + EDC ; tests de référence de diverses approches de calcul basées sur la mécanique quantique (DFT, MP2, semi-empirique) et la mécanique moléculaire (MD/MMGBSA) et les paramètres applicables, sur l'exemple du système estradiol+βCD.

Title: A study of selected endocrine disrupting chemicals and their binding to host molecules with molecular modelling

Keywords: estradiol, cyclodextrin, AMOEBA force field, force field parametrization, DFT, endocrine disrupting chemicals

Abstract : Endocrine Chemical Disruptors (EDCs) are substances that exhibit adverse effects as a consequence of an endocrine mode of action. It often includes interaction with receptors in the same way as receptor's natural ligands. Among EDCs there are Active Pharmaceutical Ingredients (APIs) such as steroid hormones. Cyclodextrins (CDs) are cyclic oligosaccharides used as drug delivery systems for APIs of a low solubility in water, and as toxin removing agents. The goal of this study was to develop different molecular modelling techniques to analyze interactions between chosen EDCs and Estrogen Receptor or CDs. Following methods have been applied: parametrization of chosen EDCs (estradiol, progesterone, biephenol A) and CD in AMOEBA polarizable force field and succeeding Molecular Dynamics simulation of the Estrogen Receptor + EDC system; benchmark tests of various Quantum Mechanics (DFT, semi-empirical) and Molecular Mechanics (MD/MMGBSA) based computation approaches and applicable parameters, on the example of estradiol+βCD system.

Institut Polytechnique de Paris
91120 Palaiseau, France