Molecular design for catalytic activities of helical chiral oligoureas

Yaidel Toledo Gonzalez

THÈSE PRÉSENTÉE POUR OBTENIR LE GRADE DE

DOCTEUR DE L'UNIVERSITÉ DE PAU ET DES PAYS DE L'ADOUR

ÉCOLE DOCTORAL DES SCIENCES EXACTES ET LEURS APPLICATIONS (ED 211)
SPÉCIALITÉ: Chimie-Physique

Par Yaidel TOLEDO GONZÁLEZ

MODÉLISATION MOLÉCULAIRE POUR L'ÉTUDE DES PROPRIÉTÉS CATALYTIQUES DE FOLDAMÈRES HÉLICOÏDAUX À BASE D'URÉE

Sous la direction de:	Prof. Philippe CARBONNIÈRE
Codirecteur:	Prof. Jean-Marc SOTIROPOULOS
Rapporteur:	Prof. Antonio MONARI
Rapporteur:	D.R. Aurélien DE LA LANDE
Examinateur:	Prof. Céline LEONARD
Examinateur:	H.D.R. Antonia MIELGO
Invité:	D.R. Gilles GUICHARD

Pau, February 8, 2023

Una vez más,
A Papi y Nena,
mis viejos por siempre;
A Tata, mi papá, Dari y Fran,
por darme vida;
A Clau,
por todo.

Résumé

Les catalyseurs de type oligomères bioinspirés à base d'urée de formes hélicoïdales (foldamères d'urée) se sont avérés être des alternatives robustes aux catalyseurs asymétriques organométalliques. Ces catalyseurs organiques permettent des réactions qui conduisent à la formation de liaisons C-C avec une énantiosélectivité extrêment élevée et des charges aussi faibles que des rapports molaires catalyseur chiral/substrats de 1:10 000. Cependant, malgré les connaissances acquises sur la performance catalytique de ces foldamères d'urée, le mécanisme réactionnel à l'échelle atomique dans lequel il intervient ainsi que l'identification des paramètres structuraux qui font leur excellente énatiosélectivité ne sont pas encore bien compris. L'objet de cette thèse est de participer à élucider ces aspects par les outils de la modélisation moléculaire au niveau quantique. Cette étude se focalise sur l'étude d'une réaction type catalysée par ces foldamères : l'addition conjuguée d'un composé 1,3-dicarbonyl pronucléophile à des nitroalcènes. À cette fin, l'étude est réalisée en trois partie : i) Une analyse de reconnaissance moléculaire des composants du système à partir d'une procédure d'optimisation globale qui a révélé la modalité d'encrage la plus probable des deux réactifs sur le foldamère. ii) La modélisation du profil énergétique de ces structures qui a révélé l'excellente énantiosélectivité du catalyseur ainsi que la chiralité des molécules formées en excellent accord avec les données expérimentales. Ainsi et en retour, ces observations conduisent à valider le mécanisme d'encrage proposé préalablement. iii) L'identification des paramètres électroniques et structuraux à la base de l'énantiocontrôle et en conséquence des propositions d'amélioration des performances par "design moléculaire". Les résultats présentés ici se veulent d'une part servir de base théorique à l'amélioration des propriétés catalytiques de ces foldamères et d'autre part sur un plan plus méthodologique à une modalité d'investigation théorique pour des processus catalytiques à portée éminemment sociétale puisque qu'elle renvoie à la proposition de solutions pour l'environnement (par exemple la synthèse de polymères biosourcée) et l'énergie (par exemple le stockage chimique de l'hydrogène).

Abstract

Bioinspired urea-based oligomers that fold with high fidelity (foldamers) have been shown to be robust alternatives to asymmetric organometallic catalysts. These organic catalysts facilitate reactions that lead to the formation of C-C bonds with extremely high enantioselectivity and loadings as low as 1:10000 chiral catalyst/substrates molar ratios. However, despite the insights gained on the catalytic performance of these urea foldamers, the atomistic scale reaction mechanism in which they are involved, as well as the identification of the structural parameters related to their excellent enatioselectivity are not yet well understood. The aim of this thesis is to participate in the elucidation of these aspects by the tools of molecular modeling at the quantum level. This study focuses on the study of a typical reaction catalyzed by these foldamers: the conjugate addition of a pronucleophilic 1,3-dicarbonyl compound to nitroalkenes. To this end, the study is carried out in three parts: i) A molecular recognition analysis of the components of the system based on a global optimization procedure which revealed the most probable anchoring modes of the two substrates on the catalyst; ii) The modeling of the energy profile of these structures which revealed the excellent enantioselectivity of the catalyst as well as the chirality of the molecules formed in excellent agreement with the experimental data. Thus, and in return, these observations lead to the validation of the previously proposed anchoring mechanism. iii) The identification of the electronic and structural parameters at the basis of the enantiocontrol and consequently proposals for improving the foldamers catalytic performance by "molecular design". The results presented here are intended, on the one hand, to serve as a theoretical basis for the improvement of the catalytic properties of these foldamers, and on the other hand, in a more methodological level, to a theoretical investigation modality for catalytic processes with eminently societal scope since it refers to the proposal of solutions for the environment (e.g. the synthesis of bio-based polymers) and energy (e.g. the chemical storage of hydrogen).

Acknowledgments

This thesis is the result of three years of research with the unconditional help of Prof. Philippe CARBONNIERE and Prof. Jean-Marc SOTIROPOULOS, at the *Institut des Sciences Analytiques et de Physico-Chimie pour l'Environnement et les Matériaux*, in the University of Pau and the Adour Countries. I would like to thank them both in the first place for their patience, dedication, advice, and always instructive discussions during this time. Also, without their precise suggestions and corrections to this manuscript, it would never have reached the stage it is today. To them my most sincere thanks. Luckily enough, our adventure does not ends here!

This work is part of the ANR project `HCO_for_LLAC`, which involves partners from the University of Bordeaux and the University of the Basque Country. From them I have also received highly estimated suggestions and feedback during the group meetings along these three years. Thanks to Gilles GUICHARD, Daniel TATON, Antonia MIELGO, and Mikel OIARBIDE for their time and company during this time. I am sure more lunches are ahead!

I hold Aurélien DE LA LANDE and Antonio MONARI in high regard, to whom I am doubly grateful. In the first place I would like to thank them for accepting the invitations to serve as reviewers of this manuscript, and for their highly appreciated comments on it. Furthermore, I would like to thank Aurélien for his hospitality, time, and lessons during my stay at its Laboratory in Orsay learning how to use the QM/MM module of deMon2k. It has great value to me. Also Antonio for the time shared at Namur, in hope we can continue those wonderful meetings. Additionally, I would like to thank Céline LEONARD for her acceptance on being part of the jury for the defense.

At the IPREM there are lots of people to acknowledge, which have spent time with me and therefore have contributed directly or indirectly to the project and my life. Les amis de l'oficine, los coterráneos venidos de tierras del cono sur, l'accueil avec Danielle, Zoyne (hablate!) who by the way warned me about the thesis offer, El Guti for his help even in this short time, to Pepe for bringing a piece of Cuba to share... and many, many more.

Of course to “La gente del Rez de Chausée”: El Javi, Fran, Celia, La Silvi, Rosanna, Gustavito, Alba, Khouloud... I do not know if I forget someone... Ah! El Mikel (desheredado por no saber decir ni pinga, helo aquí) ¡Baff que momentos aquellos cuando eramos jóvenes! Also Las Chicas del Master. Thank you all for the help, the beers, the histories, el salseo, the mountains, el Baracuba, el Covid, the lunches, the nights at the ZHC, etc. etc. etc. Wonderful three years!

To my dearest friends del exilio: Marina, Jorge, Ricar, Sol y El Rosa. Thanks for all the moments, in hope even more and with 10x frequency could be waiting for us ahead. Nos pillamos por ahí.

To those from the University which were and are always a huge support, especially El Yosva (y a Yamila por sacarlo de la mala vida), José, Manolito, Anthuan, Luisa...

A la gente del barrio y del pre, quienes nos hemos visto envueltos en esta Gran Dispersión, a decir de Virgilio Piñeira. Se les quiere.

Por supuesto a toda mi familia, quienes han dado lo que tienen y más para que hoy estemos aquí. Un besote grande para Tatica, el Pinot y Dailén, la vieja y el viejo, mi papá, tío, Fran, Clau, la mama, Yarenis, Islay, el cuiqui, Nevys, Kevin, mis tíos, tía Deisy, Zenia, Yoli, Carmen, mis primos de aquí y los de allá...

¡Llegamos, paso a paso!

Contents

Introduction

Stereoselective synthesis is defined by the International Union of Pure and Applied Chemistry (IUPAC) as "a chemical reaction (or reaction sequence) in which one or more new elements of chirality are formed in a substrate molecule and which produces the stereoisomeric (enantiomeric or diastereoisomeric) products in unequal amounts" [1]. This constitutes a widely utilized, important, and under heavy investigation process in chemistry, which finds application in plenty of every day use materials, cosmetics, drugs, etc.

The Pharmaceutical Industry is one of the fields where stereoselective synthesis plays an important role, because the different enantiomers or diastereomers of a molecule often have different biological activity. Therefore, the use of a wrong stereoisomer could lead to disastrous results.

Precisely in the progress of this industry it is found a picture of the evolution and understanding of stereoselective synthesis throughout the years. There exists reports from the tragic cases of administration of (S)-Thalidomide to pregnant women in the 1960s, and the nearly 90% of synthetic chiral drugs being racemic in the early 1990s, to the 123 billion US dollars single-enantiomer compounds market valuation by the 2000s [2], this last jump influenced by the "chiral switch" [3, 4]. By 2010, almost 70% of the new drugs introduced in the market were single-enantiomer ones, and in 2015, only one of the U.S. Food and Drug Administration (FDA) approved chiral drugs was not enantiomerically pure [5] (from a set of 45).

The problem related to the development of enantiomerically pure substances is that the enantiomers have, in theory, the same energy, therefore, favoring one or the other in a mixture is not always a simple task [6]. To synthetically obtain substances with only one stereoisomer, several approaches could be used: (i) enantioselective catalysis, which is performed using chiral catalysts [7–9]; (ii)

chiral auxiliaries, which are organic compounds coupling to a starting material to form a new compound which can then undergo diastereoselective reactions via intramolecular asymmetric induction [10]; (iii) biocatalysis, which makes use of biological compounds, ranging from isolated enzymes to living cells, to perform chemical transformations [11]; (iv) enantioselective organocatalysis, which is similar to enantioselective catalysis, but this time the catalysts is and organic compound [12, 13]; and (v) chiral pool synthesis, where a readily available chiral starting material is manipulated through successive reactions to obtain the desired target molecule. From them, one may infer that catalysis is of common use in the obtention of stereoisomerically pure substances. Indeed, catalysis is effective for a broader range of transformations than any other method of stereoselective synthesis.

Applications of enantioselective catalysis (also known as asymmetric catalysis) have considerably grown together with the efforts to alleviate the pollution and environmental disasters the extended exploitation of petroleum and its derivatives has brought with it. Petroleum sourced plastics are all around us, and their degradation process once they are wasted is extremely difficult. They generally end in landfills or in the oceans, where they split into microplastics posing a serious danger to the life of many species, including our own. Vinyl polymers are the most significant class of petroleum sourced plastics [14], with over 3000 formulations commercially available which include a wide spectrum of properties and applications [15]. The introduction of everyday use materials capables to completely degrade into nature, and be produced from renewable bio-resources, will considerably help to solve the problems derived from vinyl polymers waste, and thus contribute to a sustainable development. However, to their widespread utilization they must be competitive with current petroleum derived plastics. On these grounds considerable advances have been made with, for example, the stereoselective polymerization of the α-methylene-γ-butyrolactone (MBL) polymer, which is a 10 years old biosourced vinyl monomer with thermal properties superior to petroleum-sourced plastics [16].

Nevertheless, the use of renewable, biomass-sourced and biodegradable materials have to overcome many obstacles to extend to a large scale production, and replace its petro-chemical counterparts. To be successful, the manufacturing processes of these polymers must become more cost-effective, and they must diversify their end-of-use [17]. Some examples of degradable polymers with defined stereochemistry as replacement of polypropylene, polystyrene, and other commodity polyolefins, are polylactide, polymenthide, and polyhydroxyalkanoates. Precisely the polylactide

(PLA) is the most widely recognized renewable polymer [18], which can be obtained from corn sugar, potatoes, and sugar cane [19] by polymerization processes. In turn, polymerization is supported by the use of stereoselective organocatalysts capables to induce the necessary mechanisms of reaction. Therefore, some key factors defining the obtention or not of a competitive material, are the efficiency and efficacy of the catalysts utilized.

Examples of catalysts commonly applied in organocatalysis are the conjoined thiourea-amine [20], the (–)-sparteine [21], among others. They are small molecules with the ability to provide catalytic activities with $> 98\%$ of monomer conversion within 0.5h to 2.0h with 0.5 mol % and 1 mol % mass concentration of the catalyst, respectively [22].

Furthermore, the progressive accumulation of information in the study of enzymes has served as inspiration for the design of high performance synthetic asymmetric catalysts. These catalysts, which are small to medium size foldamers,[1] have been proven to retain interesting enzyme's properties as the modularity and tunability, and also provide attractive stereo-, chemo-, and site-selectivity [24–27]. This is the case of neutral N,N'-linked helical foldamers conformed by urea type units in their backbone, which have been reported to enantioselectively catalyze challenging C-C bond-forming reactions with concentrations 400 times lower than common small molecule bifunctional catalysts [28, 29].

These foldamers have great characteristics to perform exceptionally in biomimetic catalysis, exhibiting long-range conformational order (e.g., helical), property which can not be reproduced by small molecules. Yet, little progress has been made in the direction of enantioselective catalysis with such molecules.

All the mentioned asymmetric catalysts, from the small molecule bifunctional catalysts to the foldamers, have increasingly gained attention through the last decades due to their suitability as mono- and bis-(thio)ureas H-bond donor catalysts, and thus have been extensively studied [30–39]. However, these systems continue to be a challenge in terms of structure and function correlation. For example, questions related to the catalyst-substrates interactions, mechanisms of reaction, and improvement of the results are to the date without answers. An effort to understand the basic phenomena involved in these molecules catalytic properties could be of great help for the development and scaling up to the commercial level of the

[1] The definition of foldamers here employed is that proposed by Hill et al. [23] as "any oligomer that can fold into a well-defined conformation in solution, the structures of which are stabilized by a collection of noncovalent interactions between non-adjacent monomer units"

needed renewable, biomass-sourced, and biodegradable polymers substituting the petroleum derivatives.

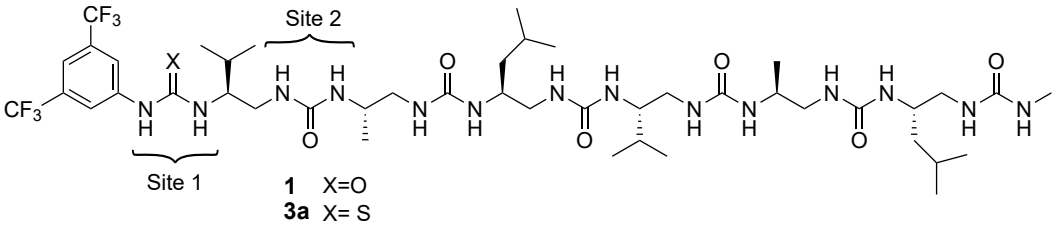
Foldamer **1** (see Figure 1a) was employed as catalyst in a conjugated addition synthesis, where it was supposed to act on nitroalkenes, and in conjunction with the influence of an achiral base on 1,3-dicarbonyl substrates would lead to γ-nitro carbonyl adducts through this synergistic nucleophile/electrophile strategy (see Figure 1b) [29]. In the presence of an achiral Brønsted base such as Et_3N or DIPEA, and using toluene as solvent, it was observed that:

1. The helix conformation of the foldamers has a direct influence on the reactivity and enantioselectivity of the reactants, being both properties very poor in the 1-mer and 2-mer foldamers (unfolded due to their length) and increasingly favorable when increasing the number of residues of the foldamer.

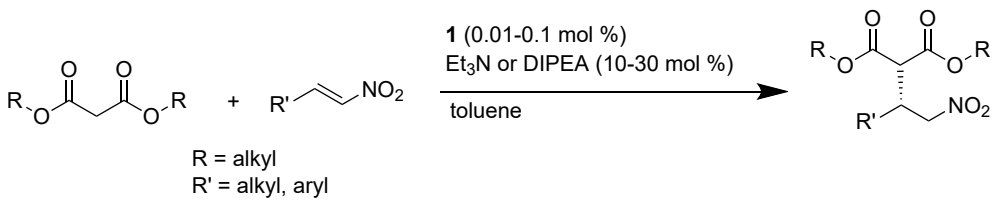
2. The number of residues is a crucial parameter: 5-mer and 6-mer foldamers presented remarkable reactivity and enantioselectivity, but shorter chain lengths failed in doing so.

3. Not only the first (thio)urea but also the second plays a fundamental role in the catalytic activity of the foldamer (marked as "site `[1]`" and "site `[2]`" in Figure 1a, respectively).

4. The needed concentration of catalyst are extremely low, in the range of 1:1000 to 1:10000 catalyst/substrate ratio.



(a) Heptaurea foldamer



(b) General reaction

Figure 1. Heptaurea representation of the foldamers studied by Bécart et al. [29] to evaluate its catalytic properties in the addition of 1,3-dicarbonyl to nitroalkenes.

However, from the obtained results further questions have raised, as for example, (i) which is the exact role of the first two ureas in the catalytic process? (ii) Why is the helical shape determinant in the catalytic process? (iii) Why only 5 and 6 residues long foldamers are able to catalyze the C-C bond formation with high enantiocontrol? (iv) Could longer foldamers (7 to 10 residues long) still retain the same selectivity than those in the range of 5 to 6 residues?

The objective of this investigation is to rationalize the role of chiral aliphatic N,N'-linked oligourea foldamers like **1** as catalyst taking as model reaction the conjugated addition of 1,3-dicarbonyl to nitroalkenes to support and explain insightful experimental findings previously reported by Bécart et al. [29]. Some authors have already employed these mentioned results in their exploration of chiral organocatalysts [40–45, for example], however no theoretical study oriented to explore the systems and essay to answer the basic questions raised by them has been found by the authors. This lack of theoretical investigations has also been recognized (and demanded) by various authors as an obstacle to the full development of these environmentally friendly technologies [46]. The use of computational chemistry could significantly help to better understand these novel systems.

In this direction, Computational Chemistry has increasingly become a valuable tool for the prediction and development of catalysts, nowadays some authors even place it as "an essential component of catalyst design" [47]. Computational chemistry is a branch of chemistry that uses computer simulation to assist in solving chemical problems. It uses methods of theoretical chemistry, incorporated into computer programs throughout numerical chemistry, to calculate the structures and properties of all sort of chemical systems. Using these techniques makes possible to describe catalytic process as, for example, the mechanisms and their selectivity for different products, trough the study of the system's transition states.

In fact, starting from the first organocatalytic reaction studied computationally back in the early 2000 [48–50], the use of computational methods has evolved, and today it is not only possible to rationalize the reactions but also predict and propose the catalysts characteristics [51]. Recent investigations prove it, with the study of the reaction between enoldiazoacetamide and N-methylindole [52], the use of catalysis in the epoxidation of olefin [53], the use of Noyori-Ikariya chiral molecular ruthenium complex as catalyst [54], the enantioselectivity relationship for the asymmetric addition of diethylzinc to benzaldehyde [55], just to cite some examples.

Nevertheless, taking into account (i) the large size of the system, (ii) the num-

ber of molecules involved, and (iii) the importance of an accurate description of the noncovalent interactions due to its key role in organocatalytic processes, many challenges arise in the investigation of systems like that represented in Figure 1. Perhaps the most important challenge to overcome is the determination of the most stable configuration of the system's components (molecules) ones with respect to the others. Furthermore, the employment of accurate computational methods usually involve a large demand of computational resources, which quickly grow with the number of atoms of the system.

In this investigation it is used a home-made Global Optimization algorithm which is capable to identify the most stable arrangements obtained from a system composed by several bodies. Furthermore, in some parts it was decided to employ Density Functional Theory with some additional ingredients as an implicit solvent model to reproduce the experimental conditions and empirical dispersion corrections to account for better description of the long range interactions between the molecules under query. In some other parts, mainly where large computational resources were needed, it was decided to keep the Density Functional Theory for the description of the reactive region of the system, and simulate the non-reactive region with a more computationally efficient semi-empirical method. Additionally to this paragraph light overview of the main ideas employed in the manuscript to computationally study the catalytic system, Chapter 1 provides a deeper analysis on the theories and methods employed.

This manuscript is structured in four chapters. Chapter 1 briefly describes the methods and theories upon which the computational chemistry calculations further developed in other chapters are supported by. The theoretical foundations of Quantum Chemistry are discussed, along with the most widely currently employed methods to numerically obtain information of the phenomenon taking place at the molecular level in quantum systems. Furthermore, discussions on the selection of the functional and basis sets adequate to describe organocatalytic processes, the use of a solvent, and long-range interaction corrections are exposed. Also, subjects on finding the global minimum of a molecular cluster and the molecular recognition process are treated. Finally, a review on the limits and previous applications of these methods to problems similar to the one being treated here has been developed, with the final results of defining which method was better suited to carry out our computations.

The specific catalytic systems under investigation is presented and characterized

in Chapter 2. There, we have used most of the theoretical annotations made in Chapter 1 to study the molecular recognition process between the catalyst and the substrates, identifying the preferred position of each molecule with respect to the others throughout a Global Optimization procedure. With that information, the analysis of the nature of the system's main interaction has also been carried out. This chapter answers to some fundamental questions that the experimental results had raised, providing insights of the molecular recognition process between catalyst and substrates. However, when dealing with enantioselective catalysis, the computation of the energetic profile and the estimation of the enantiocontrol is mandatory.

Chapter 3 is devoted to the analysis of the energetic profiles of the catalyzed reactions. In it, the structures obtained in Chapter 2 from the Global Optimization procedure have been utilized to determine its associated transition states. This study has additionally been carried out varying the number of residues of the foldamer, to identify any possible dependence between this parameter and the enantiocontrol.

Nevertheless, the identification of the key parameters influencing the enantiocontrol in the foldamers was missing at this point. To that end, several studies of different properties were carried out to correlate the enantiocontrol with geometrical parameters of the foldamer and the interactions strength between catalyst and substrates.

With the description of the catalytic system made up to this point, in Chapter 4 we undertake the dicey adventure of modifying the original foldamer to identify how changes to its backbone and side-chains may vary the outcome of the reaction and thus the foldamer's catalytic properties.

Finally, this manuscript is complemented by a series of Appendixes, some of which are dense enough to take a space on the main body. However, it was decided to isolate the information in these appendixes because despite being important to the investigation, they are indirectly related to the solution of the problems faced, and would have momentarily break the idea of the analysis.

Chapter 1

Theoretical tools and state of the art

On July 5th, 1687, was published Isaac Newton's series "Mathematical Principles of Natural Philosophy", where its Law of Universal Gravitation and Laws of Motion were presented. As these theories describe and predict so precisely the movement of macroscopic objects, and no experimental results were in disagreement with them at the time, they were taken as granted for more than two centuries. The classical behavior of a system was completely described by knowing its initial state.

However, by the end of the XIX century, some new phenomena involving microscopic objects were not finding explanations on Newton's Classical Mechanics. On December 14th, 1900, at the German Society of Physics, Max Plank presented its theory for the *balck-body radiaton* problem, suggesting that the emission and absorption of light by substance does not take place continuously, but by small quantized portions. This idea shaken the whole Classical Mechanics, and is is today known as the beginning of Quantum Mechanics: the set of laws describing the behavior of small particles as electrons and nuclei of atoms.

The use of quantum mechanics to explore chemistry problems is know as Quantum Chemistry, and it finds applications in all branches of chemistry. For example, organic chemists estimate stabilities of molecules and investigate reaction mechanisms with it, inorganic chemists use ligand field theory, analytical chemists use spectroscopy methods, and biochemists study ezyme-substrate interactions.

The present chapter concisely touches the theories, methods, and tools that will be employed throughout the rest of the manuscript to investigate the catalytic system composed by foldamers like **1**, and the triethylamine achiral Brønsted base, in challenging C-C bond formation reactions.

1.1 The Schrödinger equation

To describe the sate of a system in quantum mechanics it is necessary the use of a function of its particles' coordinates and the time called **wave function** or **state function** $\Psi = \Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots, t)$. This function contains all possible information about the system, and evolves in time according to the following expression (for a one-particle, one-dimensional system):

$$-\frac{\hbar}{i}\frac{\partial\Psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2} + V(x,t)\Psi(x,t) \tag{1.1}$$

where m is the mass of the particle, $i = \sqrt{-1}$, $\hbar \equiv h/2\pi$ (h being Plank's constant), and $V(x,t)$ represents the potential energy of the system. The generalization to a many-particles, many-dimensions system is straightforward.

When the potential energy is not dependent on time, the wave function can be separated in terms of coordinate and time as

$$\Psi(x,t) = f(t)\psi(x) \tag{1.2}$$

Substituting equation (1.2) into (1.1) results in an equation for the evolution of time

$$\frac{df(t)}{f(t)} = -\frac{iE}{\hbar}dt \tag{1.3}$$

and other for the evolution of the coordinate

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\psi(x) = E\psi(x) \tag{1.4}$$

where E is a constant with energy dimensions, which is postulated to be the total energy of the system, and the term enclosed in square brackets is an operator known as Hamiltonian $\hat{\mathcal{H}}$. Equation (1.4) is the time-independent Schrödinger equation, and it is the one we refer to when mentioning the term "Schrödinger equation" hereafter.

The objective of Quantum Mechanical Computations is to solve the Schrödineger equation $\hat{\mathcal{H}}\psi = E\psi$. In simple situations, as the particle in a one-dimensional box and the hydrogen atom, the equation of the system is separated into different uncoupled equations involving only one variable, which are much simpler to solve than equation (1.4). However, for more complex systems, it is not possible to make the variables separation procedure, and the exact solution cannot be obtained. Take a system formed by M nuclei and N electrons, the non relativistic Hamiltonian

becomes into

$$\begin{aligned}\hat{\mathcal{H}} = &-\sum_{k}^{M}\frac{\hbar^2}{2m_k}\nabla_k^2 - \sum_{i}^{N}\frac{\hbar^2}{2m_e}\nabla_i^2 + \sum_{i}^{N}\sum_{j>i}^{N}\frac{e^2}{|\mathbf{r}_i-\mathbf{r}_j|} \\ &-\sum_{k}^{M}\sum_{i}^{N}\frac{Z_k e^2}{|\mathbf{R}_k-\mathbf{r_i}|} + \sum_{k}^{M}\sum_{l>k}^{M}\frac{Z_k Z_l e^2}{|\mathbf{R}_k-\mathbf{R}_l|}\end{aligned} \tag{1.5}$$

where $\mathbf{R}_k$, m_k, and ∇_k^2 are the vector position of nuclei, the mass of the nuclei, and the Laplacian operating on nuclei coordinates, respectively. The same is applied for the electrons with $\mathbf{r}_i$, m_i, and ∇_i^2.

Equation (1.5) reflects the presence of attraction pairwise terms between the electrons and the nuclei, and repulsive nuclei-nuclei electron-electron pairwise interactions. This **correlation** between the particles is the main reason why exact solutions for systems other than the hydrogen atom are impossible to obtain. In order to extend the theory to more complex systems, it is necessary to include simplifications and approximations [56]. Nevertheless, care must be taken because the roughest the approximations are, the less accurate a method will be with respect to the exact solution of equation (1.4)

1.1.1 Born-Oppenheimer approximation

The Born-Oppenheimer (BO) approximation [57] assumes that the motion of atomic nuclei and electrons in a molecule can be separated. As the masses of the nuclei are considerably higher than the masses of the electrons, during the time of an electronic motion cycle, the change in nuclear configuration is negligible, thus, to a high-quality approximation, the electrons are moving in a field of fixed nuclei. This approximation makes possible to suppress the kinetic energy of the nuclei represented in the first term of equation (1.5), and consider as a constant the nuclei-nuclei repulsive interaction represented in the last term. As any constant added to the a Hamiltonian operator only changes the values of the total energy and has no effect on the wave functions, the nuclei repulsion term can be dropped thus obtaining the Schrödinger equation for the electronic motion

$$\hat{\mathcal{H}}_{\text{elec}}\psi_{\text{elec}}(\mathbf{r},\ \mathbf{R}) = \varepsilon_{elec}(\mathbf{R})\psi_{\text{elec}}(\mathbf{r},\ \mathbf{R}) \tag{1.6}$$

with

$$\hat{\mathcal{H}} = -\sum_{i}^{N}\frac{\hbar^2}{2m_e}\nabla_i^2 + \sum_{i}^{N}\sum_{j>i}^{N}\frac{e^2}{|\mathbf{r}_i-\mathbf{r}_j|} - \sum_{k}^{M}\sum_{i}^{N}\frac{Z_k e^2}{|\mathbf{R}_k-\mathbf{r_i}|} \tag{1.7}$$

and $\mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2, \ldots\}$, $\mathbf{r} = \{\mathbf{r}_1, \mathbf{r}_2, \ldots\}$. As can be seen both the electronic wave function ψ_{elec} and the pure electronic energy $\varepsilon_{\text{elec}}$ depend parametrically on the nuclei positions. The electronic contribution to the total energy is then obtained as

$$\varepsilon_{\text{tot}}(\mathbf{R}) = \varepsilon_{\text{elec}}(\mathbf{R}) + \sum_{k}^{M} \sum_{l>k}^{M} \frac{Z_k Z_l e^2}{|\mathbf{R}_k - \mathbf{R}_l|} \tag{1.8}$$

It is time now to consider nuclear motions. According to our picture, the electrons move much faster than the nuclei, thus it is reasonable to think that the nuclei move in an average field of the electrons in which the total energy $\varepsilon_{\text{tot}}(\mathbf{R})$ provides a potential for the nuclear motion. Hence the Schrödinger equation for the nuclear motion is

$$\hat{\mathcal{H}}_{\text{nuc}} \psi_{\text{nuc}}(\mathbf{R}) = E \psi_{\text{nuc}}(\mathbf{R}) \tag{1.9}$$

where

$$\hat{\mathcal{H}}_{\text{nuc}} = -\sum_{k}^{M} \frac{\hbar^2}{2m_k} \nabla_k^2 + \varepsilon_{\text{tot}}(\mathbf{R}) \tag{1.10}$$

and E is the total energy of the system that includes electronic, vibrational, rotational and translational energy. The nuclear wave function ψ_{nuc} describes the vibration, rotation, and translation of the system. Reached this point, we can introduce the concept of potential energy surface (PES): a PES is a surface defined by ε_{tot} over all nuclear coordinates on which the nuclei evolve. Note that hereafter the PES will be identified as V.

Briefly, within the BO approximation the problem of finding the solution to the Schrödinger equation is twofold. First, one solves the electronic equation for a grid of different nuclei orientations, thus obtaining the PES. Second, the nuclear equation is solved using the PES obtained in the previous step. Finally, the total wave function is

$$\psi(\mathbf{r}, \mathbf{R}) = \psi_{\text{elec}}(\mathbf{r}, \mathbf{R}) \cdot \psi_{\text{nuc}}(\mathbf{R}), \tag{1.11}$$

This approximation considerably simplifies the solution process of the Schrödinger equation, because $\psi_{elec}(\mathbf{r}, \mathbf{R})$ and $\psi_{nuc}(\mathbf{R})$ can be solved independently.

1.2 Approaches to describe atomistic systems

Nowadays, four families of methods have been constantly evolving to practice the computation of behaviors at the atomistic scale despite of the barriers imposed by the complexities of the Schrödinger equation, and provide adequate answers to the

questions faced by the researchers:

1. molecular mechanics (MM),
2. semiempirical (SE),
3. *ab initio* (QM), and
4. Density Functional Theory (DFT).

Molecular Mechanics (MM) methods are the less close to the exact solution of equation (1.4), but also those more efficient in computational resources. They treat molecules as composed by a collection of balls (the atoms) which are held by springs (bonds). Knowing initially the position of the atoms of a given system, and the force field which acts upon them, it is possible to determine the infinitesimal move of the particles from the resolution of Newton's equations, and therefore the future evolution of a system for several millions of such computations. Note that the energy of the system is not computed by the resolution of the Schrödinger equation, but obtained from a mathematical expression depending on the position of the atoms and the empirical estimation of the strength of their interactions, called force field. Some examples of widely utilized MM force fields are AMBER [58] (organic systems), CHARMM [59, 60] (organic systems), OPLS-AA [61, 62] (organic materials), UFF [63] (most atoms of the periodic table), and GUPTA [64] (nanoalloys).

Semiempirical calculations try to solve the Schrödinger equation, but uses a set of experimental values to parameterize the complex integrals rising from the electrons correlation. This provides more accurate results, but with higher computational resources consuption than MM. Further details are provided in section 1.3 on the following page.

On the other hand, *ab initio* (QM) calculations do solve the electronic Schrödinger equation without any parametrization, but using approximations as the Born-Oppenheimer and taking ψ as an antisymmetrized product of one-electron spin-orbitals using a finite (and hence incomplete) basis set. The results obtained by this method are the energy and a *wavefunction*. Theoretically, the *wavefunction* contains all the information related to a system, and thus all its properties can be obtained.

Finally, the Density Functional Theory (DFT) [65–68] avoid the complexities of dealing with the wave function by computing the molecular electron probability density ρ. With this information it is possible to obtain the molecular electronic energy and various properties without computing the molecular wave function [56]. The method is further detailed in section 1.5 on page 16.

Here it is pointed out that *ab initio* methods are categorized as *first principle methods*, i.e. a method developed on the basis of a physical theory and constants only, without any additional input. Similarly, as the Density Functional Theory was also developed only on the basis of physical theories, then it can be categorized under the *first principle* bracket. Nevertheless, a debate in this subject is commonly open due to the fact that currently employed functionals during practical applications of DFT include parameters which were not determined from first principles, but empirically.

In some situations, mainly where the systems under study are exigent in computational time and resources due to their complexity, the above methods can be mixed. The application of Quantum Mechanics methods (QM) in the region of the system which influences the property(ies) under observation, in conjunction with MM methods applied to a region with low influence is a viable approach to save computational resources while at the same time the results obtained are usually in agreement with more exact theoretical treatments [6]. This approach is known as QM/MM, however, QM/Semiempirical (QM/SE) are also possible viable methods which can even be more accurate than QM/MM. QM/SE methods will be employed for the analysis of several systems in the present investigation.

In next sections some of the above mentioned methods will be further described.

1.3 Semiempirical methods

Semiempirical calculations modify Hartee-Fock calculation (see section 1.4) by introducing functions with empirical parameters. Within this framework, certain pieces of information, such as two electron integrals, are approximated or completely omitted. In order to correct for the errors introduced by omitting part of the calculation, the method is parameterized by curve fitting in a few parameters or numbers, in order to give the best possible agreement with experimental data.

The merit of semiempirical calculations is that they are much faster than the *ab initio* calculations. The demerit of semiempirical calculations is that the results can be slightly defective. If the molecule being computed is similar to molecules in the database used to parameterize the method, then the results may be very good. If the molecule being computed is significantly different from anything in the parameterization set, the answers may be very poor.

One of the earliest semiempirical approaches in quantum chemistry was the π-electron method of Hückel [69], developed in the 1930s to treat planar conjugated hydrocarbons. Since then, significant advances have been made, and the methods

with practical relevance nowadays include AM1 [70], OM3 [71], PM6 [72] and PM7 [73].

AM1 methods started to develop in the mid 80' as an attempt to correct the errors observed in the core repulsion functions of the Modified Neglect of Diatomic Overlap (MNDO) [74, 75] semiempirical method. In their approach, the total energy of the molecule $E_{\text{tot}}^{\text{mol}}$ was given by the sum of the electronic E_{el} and the repulsions $E_{\text{AB}}^{\text{core}}$ between the cores of atoms A and B (see equation (1.12)),

$$E_{\text{tot}}^{\text{mol}} = E_{\text{el}} + \sum_{\text{AB}} \sum E_{\text{AB}}^{\text{core}} \tag{1.12}$$

but the various terms in the Fock matrix related to the electronic energy and the core repulsions were not evaluated analytically. These terms were determined either from experimental data or from semiempirical expressions which contained numerical parameters which can be adjusted to fit experimental data. Then, the difference between AM1 and MNDO is that the latter utilizes equation (1.13) while the former uses equation (1.14) to compute the $E_{\text{AB}}^{\text{core}}$ energy.

$$E_{\text{AB}}^{\text{core}} = Z_A Z_B \gamma_{SS} \left[1 + e^{-\alpha_A R_{AB}} + e^{\alpha_B R_{AB}}\right] \tag{1.13}$$

$$E_{\text{AB}}^{\text{core}} = Z_A Z_B \gamma_{SS} \left[1 + F(A) + F(B)\right] \tag{1.14}$$

where

$$F(A) = \exp(-\alpha_A R_{\text{AB}} + \sum_i K_{\text{A}_i} e^{L_{\text{A}_i}(R_{AB} - M_{\text{A}_i})^2}), \tag{1.15}$$

$$F(B) = \exp(-\alpha_B R_{\text{AB}} + \sum_j K_{\text{B}_j} e^{L_{\text{B}_j}(R_{AB} - M_{\text{B}_i})^2}), \tag{1.16}$$

γ_{SS} are the two-electron, two-center integrals, and Z_A and Z_B the atomic number of atoms A and B, respectively. All other values are parameterized.

On the other hand, OMx methods have an explicit representation of Pauli exchange repulsion, which was shown to improve the description of conformational properties and noncovalent interactions of the systems. Furthermore, the OMx-D [76] strategy is to keep the OMx formalism and parameters unchanged, adding the effects of empirical dispersion terms (see section 1.8 on page 22).

Finally, the PM6 and PM7 methods cover essentially the whole periodic table and can be used to compute both molecular and solid-state properties. They are basically a MNDO method with the E_{AB}^{core} energy improved, just like AM1. In the

case of PM6, for example, the core repulsion functions are defined by equation (1.17).

$$E_{\mathrm{AB}}^{\mathrm{core}} = Z_A Z_B \gamma_{SS} \left[1 + x_{AB}\, e^{-\alpha_A (R_{AB} + 0.0003\, R_{AB}^6)}\right] \tag{1.17}$$

1.4 Hartree-Fock method

One of the simplest and widely known *ab initio* method to solve the electronic Schrödinger equation $\psi_{elec}(\mathbf{r}, \mathbf{R})$ is the Hartree-Fock (HF) method [77]. It is based on a simple approximation in which the electronic wave function is given by a Slater determinant [78] of N spin-orbitals χ_i

$$\psi^0(\boldsymbol{\chi}_1, \boldsymbol{\chi}_2, \ldots, \boldsymbol{\chi}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\boldsymbol{\chi}_1) & \chi_1(\boldsymbol{\chi}_2) & \cdots & \chi_1(\boldsymbol{\chi}_N) \\ \chi_2(\boldsymbol{\chi}_1) & \chi_2(\boldsymbol{\chi}_2) & \cdots & \chi_2(\boldsymbol{\chi}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_N(\boldsymbol{\chi}_1) & \chi_N(\boldsymbol{\chi}_2) & \cdots & \chi_N(\boldsymbol{\chi}_N) \end{vmatrix} \tag{1.18}$$

where the variables $\boldsymbol{\chi}$ include both coordinates of space and spin.

The differential equations for finding the Hartree-Fock orbitals have the form

$$\hat{\mathcal{F}}(i)\chi_i = \epsilon_i \chi_i \tag{1.19}$$

where χ_i is the ith spin-orbital, the operator $\hat{\mathcal{F}}(i)$ is the effective mono-electron Hartree-Fock Hamiltonian known as the **Hartree-Fock operator**, and the eigenvalue ϵ_i is the orbital energy of spin-orbital i.

$$\hat{\mathcal{F}}(i) = -\frac{1}{2}\nabla_i^2 - \sum_k^M \frac{Z_k}{\mathbf{r}_i - \mathbf{R}_k} + \nu^{HF}(i) \tag{1.20}$$

The term $\nu^{HF}(i)$ is the average potential experienced by the ith electron due to the presence of the other electrons. Essentially, the HF method replaces the many-electron problem by a one-electron problem in which the electron-electron repulsion is treated in an averaged way. Since the HF potential $\nu^{HF}(i)$ depends on the spin-orbitals of the other electrons, equation (1.19) is not linear and must be solved iteratively. This iterative procedure is called the self-consistent field method.

In 1951, Roothaan proposed representing the Hartree-Fock orbitals as linear combinations of a complete set of known functions, called basis functions. The accuracy of the HF calculation will be highly determined by the size of the basis set

selected to expand the spin-orbitals, but even for an infinite basis set there will be a difference between HF limit and the exact energy of the system. This energy is know as **correlation energy**, and usually represents the 1% of the total energy.

$$E_{corr} = E_{exact} - E_{HF} \tag{1.21}$$

This quantity is considerably large for the study of chemical processes, specially when weak interactions as Van der Waal's are predominant in the system. There exist many alternatives to surpass this inconvenient, DFT being one of them.

1.5 Density Functional Theory

In the basic formulation of DFT the energy functional is written as

$$E[\rho] = T_e[\rho] + V_{Ne}[\rho] + V_{ee}[\rho] \tag{1.22}$$

where T_e is the sum of the electronic kinetic energy, V_{Ne} is the interaction energy of the electronic cloud with an external potential, and V_{ee} is the electron-electron repulsion. However, only the second term in equation (1.22) can be easily calculated:

$$V_{Ne}[\rho] = \int \rho(\mathbf{r})\, \nu(\mathbf{r})\, d\mathbf{r} \tag{1.23}$$

being $\nu(\mathbf{r})$ generally the potential associated with the atomic nuclei.

An strategy to overcome this issue is to re-write equation (1.22) as

$$E[\rho] = T_S[\rho] + V_{Ne}[\rho] + J[\rho] + (T_e[\rho] - T_S[\rho] + V_{ee}[\rho] - J[\rho]) \tag{1.24}$$

introducing the functional $J[\rho]$ which represents the interaction energy of a classical electron cloud with itself:

$$J[\rho] = \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\, \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}\, d\mathbf{r}' \tag{1.25}$$

$T_s[\rho]$ is the kinetic energy of electrons in a fictitious system which has the same electron density ρ as the real system, but in which there is no electron-electron interactions. Also, the term in parenthesis in equation (1.24) is defined as the exchange-correlation energy functional of an interacting system with density $\rho(\mathbf{r})$:

$$E_{xc}[\rho] = T_e[\rho] - T_S[\rho] + V_{ee}[\rho] - J[\rho] = \int \rho(\mathbf{r})\, \epsilon_{xc}(\rho(\mathbf{r}))\, d\mathbf{r} \tag{1.26}$$

where $\epsilon_{xc}(\rho)$ is the exchange-correlation energy per electron of a uniform electron gas of density ρ.

Finally, substituting equations (1.23), (1.25), and (1.26) into equation (1.24), and subject to the condition

$$\int \delta\rho(\mathbf{r})\, d\mathbf{r} = 0 \tag{1.27}$$

it's possible to obtain

$$\int \delta\rho(\mathbf{r}) \left\{ \varphi(\mathbf{r}) + \frac{\delta T_S[\rho]}{\delta\rho(\mathbf{r})} + \mu_{xc}(\rho(\mathbf{r})) \right\} d\mathbf{r} = 0, \tag{1.28}$$

where

$$\varphi(\mathbf{r}) = \nu(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r} \tag{1.29}$$

and

$$\mu_{xc}(\rho) = \frac{d\left(\rho\, \epsilon_{xc}(\rho)\right)}{d\rho} \tag{1.30}$$

is the exchange and correlation contribution to the chemical potential of a uniform gas of density ρ.

In agreement with Hohenberg et al. [65], given φ and μ, it's possible to obtain the $\rho(\mathbf{r})$ which satisfy these equations by solving the one-particle Schrödinger equation

$$\left\{ -\frac{1}{2}\nabla^2 + [\varphi(\mathbf{r}) + \mu_{xc}(\rho(\mathbf{r}))] \right\} \psi_i^{KS}(\mathbf{r}) = \epsilon_i\, \psi_i^{KS}(\mathbf{r}) \tag{1.31}$$

considering

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i^{KS}(\mathbf{r})|^2 \tag{1.32}$$

and N the number of electrons in the system.

All this formalism presented so far has to be solved self-consistently, beginning with a guessed value of $\rho(\mathbf{r})$ which makes possible to construct $\varphi(\mathbf{r})$ and μ_{xc} from equations (1.29) and (1.30), respectively. Then, by using equations (1.31) and (1.32) a new and improved $\rho(\mathbf{r})$ is determined, leading to another cycle of the self-consistent field.

Note that the KS orbitals ψ^{KS} are not real orbitals, and thus they do not correspond to any real physical system. Their only function in the theory is to provide an adequate mapping between kinetic energy and density.

However, the exchange-correlation energy ϵ_{xc} is still missing from the analysis, and it is needed to compute equation (1.30). Clearly stated, the analytical form of this quantity is, to this day, unknown. In fact, due to the complexity of the

exchange-correlation term, only guesses have been made, and nobody knows what the exact expression for ϵ_{xc} is. Let us explore some of these approximations.

Local density approximation (LDA)

This approximation is based on the assumption that electrons could be analyzed as an homogeneous gas in a box with periodic boundary conditions. This way it is possible to establish a dependence of ϵ_{xc} on the dimensions of the box and the properties of the gas. Note that the electron density in a molecule is inhomogeneous, but within an infinitesimal of the volume it can be assumed as homogeneous. Then, the exchange-correlation energy can be obtained as a product of the infinitesimal volumes and the exchange-correlation energy density from the homogeneous gas theory [79, pp. 687].

Due to the local character of the approximation is why it is known as Local Density Approximation (when the ρ dependence is used), or Local Spin Density Approximation (LSDA) when the dependence is further split into α and β spin functions ρ_α and ρ_β, respectively.

General gradient approximation (GGA)

The following approximations take a step further compared to LDA, and assume that ϵ_{xc} may not have only a local dependence, but also may depend on the vicinity of the analyzed point. This means that $\rho(\mathbf{r})$ will also be dependent of its gradient at point $\mathbf{r}$:

$$\epsilon_{xc}^{GGA} = \epsilon_{xc}^{LSDA} + \int B_{xc}(\rho_\alpha,\, \rho_\beta, \nabla\rho_\alpha,\, \nabla\rho_\beta)\, d\mathbf{r} \tag{1.33}$$

where B_{xc} is carefully selected to maximize correspondence between experiment and theory [80].

The GGA approximation poorly describes the Van der Waals and other long range interactions due to its quasi-local character. This represents a problem when systems as biomolecules and catalyst are studied, due to the importance of these weak interactions for the associated chemical processes [47, 81, 82].

Hybrid functionals

In order to improve the description of chemical systems, and to surpass the inconveniences of the HF and DFT methods, hybrid approaches have been developed [83]. The hybrid functionals are composed of a combination of exchange-correlation terms

from other functionals and theories, as well as parametrizations from experimental observations. Despite this mixed nature might rise some suspicion, they have been proven as a pragmatic solution capable to describe at an acceptable extent complex chemical processes in organic molecules.

The B3LYP [84–87] exchange-correlation functional posses a mixed nature, as it is composed by small portions of energies computed with different methods [79, pp. 689].

$$\epsilon_{xc}^{B3LYP} = \epsilon_{xc}^{LSDA} + 0.20(E_x^{HF} - E_x^{LSDA}) + 0.72E_x^{B88} + 0.81E_c^{LYP} + 0.19E_c^{VWN} \quad (1.34)$$

It includes LSDA energy, the Hartee-Fock exchange energy (E_x^{HF})[1], the Becke [88] exchange potential (E_x^{B88}), the Lee-Young-Parr correlation potential (E_c^{LYP}) [89, 90], and a further addition of the Vosko-Wilk-Nusair potential (E_c^{VWN}) [91].

Another widely used functional is M06-2X [92], which is a member of successive families of functionals developed by Truhlar and co-workers. It is a highly parameterized hybrid meta-GGA functional intended for main-group thermochemistry and noncovalent interactions. The expression of the hybrid exchange-correlation energy can be generally written as follows:

$$E_{xc}^{hyb} = \frac{X}{100}E_x^{HF} + \left(1 - \frac{X}{100}\right)E_x^{DFT} + E_c^{DFT} \quad (1.35)$$

where E_x^{HF} is the nonlocal Hartree-Fock exchange energy, X is the percentage of Hartee-Fock exchange in the hybrid functional, E_x^{DFT} is the local DFT exchange energy, and E_c^{DFT} is the local DFT correlation energy.

For the specific case of the M06-2X functional, $X = 54$, the correlation energy is the sum of opposite-spin ($\alpha\beta$) and parallel-spin ($\alpha\alpha$ and $\beta\beta$) electrons terms, as represented in equation (1.36)

$$E_c^{M06-2X} = E_c^{\alpha\beta} + E_c^{\alpha\alpha} + E_c^{\beta\beta}, \quad (1.36)$$

being $E_c^{\sigma\sigma}$ highly parameterized integrals [92].

Finally, the exchange energy is defined by the integral represented in equation (1.37)

$$E_x^{M06-2X} = \sum_{\sigma} \int \left[F_{x\sigma}^{PBE}(\rho_\sigma, \nabla\rho_\sigma)\, f(w_\sigma)\right] dr \quad (1.37)$$

where $F_{x\sigma}^{PBE}(\rho_\sigma, \nabla\rho_\sigma)$ is the exchange energy density of the PBE [80] exchange

[1] In fact, this HF energy term is in reality the KS exchange energy, as the KS Slater determinant is used and not HF's.

model, and $f(w_\sigma)$ is the spin kinetic-energy-density enhancement factor.

1.6 Basis functions

Most quantum mechanics methods begin the calculation by choosing a set of n basis functions $\{\chi_j\}$, which are used to express the Molecular Orbitals (MOs) ϕ_i as a linear combination of atomic orbitals (LCAO)

$$\phi_i = \sum_j^n c_{ij}\chi_j \tag{1.38}$$

where c is a coefficient associated to each basis function. The functions $\{\chi_j\}$ are commonly represented using either Slater type orbitals (STOs) or Gaussian type orbitals (GTOs). The use of each type of orbital depends on the system's particularities. STOs are physically the best choice as they are solutions to the Schrödinger equation of hydrogen-like atoms, and decay exponentially far away from the nucleus. However, calculations of integrals over STO functions require considerable computational time. On the other hand, GTOs lack physical meaning but are more computationally efficient.

Nowadays, there exist hundreds of basis sets, and an agreed terminology is used to describe and identify them:

- A **minimal** basis set consists of one STO for each inner-shell and valence-shell AO of each atom.
- A **double-zeta** (DZ) basis set is obtained by replacing each STO of a minimal basis set by two STOs that differ in their orbital exponents ζ.
- A **triple-zeta** (TZ) basis set replaces each STO of a minimal basis set by three STOs that differ in their orbital exponents.
- A **split-valence** (SV) basis set uses two (or more) STOs for each valence AO but only one STO for each inner-shell (core) AO. An SV basis set is minimal for inner-shell AOs and double zeta (or triple zeta or ...) for the valence AOs.
- Split-valence sets are called **valence double zeta** (VDZ), **valence triple-zeta** (VTZ), ... according to the number of STOs used for each valence AO.

Other terms can be added to the basis sets to account more quantum effects. For example the polarization of the AO due to their distortion in shape and centers

of charge are taken into account adding STOs basis functions whose l quantum numbers are greater than the maximum l of the valence shell of the ground-state atom. This is represented by one or two stars accompanying the basis set name, like 6-31G** (note that 6-32G(d,p) is an equivalent notation).

It is also possible to improve the description of anions, lone pairs, Hydrogen Bonds, etc., by improving the description of the AO at large distances from the nuclei. This is achieved with the use of diffuse functions, which are denoted with one or two plus sign by the side of the basis function name (6-31+G, for example). One "+" sign means that diffuse functions will be included in all atoms but hydrogen, whereas two "+" signs means that also in the hydrogen are included the diffuse functions.

1.7 Basis set superposition error

The energy resulting from the interactions between the fragments of a system is known as binding energy BE. Let's imagine a system composed by two fragments A and B, then the binding energy can be expressed as

$$BE = E_{AB}^{ab} - (E_A^a + E_B^b) \tag{1.39}$$

where E_{AB}^{ab} is the energy of the dimer AB at its minimum energy computed using the total basis set (denoted as ab). E_A^a and E_B^b are the minimum energies of the monomers A and B, computed utilizing only the basis set of each fragment a and b, respectively.

However, in the study of weakly bound clusters, an artificial strengthening of intermolecular interactions is often observed, and the BE can not be correctly predicted: this phenomena is known as "basis set superposition error" (BSSE) [93]. As a monomer A approaches monomer B, the dimer can be artificially stabilized as monomer A utilizes the extra basis functions from monomer B to describe its electron distribution, and vice versa. Clearly, this only takes place when the monomers are at a relatively close distance, as at long distances the overlap integrals are too small and thus the stabilization effect is lost. This distance dependent inconsistency is the source of the BSSE. Even if this inconsistency could be eliminated, there would remain errors due to the fact that the basis set is incomplete (known as "basis set incompleteness error" (BSIE)). Note that the only way to perfectly prevent this errors is using a complete basis set, which is not possible. To solve this problem, Boys et al. [94] proposed a method called "Counterpoise correction" (CP).

The BSSE can be computed as

$$BSSE = E^{ab}_{A^*} + E^{ab}_{B^*} - (E^{a}_{A^*} + E^{b}_{B^*}) \tag{1.40}$$

where $E^{ab}_{A^*}$ is the energy of monomer A at its equilibrium geometry in the dimer AB, computed using the dimer basis set ab and taking the monomer B as an arrangement of ghost atoms (similarly for the $E^{ab}_{B^*}$ term). $E^{a}_{A^*}$ is the energy of the monomer A at its equilibrium geometry in the dimer AB using the basis set a (similarly for the $E^{b}_{B^*}$ term).

Then, the corrected binding energy using the Counterpoise method and considering evolution in the monomers geometries as they approach to each other can be expressed as

$$\begin{aligned} BE^{CP} &= BE - BSSE \\ &= E^{ab}_{AB} - (E^{a}_{A} + E^{b}_{B}) - (E^{ab}_{A^*} + E^{ab}_{B^*}) + (E^{a}_{A^*} + E^{b}_{B^*}) \end{aligned} \tag{1.41}$$

Equation (1.41) can be simplified if the geometries of the monomers do not evolve after the adsorption, and it's re-written as:

$$BE^{CP}_{no\ evol} = E^{ab}_{AB} - (E^{ab}_{A^*} + E^{ab}_{B^*}) \tag{1.42}$$

1.8 Empirical dispersion corrections

Despite the selection of a large basis set, DFT methods are not capable to correctly describe noncovalent interactions, which are related to long range electron-correlation. However, chemical phenomena as enantioselectivity could be related to long range interactions between molecules bulky groups [6]: attractive (p-stacking, van der Waals, etc.) or repulsive (conventional steric repulsion). It is clear that further improvements to DFT in order correct the dispersion-related energy are needed, and the use of dispersion-corrected models is mandatory [95–98] for an improved description of noncovalent interactions (specially Van der Waals forces).

One approach to include dispersion correction in DFT computations is the use of the M06 family methods, which are heavily parameterized to describe these kind of interactions. Another approach is to add a dispersion energy term to the energy computed with any other DFT functional.

Stefan Grimme et. al. [99–101] have proposed a dispersion corrected energy de-

scribed by the following equation

$$E_{DFT-D3(BJ)} = E_{KS-DFT} - E_{disp}, \tag{1.43}$$

where E_{KS-DFT} is the usual self-consistent KS energy obtained by the DFT method used, and E_{disp} is the dispersion correction as a sum of two- and three-body energies, then

$$E_{disp} = \frac{1}{2} \sum_{n=6,8} \sum_{A \neq B} s_n \frac{C_n^{AB}}{R_{AB}^n} f_{d,n}(R_{AB}) + \frac{1}{6} \sum_{A \neq B \neq C} \frac{C_9^{ABC}}{R_{ABC}^9} f_{d,9}(R_{ABC}, \theta_{ABC}) \tag{1.44}$$

In equation (1.44), Becke-Johnson (BJ) damping functions $f_{d,n}$ [102–104] are utilized to avoid (i) near singularities for small R_{AB} and (ii) double-counting effects of correlation at intermediate distances. Furthermore, s_n are the scaling factors; C_n^{AB} are the dispersion coefficients of nth order between atoms A and B (similarly for C_n^{ABC} in the case of three atoms); R_{AB} is the distance between atoms A and B; θ_{ABC} is the internal angle of the triangle ABC; and R_{ABC} is the mean distance of the three distances between atoms A, B, and C. The coefficients C_6^{AB} are estimated from a set of Time Dependent Density Functional Theory (TD-DFT) [105] computations, and the C_8^{AB} coefficient are computed from C_6^{AB} values.

The damping functions are described by

$$\begin{aligned} f_{d,n=6,8}(R_{AB}) &= \frac{R_{AB}^n}{R_{AB}^n + f^n} \\ f^n &= a_1 \sqrt{\frac{C_8^{AB}}{C_6^{AB}}} + a_2, \end{aligned} \tag{1.45}$$

where a_1 and a_2 are free fit parameters introduced by BJ.

In the literature, these corrections are specified following the methodology employed. For example, if the hybrid B3LYP functional is corrected with the third revision of Grimme's dispersion terms and the original damping is used the methodology would be referred as B3LYP-D3. Furthermore, if the BJ damping functions are involved, it would be called B3LYP-D3(BJ).

1.9 Method and the basis set selection

With all the possible combinations of methods and basis sets currently available to carry out DFT computations, the correct selection to describe the system under investigation can be a tricky task, and at the moment there is no universal

method/basis set combination.

In the study of small metal-catalytic systems, for example, Dub et al. [54] used M06-2X with Grimme's D3 correction and the def2-TZVP basis set [106, 107] to study the Noyori-Ikariya asymmetric transfer hydrogenation of ketones; Wang et al. [55] used HF/6-31G and a single point refinement with B3LYP/6-31+G** for the study of the asymmetric addition of diethylzinc to benzaldehyde; Pounder et al. [108] used B3LYP-D3(BJ) functional with the double-ζ def2SVP basis set [106, 107] in the study of the mechanism and origin of enantioselectivity in the rhodium-catalyzed asymmetric ring-opening reactions of oxabicyclic alkenes with organoboronic acids; and Wen et al. [52] studied the enantioselectivity of the asymmetric [3+2]-annulation between N-methylindole and enoldiazoacetamide catalyzed by prolinate-coordinated dirhodium using a DFT/Semiempirical approach, with B3LYP-D3(BJ)/6-31G(d) in the reactive region of the system and PM6-D3 in the other atoms.

In organocatalysis, Wåhlander et al. [109] have used M06/6-31+G* and a single point energy refinement with 6-311+G* basis set in the investigation of phosphoramides as an interesting alternative to the thiourea motif in organocatalysts; Sakai et al. [110] employed the B3LYP-D3(BJ)/6-311++G** and M06-2X/6-31G** chemical models to develop studies on the asymmetric Michael Addition of α-aminomaleimides to β-nitrostyrenes using an organocatalyst derived from *Cinchona* alkaloid; and the list grows in number and variety [111–118, for example].

According to the studies presented in the previous paragraphs, combinations of B3LYP, the M06 family, and double- and triple-ζ basis sets are without a doubt the more recurrent model chemistry used in organo- and metal-catalysis. Furthermore, Burns et al. [119] carried out an extensive benchmarking of several model chemistry tailored mainly to noncovalent interactions, arriving at very similar conclusions.

Finally, regarding the selection of the basis set, Witte et al. [120] proved that even the massive aug-cc-pV5Z basis was plagued by BSSE. Thus, a double-ζ basis set, including polarization and diffuse terms, in conjunction with D3 corrected functionals will be enough for the description of catalytic and noncovalent interactions dominated systems.

1.10 The influence of a solvent

In chemistry, a solution is a special type of homogeneous mixture composed of two or more substances. In such a mixture, a solute is a substance dissolved in another substance, known as a solvent. For example, the homogeneous catalysis is a process where the catalysts and the substrates are in the same phase, and they both

constitutes a solute which is dissolved into a solvent. The selection of the solvent has the capacity to influence properties such as chemical reactivity and molecular association throughout solubility, stability and reaction rates. Thus, choosing the appropriate solvent allows for thermodynamic and kinetic control over a chemical reaction, which is closely related to the performance of a catalyst.

To deal with the influence of a solvent in Computational Chemistry, two main strategies can be followed. First, the solvent molecules can be explicitly represented in the system, which of course will considerably increase the time and resources in quantum computations, therefore, this approach is generally reserved for Molecular Mechanics treatments of the system. However, there is a second possibility, which do not represents the solvent explicitly, but treats it as a continuum medium influencing the solute. This way, the electronic structure problem for a molecule in a liquid is reduced to the size of the solute of interest. Continuum solvation models represent a solvated molecule inside a molecule-sized and -shaped electrostatic cavity which contains the largest possible part of the solute charge distribution within it, excluding the solvent [121]. In the electrostatic theory of dielectric media, the medium is associated with a relative permittivity, which is a scalar function of position for isotropic nonhomogeneous media. In practice, the relative permittivity is usually equated to the bulk solvent static dielectric constant outside the solute cavity and to a smaller value inside the cavity.

Several methods have been designed to develop the continuum solvation formalism in practical computations [122], as for example the Polarizable Continuum Model (PCM) [123–125], the Onsager method [121, 126–130] and the SMD method [131].

1.11 Global optimization and organocatalysis

In the theoretical investigation of relatively small molecules, a good initial guess of the system's geometry is easily constructed with current state of the art software. This initial guess can be further perfected through a local minimum optimization procedure applied on the potential energy surface obtained from the atomic coordinates of the molecule. However, when an assembly of few to several millions of structural units such as independent atoms or molecules is meant to be studied (such assembly is known as chemical cluster [132, 133]), the initial position of each structural unit with respect to the others can not be guessed any more, as it would constitute a significant oversimplification of the problem.

In a cluster, there exists numerous ways of arranging the structural units that

leads to local minimum on the PES. The most abundant isomer in the ensemble of structures at low temperature will be the local minimum with lowest energy, which is known as global minimum. Finding the global minimum of a chemical cluster is what Global Optimization (GO) methods [134] are designed to accomplish.

As the number of minima rises exponentially with increasing cluster size, finding the global minimum can be very difficult. It has been found that "traditional" Monte Carlo (MC) and Molecular Dynamics (MD) Simulated Annealing (SA) approaches often encounter difficulties finding global minima for particular types of interatomic interactions such as the short ranged Morse potential [135]. For this reason, other types of algorithms, like genetic and basin hopping algorithms [136, 137], for example, have found increasing use in the area of finding global minima for clusters.

A global minimization should be able to find the globally minimal function value, irrespective of the starting point. To be sure of having found the global optimum of a function, one must have visited all search space. For continuous variables, and without further "tricks", this would entail an infinite number of function evaluations already in a one-dimensional case. Hence, global optimization is all about finding clever ways to "judge" finite patches of search space without visiting all points in them.

In practice, it turns out that deterministic global minimization often is extremely expensive, to the point of infeasibility [134]. Fortunately, stochastic algorithms using certain problem-solving heuristics can frequently be employed to quickly find minima with function values not too far above the global minimum in a small fraction of the time needed for a deterministic search for the true global minimum.

Marchal et al. [138] proposed an algorithm for finding the most stable isomers for any clusters, called Global Search Algorithm for Minima exploration (GSAM). It has two main steps: (i) the generation of the initial guesses, and (ii) the selection scheme, and full optimization of the stable structures followed by the computation of their vibrational properties. The scheme used in this study to obtain the initial guesses is based on the generation of several possible shapes for a given cluster, followed by the selection part of the algorithm applied to those generated structures.

The selection part is composed by three partial optimizations, each of them having normally allowed between 30 to 50 optimization cycles following the Berny optimization [139, 140] algorithm as implemented in the Gaussian 16 Rev. C.01 code [141]. Immediately after each optimization step, two discarding processes are applied to the structures. First, the structure with lower energy is taken as reference, and an energy threshold is applied to all other structures, removing those that do

not meet the established criteria. Generally, the energy threshold starts at around 40 $kcal/mol$ in the first partial optimization, and it is decreased to $\sim$ 20 $kcal/mol$ and 10 $kcal/mol$ in the second and the third partial optimizations, respectively. It is generally observed that this process removes at least the 30% of structures that entered it.

In second place, the normed quadratic difference between two geometries inside a given set of structures (D_{ij}) is computed following equation (1.46), furthermore discarding a substantial number of structures which would have lead to the same stationary point ($\sim$ 5%), based on the criteria presented in equation (1.47) and the Minkowsky metric [142].

$$D_{ij} = \left(\sum_{\alpha} \frac{|Q_{i\alpha} - Q_{j\alpha}|^2}{\max_{kl} |Q_{k\alpha} - Q_{l\alpha}|^2} \right)^{1/2} \tag{1.46}$$

$$\text{structure i with respect to structure j} \begin{cases} \text{discarded if } D_{ij} < 1\% \\ \text{selected if } D_{ij} > 1\% \end{cases} \tag{1.47}$$

The remaining configurations at the end of the 3 partial optimization processes are then fully optimized and their harmonic frequencies are computed in order to ensure their stability.

This GSAM algorithm has been implemented in the house made code also called GSAM (Inter Deposit Digital Number: FR.001.450001.000.S.P.2010.000.31235), and has been proven as an efficient global search algorithm of minima exploration for finding the most stable isomers of chemical clusters, validating the algorithm and code with a test-case in the Si_n (n=3,15) cluster [138]. Further publications [143–146] related to the structures of nucleic acid bases demonstrated that the algorithm is indeed well suited for the study of small to medium organic systems, successfully finding the global minimum for them. The GSAM code will be used in the present study to find the most stable arrangements of our targeted molecular cluster, which is composed by four fragments: the catalyst, the achiral Brønsted base, and the two substrates. Note that precisely due to the small number of fragments in our system, Global Optimization approaches like the Genetic Algorithm [147] could not be efficiently used, as they are designed for larger numbers of fragments.

Despite (or maybe due to) GO of clusters employing *first principle methods* for the computation of energy is much more challenging than the local optimization of chemical structures, the field have witnessed a vertiginous development in the last 30 years [148]. Nowadays, several software, in addition to GSAM, include the ability

to carry out GO: PDECO [149], GEGA [150], CALYPSO [151], and ABCluster [152, 153] are just some examples.

The application of these algorithms have been gratefully welcomed in several research fields as the exploration of new materials [154–157], the calculation of thermodynamic properties [158–160], and revealing catalysis mechanisms both in homogeneous [161–163] and heterogeneous [164–166] systems. Normally the global optimization process is carried out employing a force field to speed up the localization of the global minimum, followed by a quantum mechanics computation which refines the final structures for the determination of properties.

Additionally, sub-fields of global optimization have evolved, being Molecular Docking one of the most extensively reported in the bibliography. Molecular Docking is supported by the use of force fields in the computation of energy, and therefore finds applications in the study of nanoparticles [167], drug discovery [168], and many other areas related to large molecular systems. Some of the software employed to carry out molecular docking studies are GOLD [169, 170], AutoDock [171], and DARWIN [172].

In organocatalysis, Phillips et al. [173] made recently an extensive review on the relation between noncovalent interactions and catalytic properties. Furthermore, in 2015 Žabka et al. [174] published another review with the aim to give the investigators an overview of the methods utilized for mechanistic investigations in hydrogen-bonding organocatalysis. However, no reference to global optimizations were found in these reviews despite relatively complex catalytic systems were explored.

When studying the catalyst-substrates interactions, the chemical intuition in the positioning of the substrates with respect to chiral primary amine-guanidines lead Avila et al. [175] to theoretically support their investigation. They conclude that primary amine-guanidines, prepared by a simple monoguanylation of enantiomerically pure trans-cyclohexane-1,2-diamines act as organocatalysts in the enantioselective conjugate addition of isobutyraldehyde to nitroalkenes leading to enantiomerically enriched γ-nitroaldehydes. However, no global optimization techniques were utilized in the study for the identification of the PES global minimum.

Also Ding et al. [176] have recently demonstrated that H-bond donors can play a key role in controlling diastereoselectivity in a catalytic enantioselective reaction. They studied a highly enantioselective diastereodivergent asymmetric Michael addition of α-azido ketones to nitroolefins. Here, despite they theoretically studied reactions involved the catalyst and two substrates, one more time the global minimum was assumed from chemical intuition, and no global optimization was carried out. This pattern seems to repeat indefinitely in the literature [177–180].

A MM conformational search in the most important dihedral angles of the molecule to explore the global minimum is generally the method employed for investigations tailored to the study of the catalyst [181–185]. However, works where the global minimum is searched not only for the catalyst but for the catalyst-substrates system at the QM level of theory are missing from the literature, to the knowledge of the author. In that case, a conformational search generally would be an oversimplification as the system could have non-intuitive global minimum. A GO procedure will undoubtedly be the choice for relatively complex chemical clusters.

Furthermore, in the investigation of catalytic systems MM is not descriptive enough for the noncovalent interactions, which dominate and determine the molecular recognition process in organocatalysis. For this reason a DFT based GO is compulsory to obtain adequate results.

1.12 Procedure to identify the transition states

Once the reactants are identified by a Global Optimization and/or a Conformational Search process, to find the transition states an approximate geometry of them could be predicted from chemical intuition. Latter requesting an optimization to a first-order saddle point in the software employed would lead to the localization of the system's correct transition state geometry. However, taking into account the complexity of the systems here investigated and the number of bonds, angles, and dihedrals involved in the description of their geometry, predicting a transition state structure which is actually close to the real one is a very unlikely task to succeed on.

Nevertheless, a second approach is to perform a relaxed scan of the potential energy surface (V) along the two reactive atoms, which in our case will increasingly get closer the reactive carbons by steps of -0.05Å, generally starting from a distance around the 3.7Å (which is characteristic of the reactants geometries found in Chapter 2), and with a lower boundary correspondent to a short carbon-carbon bond (around 1.20Å). Successive shortenings of the C-C distance will start to transit an energy profile similar to that represented in Figure 1.1. This process will scale upwards the activation energy barrier, and once the quadratic region is reached after the change in convexity of the curve, the maximum energy geometry can be selected as a very close "guessed" structure to the system's transition state. Then, the first order saddle point of the potential energy surface is found from a Newton-Raphson procedure (see below) using the "guessed" structure as initial point.

The referred adjustment of the geometry is developed with an improved version

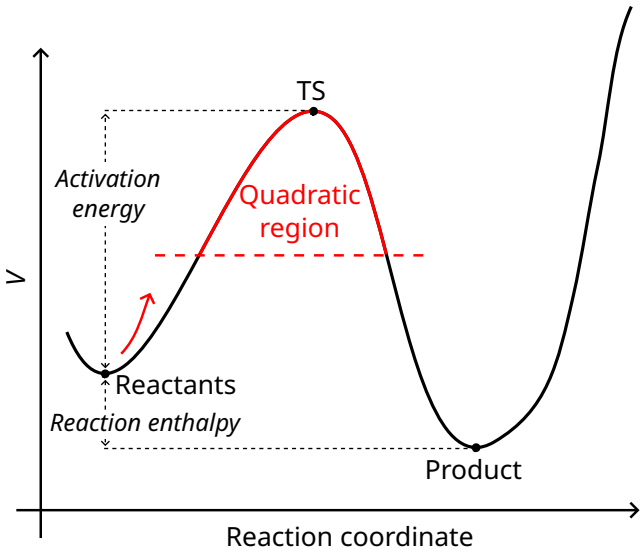


FIGURE 1.1. Schematic representation in 2 dimensions of the reaction coordinate of a generic simple reaction.

of the Berny algorithm [139, 140] in the Gaussian 16 Rev. C.01 software, which follows the general principles of finding a stationary point. In the first place, the Self Consistent Field (SCF) [186] (or some other) method to approximately solve the Schrödinger equation to find V at the guessed initial geometry is used. With these calculated values, the $3N-6$ nuclear coordinates are changed to a new set that is likely to be closer to the stationary point than the initial set, and the SCF V, its gradient $\vec{\nabla}V$, and its Hessian[2] are calculated at this new structure. Using the results of the new calculation, a further improved set of nuclear coordinates is calculated, and the SCF calculation is repeated at the new geometry. The process is repeated until $\vec{\nabla}V$ differs negligibly from zero, indicating that a stationary point (at which $\vec{\nabla}V$ is zero) may have been found.

An efficient way to find if $\vec{\nabla}V \approx 0$ is the Newton-Raphson method, which approximates the function by a Taylor-series expansion that is terminated after the quadratic terms, and uses accurately evaluated first and second partial derivatives of the function (which occur in the linear and quadratic terms of the Taylor series). However, due to the extremely large time required for the calculation of the Hessian matrix in first principle computations, modifications to the Newton-Raphson method are introduced in order to manage these inefficiencies. These methods are known as quasi-Newton-Raphson, and one does not calculate the Hessian directly, but instead starts with an estimation of it and gradually improves this estimation

[2]The Hessian of a function $f : \mathbb{R}^n \to \mathbb{R}$, with independent variable a vector $\mathbf{x} \in \mathbb{R}^n$ in such a way that $f(\mathbf{x}) \in \mathbb{R}$ and all its second derivatives exists, is a square $n \times n$ matrix formed by the set of f second derivatives $(\mathbf{H}_f)_{i,j} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}$.

using gradient information calculated at each step in the optimization cycle.

Finally, to identify what kind of stationary point has been found, it is essential to test the nature of the stationary point found by the geometry optimization. This is done by doing a vibrational-frequency calculation at the geometry found (i.e. using the Hessian matrix information). For a minimum, all the calculated frequencies will be real. For a first-order saddle point (transition state), one calculated frequency will be imaginary [56].

Now, if a close structure to the products of the reaction is known in addition to that of the reactants, it is also possible to employ in Gaussian either the QST2 or the QST3 (which additionally requires an initial guess of the transition state structure). These methods, implemented by H. B. Schlegel and coworkers [187, 188], uses a quadratic synchronous transit approach to get closer to the quadratic region of the transition state and then uses a quasi-Newton or eigenvector-following algorithm to complete the optimization. Nevertheless, practice and previous experiences in our group shows that the application of QST2 or QST3 methods is frequently not fruitful, because these methods only converge efficiently when provided with an empirical estimate of the Hessian and suitable starting structures in the Gaussian implementation, which are difficult to guess.

In this investigation, the approach utilizing the relaxed scan and the further user defined input of the guessed structure of the transition state resulting from it to the Berny algorithm implemented in the Gaussian 16 Rev. C.01 package will be employed.

1.13 Interaction forces in organocatalysis

Molecular structure is determined by covalent, noncovalent, and electrostatic interactions, the latter two being the driving force in most biochemical processes. Two or more atoms can share electrons to form covalent bonds, building molecules with totally different physical and chemical properties than the conforming parts. This kind of interactions were described by Lewis [189] even before the quantum theory of chemical bonds, and today we have a correct understanding of the underlying processes in the formation and rupture of covalent chemical bonds. On the other hand, two atoms/molecules can interact through a nonreactive channel, where bonds are neither made nor broken. This last kind of interactions is known as noncovalent interactions, and were first identified by van der Waals [190] in 1873. They play a determinant role [191] in the structure of liquids, solvation phenomena, structure of biomolecules such as DNA and proteins, molecular recognition processes, etc.

For example, in organocatalysis noncovalent interactions determine the outcome of organocatalytic reactions [82, 192–195], as they determine the molecular recognition process between substrates and catalysts. This class of interactions comprises a wide range of energies, and mainly encompasses hydrogen bonding, Van der Waals forces, and π-effects. Their difference with covalent interactions is that they do not involve the sharing of electrons between the atoms.

Hydrogen bonding

The dipole-dipole interaction established between an electronegative atom and an hydrogen atom which is bind to another electronegative atom is known as hydrogen bond (HB or H-bond). This interaction has predominantly an electrostatic nature, and its energy range is lower than covalent interactions, but higher than Van der Waals interactions. Nowadays the HB's nature is subject of an ongoing debate [196], and it may be seen as a complex interplay between different energy components: dispersion, orbital interactions, electrostatic interactions, etc.

Van der Waals interactions

A description for Van der Waals forces was first presented by J. D. van der Waals on its doctoral dissertation in 1873 [190]. They are a long range force present in all systems, and involves permanent or induced dipoles (or multipoles). Dipole dipole, dipole-induced dipole, and London dispersion are examples of Van der Waals forces.

π-effects

π-effects can be broken down into numerous categories, including π-π interactions, cation-π and anion-π interactions, and polar-π interactions. In general, π-effects are associated with the interactions of molecules with the π-systems of conjugated molecules such as benzene. These interactions are usually observed when the ring is rich (or lacking) in π electrons, and the second molecule (which can be another ring) is lacking (rich) in π electrons, and they approach enough each other.

1.14 Molecular recognition analysis and techniques

Theoretical computations from *first principle methods* require a post treatment of the estimated wave function to turn these computations into parameters able to

rationalize and understand the chemical behavior of the system. Usually, the information needed is no more than energies, molecular properties, and a description of the intra- and inter-molecular interactions.

Differently from covalent interactions, which are mostly established in the vicinity of a 2Å distance, the noncovalent interactions act on distances as large as tens of Angstroms, and their identification is not straightforward due to long-range intrinsic error related to current computational methods.

Several theories and methods have been developed to reveal the presence of noncovalent interactions in the analysis of molecular systems, as well as to identify, measure, and characterize their energy components. Computation of the electrostatic potential of the molecules, Non-Covalent Interactions (NCI) [81], Integrated Gradient Model (IGM) [197], Quantum Theory of Atoms in Molecules (QTAIM) [198–200], and Natural Bond Orbitals (NBO) [201–207] are just a few of the options, each of them having advantages, disadvantages, and limitations. A more detailed description of these methodologies can be found below, in section "Revealing noncovalent interactions".

Electrostatic potential

The electrostatic potential expresses the net electrical effect of the nuclei and electrons of a molecule at any point $\mathbf{r}$, and it is given rigorously by equation (1.48):

$$V^e(\mathbf{r}) = \sum_A \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|} - \int \frac{\rho(\mathbf{r}')\, d\mathbf{r}'}{|\mathbf{r}' - \mathbf{r}|} \tag{1.48}$$

being Z_A the charge on nucleus A, located at $\mathbf{R}_A$, and $\rho(\mathbf{r}')$ the total electronic density.

The sign of $V^e(\mathbf{r})$ in any particular region depends upon whether the effects of the nuclei or the electrons are dominant there. Sites reactive toward electrophiles can be identified and ranked by means of the local minima (V^e_{min}), whereas the maxima of $V^e(\mathbf{r})$ on the molecular surface ($V^e_{S,max}$) serve the same purpose for nucleophilic attack. Several investigations [208–210] have shown that V^e_{min} and $V^e_{S,max}$ correlate well with hydrogen-bond-accepting and -donating tendencies, respectively, in solute-solvent interactions. Also, they identified a set of molecular descriptors which can be related to the Electrostatic Potential Surface (ESP) values on the surface of the molecule: density, boiling point, surface tension, heats of vaporization and sublimation, impact sensitivity, diffusion constant, viscosity, solubility, solvation energy, etc.

Furthermore, Murray et al. [211] proved that the electrostatic potential on a molecular surface can also be used to characterize a quantitative measure of the degree of charge separation in the molecule, and its tendency to interact with other molecules.

Revealing noncovalent interactions

Many methods exists to evaluate the hydrogen bonds strength and separate each component contribution to its total energy, for example NBO, QTAIM, NCI, Symmtry-adapted perturbation theory SAPT [212–216], and more. However each of them has advantage, drawbacks, and specific applications.

For the description of hydrogen bonds, the use of NBO is not accurate in common situations, because the main nature of these bonds is electrostatic and this component can not be faithfully revealed by the $E(2)$ energy resulting from Second Order Perturbation Theory Analysis of the Fock Matrix in NBO Basis. Additionally, it has been proved by Stone [217] that the charge-transfer component of HBs cannot be correctly described by $E(2)$ value. However, it can still be used to measure the relative strength of stabilizing interactions due to orbitals overlap. For example, in the case of analyzing HBs, the interaction between the lone pair (LP, in NBO nomenclature) of an Oxygen donor and the σ^* orbital of a N–H bond (BD*, in NBO nomenclature) are those we would be interested in to search for $E(2)$ energies.

On the other hand, methods like NCI and IGM will provide accurate results describing noncovalent interactions and hydrogen bonds. The noncovalent interaction (NCI) method, which is also known as reduced density gradient (RDG) method, is very popular for studying weak interaction. In its approach, the first step is to find a way to distinguish weak interaction regions from other regions in the molecular system . From the table 1.1 it can be identified that if only the regions where the value of reduced density gradient function (see equation (1.49)) is in the range of $0 \sim$ medium are preserved, then "Around nuclei" and "Boundary of molecule" regions will be shielded.

For the remaining regions ("Around chemical bond" and "Weak interaction region"), if only the region where $\rho(\mathbf{r})$ is small is kept, then only weak interaction region will be revealed. With this region isolated, one can then assign a color scale based on the values of ρ and discriminate between strong attraction, Van der Waals interactions, and strong repulsion.

$$\mathrm{RDG}(\mathbf{r}) = \frac{1}{2(3\pi^2)^{1/3}} \cdot \frac{|\nabla\rho(\mathbf{r})|}{\rho(\mathbf{r})^{4/3}} \tag{1.49}$$

	Around nuclei	Around chemical bond	Weak interactions region	Boundary of molecule
$\|\nabla\rho(\mathbf{r})\|$	Large	$0 \sim$ minor	$0 \sim$ small	Very small $\sim$ small
$\rho(\mathbf{r})$	Large	Medium	Small	$0 \sim$ small
RDG($\mathbf{r}$)	Medium	$0 \sim$ minor	$0 \sim$ medium	Medium $\sim$ very large

Table 1.1. Criteria followed to isolate the noncovalent interactions from other kind of interactions present in a molecular system using the NCI method.

To describe the Integrated Gradient Model method, let's analyze the H_2 molecule. The atomic density in free-state of each atom along the molecular axis is shown as represented in Figure 1.2a. There, the gradient of the atomic density of the two atoms in the interatomic region have opposite signs, therefore, in the gradient of promolecular density (the curve g in Figure 1.2b), the contribution from the two atoms largely cancel with each other in the region between the two atoms. Note that at the midpoint of the two hydrogens, g is exactly zero, such point corresponds to bond critical point (BCP, further described below) in AIM theory under promolecular density.

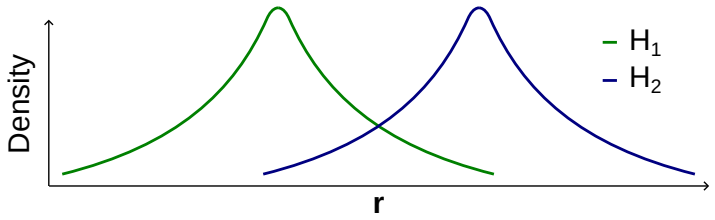
In Figure 1.2b, g^{IGM} is the IGM type of gradient, and it is calculated as sum of absolute value of density gradient of each atom in their free-states; in other words, phase is completely ignored and thus the density gradient originating from various atoms never cancel with each other. Due to this feature, g^{IGM} is upper limit of g.

Furthermore, the δg function in Figure 1.2b is defined as the difference between g^{IGM} and g. It can be seen that δg is non-zero in the atom interaction region, and has maximum value at the BCP position. Clearly, δg could be used to reveal interaction regions just like the reduced density gradient, and it is precisely the approach followed by the IGM method. In addition, the magnitude of δg in the interaction region has close relationship with the interaction strength. In general case, g^{IGM} and δg may be defined as follows:

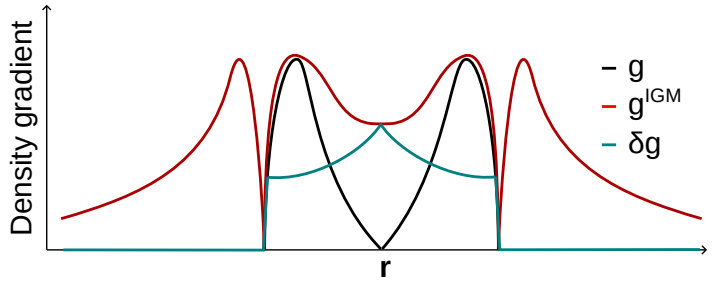
$$g(\mathbf{r}) = \left| \sum_i \nabla\rho_i(\mathbf{r}) \right| \tag{1.50}$$

$$g^{IGM}(\mathbf{r}) = \left| \sum_i \mathrm{abs}\left[\nabla\rho_i(\mathbf{r})\right] \right| \tag{1.51}$$

$$\delta g(\mathbf{r}) = g^{IGM}(\mathbf{r}) - g(\mathbf{r}) \tag{1.52}$$



(a) Atomic density



(b) Density gradients

FIGURE 1.2. Schematic representation of the treatment to the density in the Integrated Gradient Model.

In other direction, Bader proposed a topological analysis for analyzing electron density in "atoms in molecules" (AIM) theory, which is also known as "the quantum theory of atoms in molecules" (QTAIM) in its work "Atoms in molecules: a quantum theory" [198]. Further developments have extended this technique to other real space functions, like the first topology analysis research of electron localization function for small molecules presented by Silvi et al. [218]. In topology analysis language, the points at where the gradient norm of a function is zero (except at $\pm\infty$) are called critical points (CPs). Furthermore, CPs can be classified into four categories accordingly to how many eigenvalues of the Hessian matrix of the real space function are negative.

(3,-3): All three eigenvalues of the Hessian matrix of the function are negative, namely the local maximum. For the electron density analysis and for heavy atoms, the position of (3,-3) are nearly identical to nuclear positions, hence (3,-3) is also called nuclear critical point (NCP).

(3,-1): Two eigenvalues of Hessian matrix of the function are negative, namely the second-order saddle point. For electron density analysis, (3,-1) generally appears between attractive atom pairs and hence commonly called as bond critical point (BCP).

(3,+1): Only one eigenvalue of Hessian matrix of the function is negative, namely first-order saddle point (like transition state in potential energy surface). For electron density analysis, (3,+1) generally appears in the center of ring system

and displays steric effect, hence (3,+1) is often named as ring critical point (RCP).

(3,+3): None of eigenvalues of Hessian matrix of the function are negative, namely the local minimum. For electron density analysis, (3,+3) generally appears in the center of cage system (e.g. pyramid P_4 molecule), hence is often referred to as cage critical point (CCP).

In this work, our attention will be mainly attracted by BCP, as they are capable to identify the attraction between two atoms, and therefore the noncovalent interactions and hydrogen bonds. In fact, Emamian et al. [219] showed that the strength and energy composition of hydrogen bonds can be predicted based on electron density at the Bond Critical point (BCP) for a broad range of neutral and charged molecules. However, in some intramolecular weak H-bonds this method will not provide accurate results, as the BCPs are not always clearly identified [220], reason why it must be followed for other analysis capable to corroborate its results. Nevertheless, QTAIM has a strong practical significance in most cases and thus constitutes one of the most commonly used methods for H-bonds analysis.

The method proposed by Emamian et al. is reliable and the cost is very low. It is based in two equations for predicting H-bond binding energy (BE). For neutral complexes it is

$$\mathrm{BE} = -223.08 \times \rho_{\mathrm{bcp}} + 0.7423, \qquad (\mathrm{MAPE} = 14.7\%) \tag{1.53}$$

and for charged complexes

$$\mathrm{BE} = -332.34 \times \rho_{\mathrm{bcp}} - 1.0661, \qquad (\mathrm{MAPE} = 10.0\%) \tag{1.54}$$

being BE expressed in $[kcal/mol]$ and ρ_{bcp} in $[a.u.]$.

Here the mean absolute percentage error (MAPE) is used to portray the magnitude of error in the established linear regressions, and it is defined as

$$\mathrm{MAPE} = \frac{\sum \left| \dfrac{y_{act} - y_{pre}}{y_{act}} \right|}{n} \times 100\% \tag{1.55}$$

The criteria used to define the strength of the H-bond is presented in Table 1.2 for both neutral and charged systems.

However, this method to determine H-bond strength has some drawbacks. For example, the author selected a benchmark of small molecules where the interactions

Type of complex	Strength	BE	Major nature
Neutral	Very weak	$> -2.5\,kcal/mol$	Dispersion + Electrostatics
	Weak to medium	from -2.5 to $-14.0\,kcal/mol$	Electrostatics
Charged	Medium	from -11.0 to $-15.0\,kcal/mol$	Electrostatics
	Strong	$< -15.0\,kcal/mol$	Electrostatics + Induction

Table 1.2. H-bonds classification criteria for both neutral and charged complexes. Taken from [219].

between sites were only H-bonds, which could increase the error for larger molecules. Also the computations were made with the B3LYP-D3(BJ)/ma-TZVP methodology, thus the use of a much poorer model chemistry will also increase the error bar.

The second drawback is easily solved, by using an appropriate methodology. The first drawback, on the other hand, could be considerable difficult for large and complex systems. One viable approach is to determine the main nature of the interaction between two atoms using the ρ_{BCP} obtained from the QTAIM analysis, employing the Nakanishi criteria [221, 222], and then compute the H-bond binding energy only on those interactions which result as H-bonds.

Furthermore, if a deeper investigation of the interaction is needed, the SAPT method has to be used. It decompose the energy into four main components: exchange, electrostatic, dispersion, and induction. This way the contribution of each component can be measured and the nature of the interaction completely characterized.

Chapter 2

Molecular recognition processes

The objective of Chapter 2 is to describe at the DFT level of theory the molecular recognition processes between the foldamers and the substrates (see Figure 2.1), utilizing the tools provided by Computational Chemistry presented in Chapter 1.

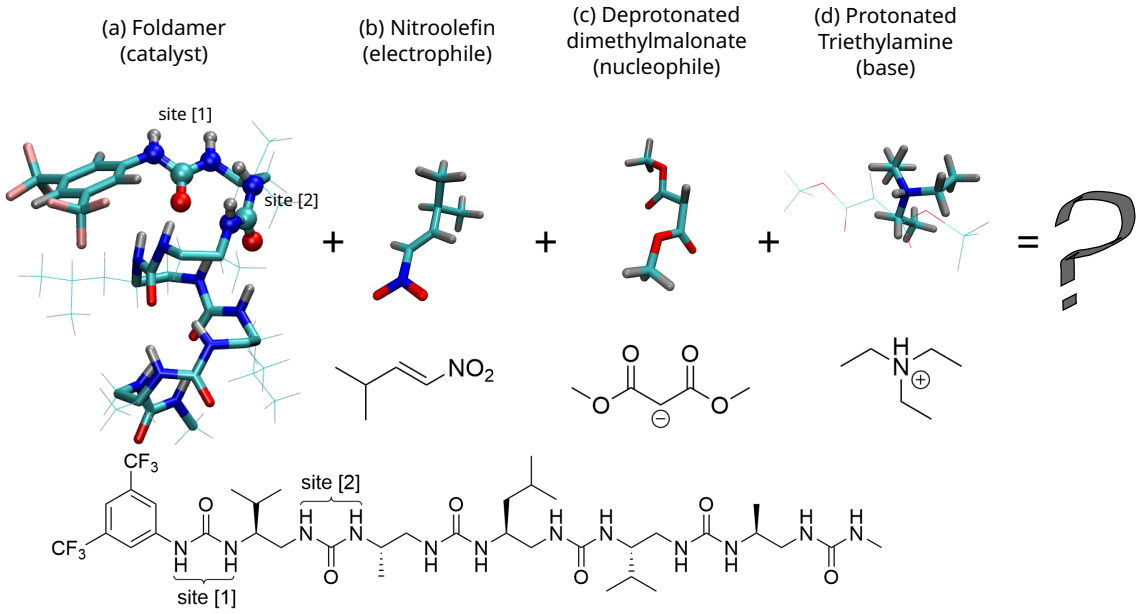


Figure 2.1. Structure of the foldamer-based chiral catalyst (hexaurea **2**) is represented in subfigure (a). The studied reaction involves the (1*E*)-3-methyl-1-nitrobut-1-ene as nitroolefin (b) and the deprotonated dimethylmalonate (c), the last being deprotonated by the triethylamine (d).

Published

Yaidel Toledo-González, Jean-Marc Sotiropoulos, Diane Bécart, Gilles Guichard, and Philippe Carbonnière. *Insight into Substrate Recognition by Urea-Based Helical Foldamer Catalysts Using a DFT Global Optimization Approach.* The Journal of Organic Chemistry 2022 87 (16), 10726-10735. DOI: 10.1021/acs.joc.2c00562

2.1 Introduction

Defined sequences and folding (e.g. protein tertiary structure) are two of the main attributes of biopolymers that determine their specific and highly diverse functionalities (transport, sensing, signaling, energy storage, catalysis, etc.). In enzymes, precise organization of active-site side chains in the 3D space, participation of co-factors, electrostatic interactions are inherent elements mediated by a folded yet dynamic backbone which contribute to facilitate the interaction with the substrate and stabilize transition states, ultimately leading to catalytic activity [223, 224]. Inspired by protein structures, small and medium size synthetic folded strands have received increasing interest as chiral catalysts because, on the one hand, they are modular and scalable and, on the other hand, they may retain some key features of enzymes such as high reactivity and stereo-selectivity but also chemo-selectivity and site-selectivity. This is particularly true for synthetic α-peptides which were found to be well-suited as chiral catalysts for a broad range of asymmetric transformations [24–27]. Concurently, foldamers (non natural oligomers with defined folding patterns) which exhibit both sequence tunability and folding predictability have gained increasing interest as scaffolds for catalyst design [225]. In these systems, the active site generally consists of an array of spatially arranged functional side chains and can also involve main-chain functional groups well-organized through folding. This is the case of chiral aliphatic N,N'-linked oligoureas, a class of helical foldamers previously reported to catalyze enantioselective C-C bond-forming reactions, in which main chain (thio)urea NH groups located at the positive pole of the helix macrodipole are available to bind and activate substrates [28, 29]. This system actually shows similarities with the well-studied helical polypeptides developed as catalysts for Juliá-Colonna epoxidations in which the four N-terminal amide NH groups not engaged in intramolecular H-bonds are involved as H-bond donor groups with the substrates [226, 227].

Oligourea hexamer **1** which contains seven urea linkages (Figure 2.2) was found to catalyze the conjugate addition of 1,3-dicarbonyl substrates to nitroalkenes in high yield and enantioselectivity at remarkably low catalyst/substrate molar ratios in the presence of an achiral Brønsted base such as Et_3N or DIPEA [29]. In this paper, it was observed that the chain length of the foldamer was a crucial parameter in terms of reactivity and enantiocontrol. Whereas the helically folded 6-residue long oligourea (**1**) and the related 5-mer (**2**) show remarkable catalytic properties, oligomers with shorter chain lengths were much less effective in terms of enantiocontrol. Structure-activity relationship studies further confirmed that the helix conformation of the oligourea foldamer is a major determinant of their catalytic

activities, and that the second (thio)urea site is as important as the first. Related hybrid oligomers consisting of a short oligourea segments fused at the N-terminus of a short helical peptide have also been used to catalyze similar reactions [42, 228, 229].

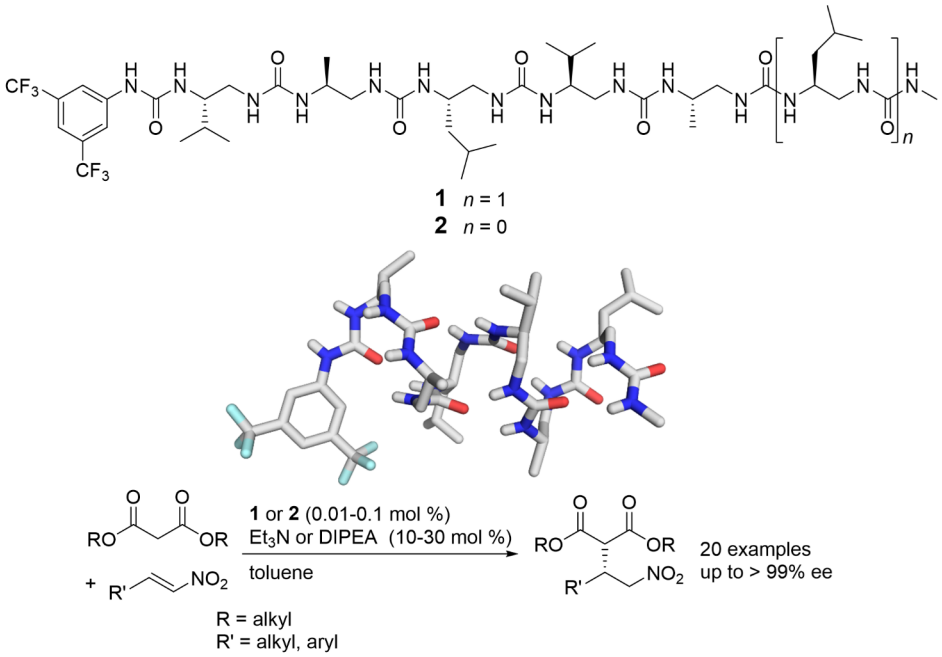


Figure 2.2. Formulae of helical foldamers **1** and **2**, X-ray structure of **1** and their use as H-bond donor chiral catalysts in the enantioselective conjugate addition of 1,3-dicarbonyl compounds to nitroalkenes [29].

Mono- and bis-(thio)urea H-bond donor catalysts [31–39] have been studied extensively and in many cases, detailed computational investigations have been conducted on these small molecules to explore the catalytic processes at the atomistic level, mainly by finding the related transitions states [36–38, 109–112, 114–116]. In contrast, the mechanism of synergistic activation with oligourea foldamer-based catalysts has not yet been investigated from a theoretical point of view to the knowledge of the author. Our aim in this chapter is thus to use computational chemistry to study the catalytic-binary system composed by the chiral H-bond donor foldamer catalyst/achiral Brønsted base, by asking the following questions: What kind of interactions occur between the reactants and the foldamer catalyst? What is the role of the helical shape in the catalytic process? Which is the influence of the catalyst on the capability of the reactants to react? To that end, the properties of the foldamer catalyst as a function of its chain length, as well as the interactions between the catalyst and reaction substrates were carefully investigated at the DFT level of theory.

Finally, it has to be pointed out that when enantioselective catalysis is the subject of an investigation, the Transition States related to each enantiomer are definitely a compulsory requirement. However, in molecular clusters, preliminary studies capable to identify the key Transition State are required; once it is located the calculation can be often limited to computing this particular species [6].

The molecular recognition is of central importance in bio- and organocatalysis, relying on non-covalent interactions such as hydrogen bonding, stacking, van der Waals interactions, etc. [230]. A deep understanding of the system's interactions and arrangements between the helical foldamer, the Brønsted base, and the two substrates is critical to obtain comprehensively characterization of the problem we are facing. From this information it will be possible to interrogate the system about enantiomeric excess (*ee*), reaction pathways, etc.

2.2 The computational method employed

The heptaurea (6-mer) was optimized from the experimental X-ray structures (CCDC accession codes 1534525) [29]. Starting with this 6-mer, the other structures were deduced as follows: from the heptaurea, we removed one residue at a time and optimized the resulting structure, obtaining this way the hexaurea (5-mer, **2**), pentaurea (4-mer), and so on. We then duplicated the backbone of the heptaurea and connected the two resulting helices, to generate the 12-mer molecule. Finally, we followed the previously mentioned procedure of removing residues one by one and optimizing until we obtained the octaurea (7-mer), completing this way a set of 12 foldamers (see Figure A.1 in Appendix A for their formulae).

The most stable binding sites were found from Global Optimization procedures (see section 1.11) employing the hexaurea **2**. Some additional possible conformers were found from the most stable species previously obtained, mainly focusing on the functional groups close to the reactive region of the catalysts, as for example the rotation of the 3,5-bis(trifluoromethyl)phenyl group. Molecular geometry optimizations were performed with the Gaussian 16 Rev. C.01 software [141], using DFT with the B3LYP-D3 functional and 6-31G(d,p) basis set. Additionally, a diffuse function was specified for O, N, F, and the negative charged carbon of the malonate anion atoms (hereafter 6-31"+"G(d,p)). The SMD variation of the Polarizable Continuum Model (PCM) using the integral equation formalism variant (IEFPCM) solvent model [131] was set in all the optimizations with toluene, and Solvent Accesible Surface (SAS) representing the solute-solvent boundary. The temperature in all simulations was set to 298.15 K, in agreement with the experiments.

The lowest energy structures of the systems composed of the malonote (i.e. dimethylmalonate) and/or nitroalkene (i.e. (1E)-3-methyl-1-nitrobut-1-ene) binding onto the hexaurea catalyst **2** [29] were obtained by the Global Search Algorithm of Minima (GSAM) interfaced with the Gaussian 16 program for computations of the energies and gradients. The GSAM procedure consists of a random process of structure generation followed by an iterative process. For each iteration the structures are partially optimized (typically 30 cycles of optimization) and sorted from their relative energy with respect to the most stable structure of the set and refined from their topological differences. Generally, the iterative process consists in 4 steps in which the relative energetic criterion that selects the structures for the next step is successively tightened, typically from 100 to 20 $kcal/mol$. This allows to decrease by 30-60% the number of the structures for the next step of optimization. The final structures are then fully optimized. Our choice to apply the B3LYP-D3 functional, together with the 6-31"+"G(d,p) basis set, stems from our previous experience, which showed that this combination leads to good agreement between experimental and theoretical vibrational spectra of different types of hydrates [143, 145] and structure of micro hydrated nucleic acid basis [144]. Nevertheless, regarding the effect of basis sets [120], functionals [119] and the estimation of interatomic molecular interactions by DFT, all the energies of the low-lying isomers reported in the present manuscript were fully re-optimized with the B3LYP-D3/6-311"+"G(d,p) and M06-2X/6-31"+"G(d,p) model chemistry.

Moreover IGM and QTAIM analysis were performed with Multiwfn [231] software, while NBO analysis was perform with Gaussian software.

2.3 Towards the global minimum of the system

The global optimization search (see section 1.11) which involves the helical foldamer, both reactants ((1E)-3-methyl-1-nitrobut-1-ene as electrophile and deprotonated dimethylmalonate as nucleophile) and the achiral base (triethylamine) was first performed using the hexaurea **2** (Figure 2.1). As depicted in Figures 2.3 and 2.4, which show the most energetically stable conformers, this numerical process found that only the two first urea groups (denoted as sites `[1]` and `[2]` in the figures) can accommodate the reactants.

Among the possible configurations, the site `[2]` is generally more favorable for the electrophilic species ((1E)-3-methyl-1-nitrobut-1-ene) and leads to a stabilization of about 18 $kcal/mol$ (Figure 2.3) while the nucleophilic species (deprotonated dimethylmalonate) tends to be astride on the two sites with a stabilization of about

31 *kcal/mol* (Figure 2.4). However, when the two reactants are considered together for binding on the foldamer, the nucleophilic species appears to bind clearly and exclusively on site [1], as stated in Figure 2.5. This clearly reveals a bifunctional activation of the catalyst from the two first sites.

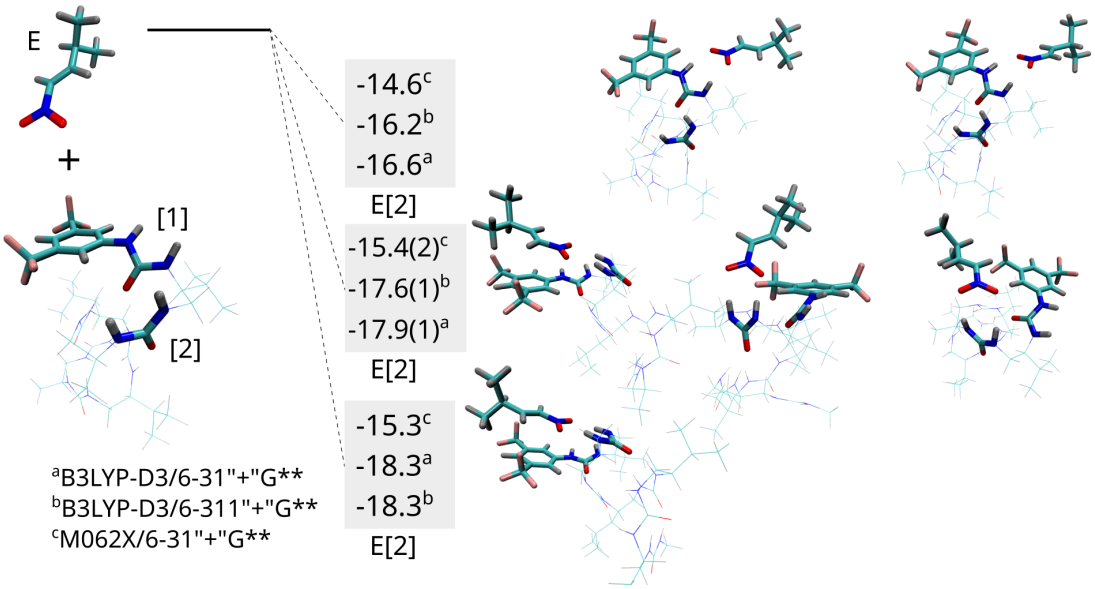


Figure 2.3. Binding energies (in kcal/mol) of the nitroalkene (electrophile) on the hexaurea catalyst **2**. Here [1] and [2] represent the two binding sites of the catalyst.

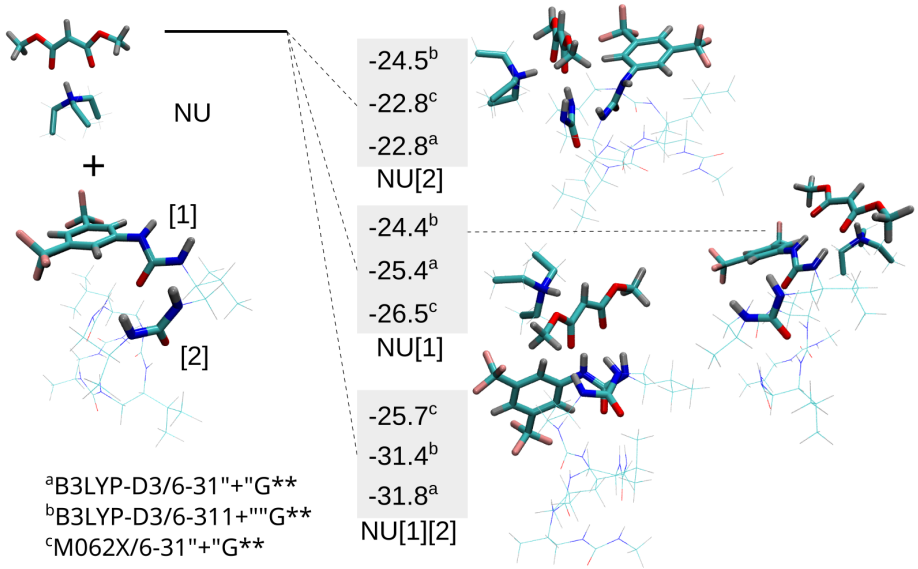


Figure 2.4. Binding energies (in kcal/mol) of the deprotonated malonate (nucleophile) on the hexaurea catalyst **2**. Here [1] and [2] represent the two binding sites of the catalyst.

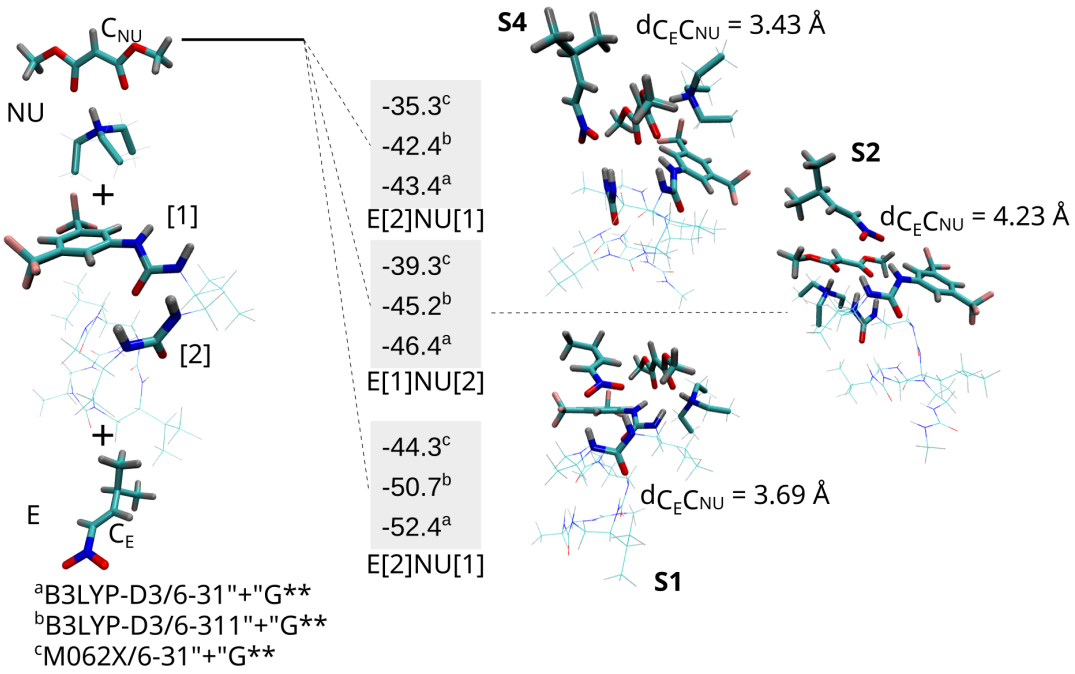


Figure 2.5. Binding energies (in kcal/mol) of the deprotonated malonate (nucleophile) and the nitroalkene (electrophile) on the hexaurea catalyst **2**. Here `[1]` and `[2]` represent the two binding sites of the catalyst.

The three structures of Figure 2.5 were obtained (i) by considering the most stable situation of Figure 2.4 (nucleophile in astride position) in which the nucleophile species was placed on sites `[1]` and `[2]` and (ii) by considering the most stable structure of Figure 2.3 in which the electrophilic species binds on site `[2]`. When the electrophile is placed in the vicinity of site `[2]`, it is then observed that the nucleophile moves form the astride position to site `[1]` during the local optimization process. Note that the three model chemistry used revealed the same pattern: the B3LYP-D3 results overbound for a few *kcal/mol* with respect to their M06-2X counterparts, and no sensitive deviations are found with the use of a triple zeta basis set as already found in the literature [119, 120]. Moreover, our computations show that the deprotonation of the malonate by the achiral base is sensitively favored in presence of the catalyst at 300 K (see Figure 2.6).

Overall, these results are consistent with experimental ^{1}H-NMR data of **1** obtained upon successive addition of malonate, triethylamine, and nitrostyrene by Diane Bécart, Gilles Guichard, and coworkers at the University of Bordeaux (see Figures A.4–A.8 in Appendix A). This shows that under the conditions used (for solubility issues, spectra were recorded in a mixture of acetonitrile-[d_3]/DMSO-[d_6] (88:12 v/v)) : (i) the deprotonated malonate interacts with the catalyst, more than the cognate malonate, (ii) this interaction involves site `[1]`, and (iii) triethylamine

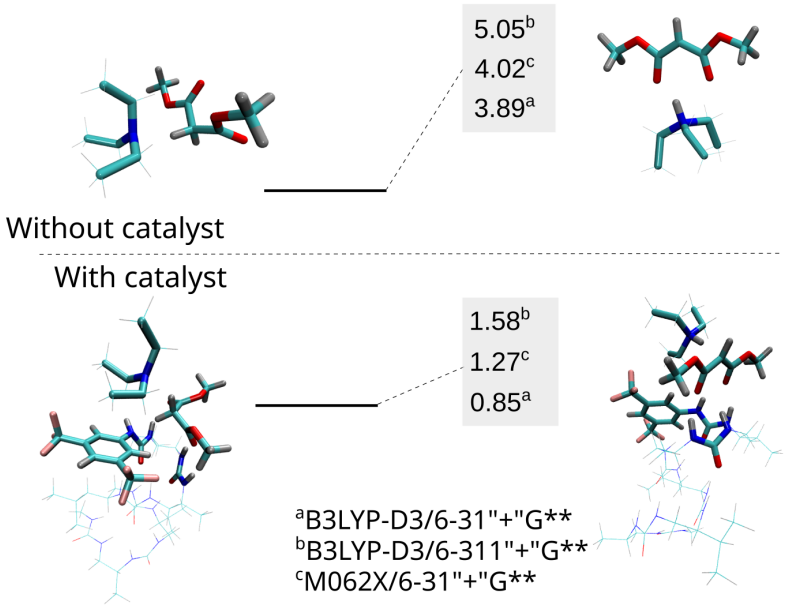


Figure 2.6. Discrepancies (in $kcal/mol$) between energies of the malonate and its deprotonated form in presence of its base, with and without the catalyst.

critically controls the interaction process because nitrostyrene and malonate alone do not lead to significant modification of the spectrum of **1**.

Furthermore, one can note that the most energetically favored structural position may also, from a structural point of view, favor C-C bond formation reaction between the two reactants since the corresponding C-C distance for this optimized structure is calculated at 3.69 Å.

Concerning the Basis Set Superposition Error (BSSE) associated to the B3LYP-D3/6-31“+”G(d,p) results presented in Figures 2.3–2.5, it was computed an average error of 7 $kcal/mol$ (see Table 2.1). Despite the BSSE value is relatively high, it is almost the same in all cases. Therefore, subtracting the BSSE energy from the Binding Energy reported in Figures 2.3–2.5 will result in almost the same binding energy values, and exactly the same order of the structures in terms of relative energy.

Finally, we tested the influence of the diffuse function on the results. To this end, we computed the values in Figure 2.5 with the B3LYP-D3/6-31G(d,p) methodology (see the last column in Table 2.1). It was observed that if the diffuse function is not present the changes in Binding Energy are considerable if compared with those obtained by B3LYP-D3/6-31“+”G(d,p). Taking also into account that the results with B3LYP-D3/6-311“+”G(d,p) fit their double-ζ counterpart, the fundamental role of the diffuse function on the correct description of these systems is revealed.

No. structure	B3LYP-D3/6-31“+”G(d,p)	BSSE	B3LYP-D3/6-31G(d,p)
1	-52.38	7.50	-64.86
2	-46.39	7.11	-57.37
3	-45.53	6.56	-58.92
4	-43.40	6.49	-54.17

Table 2.1. BSSE of the B3LYP-D3/6-31“+”G(d,p) model chemistry, and the influence of the diffuse function on the results presented in Figure 2.5 (structure 1 being the most stable). All quantities are in $kcal/mol$. Note that there are 4 structures reported in this Table, and only 3 structures appears in Figure 2.5. This is because structure 3 in this table has been removed from the figure in order to gain in clarity.

2.4 Nature of the interactions

We next wondered about the nature of the stabilizing interactions between the reactant and the catalyst, i.e. whether the interactions are more of orbital or electrostatic nature.

2.4.1 Non-covalent interactions

Figure 2.7 illustrates the non-covalent interactions between the foldamer and the different substrates obtained by IGM and QTAIM methods (see section 1.14). Four complexes were considered: a) the most stable structure of the nitroalkene on the catalyst, binding on site `[2]` ; b) the most stable structure of the deprotonated malonate on the catalyst, astride on the sites `[1]` and `[2]` ; c) the second most stable structure of the deprotonated malonate, binding on the site `[1]`; d) the most stable structure of the nitroalkene binding on site `[2]` and the malonate binding on site `[1]`.

The nitroalkene was found to bind to site `[2]` by means of two H-bonds, while a region of Van der Waals interactions rises around the 3,5-bis(trifluoromethyl)phenyl group (see Figure 2.7a). Differently, the malonate ester binds to the catalyst exclusively using an array of H-bonds, i.e. five for the complex shown in Figure 2.7b and two for the complex shown in Figure 2.7c. In any case, it is observed that the average interaction energy by hydrogen bond is about 4 $kcal/mol$, which is relatively strong and would have predominantly an electrostatic nature, as generally observed in biomolecules [81]. The results obtained with M06-2X were equivalent (see Figure 2.7d).

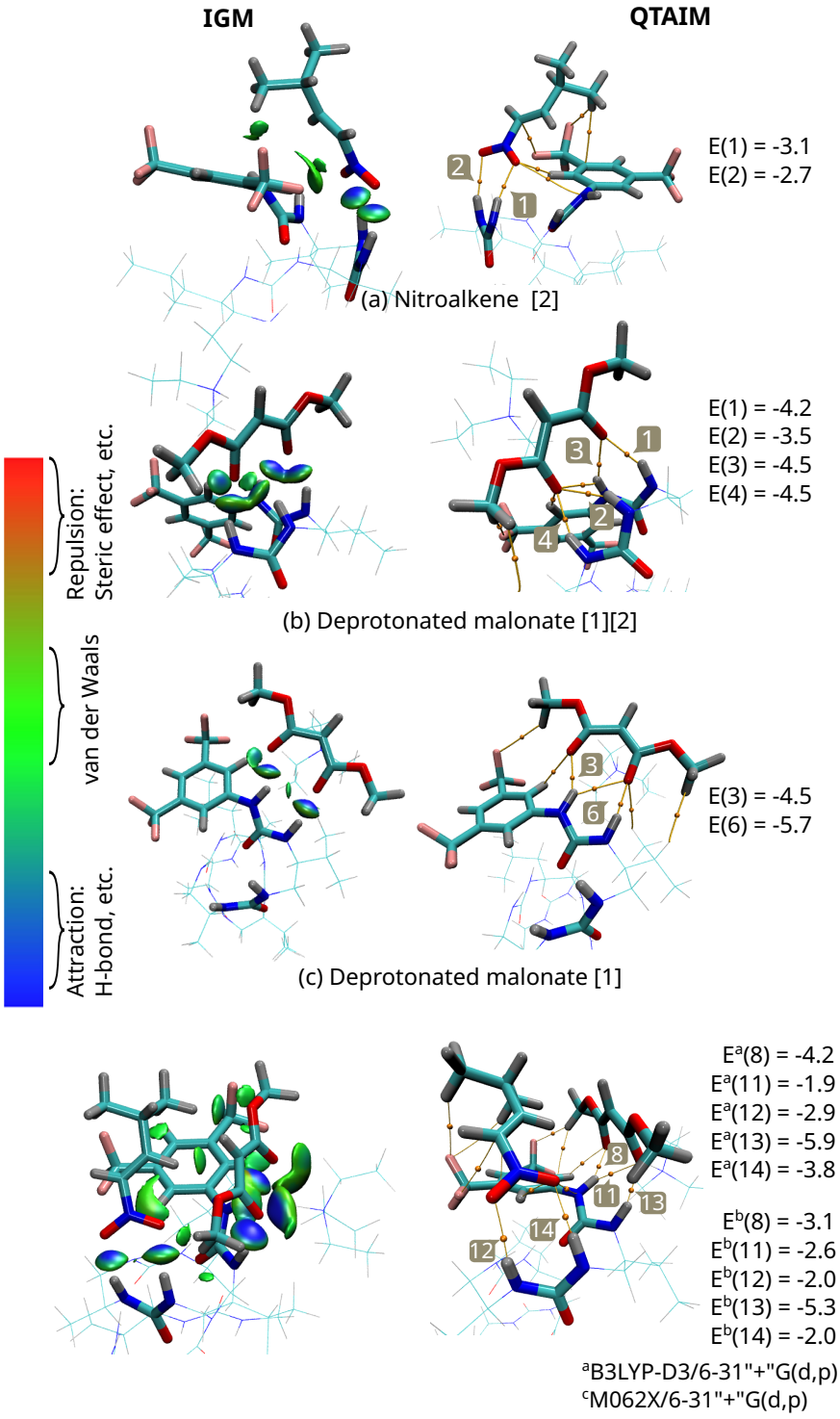


Figure 2.7. Non-covalent interactions in the most stable conformations found for the hexaurea foldamer **2** interacting with the substrates. Along with the QTAIM analysis the H-bond binding energies E (in $kcal/mol$) deduced from the electron density at the Bond Critical Point (BCP) following the method proposed by Emamian et al. [219], and discussed in section 1.14 on page 32, has been reported.

2.4.2 Orbital interaction

Concerning the investigation of the orbital interactions and according to the Frontier Orbital Theory [232], the energy of the HOMO of each substrate were compared to the LUMO of the oligourea foldamer electron acceptor (see Figure A.3). A HOMO(substrates) - LUMO(catalyst) energy gap of about 3.5 eV and 6 eV is observed for the malonate and the nitroalkene respectively. This value is not sensitively modified by the number of the residues of the foldamer (from monomer to hexamer) as shown in Figure A.3, thus suggesting that the change of affinity between the substrates and the foldamer with respect to the number of residues is not governed by orbital interactions. Although donor orbitals of the substrates and acceptor orbitals of the catalyst are quite far in terms of energy, it is observed (not depicted in the figure) a very small mixing between the orbitals of the two kinds that cannot be ignored since the decrease of each HOMO energy of the substrates when connected to the foldamer is about 0.45 eV. This decrease is constant whatever the number of residues of the foldamer.

Furthermore, to evaluate the stabilization energy due to orbital interactions, a Second Order Perturbation Theory Analysis of Fock Matrix in NBO Basis was carried out in the 3-mer (which gives low selectivity [29]), 5-mer (which gives high selectivity [29]), and 9-mer foldamers (which has not been studied experimentally yet) interacting with both substrates (see Figure 2.8 and Table 2.2). It is observed that the stabilizing interactions between the substrates and the foldamer are similar in the three systems, with two exceptions localized in the trimer: O86 to N15–H16, and O124 to N64–H65, which are $\sim$ 4 $kcal/mol$ and $\sim$ 3 $kcal/mol$ higher than their corresponding in pentamer and nonamer, respectively. Disregarding those two exceptions, only one interaction has a considerable stabilization energy due to orbital interactions in the three systems, and it is localized in one of the two acceptor dimethylmalonate's oxygen (see for example the interaction between O128 and N34–H35 in the pentamer of Figure 2.8b). Furthermore, it is interesting to point out that the stabilizing interactions are constant and do not depend on the catalyst's length.

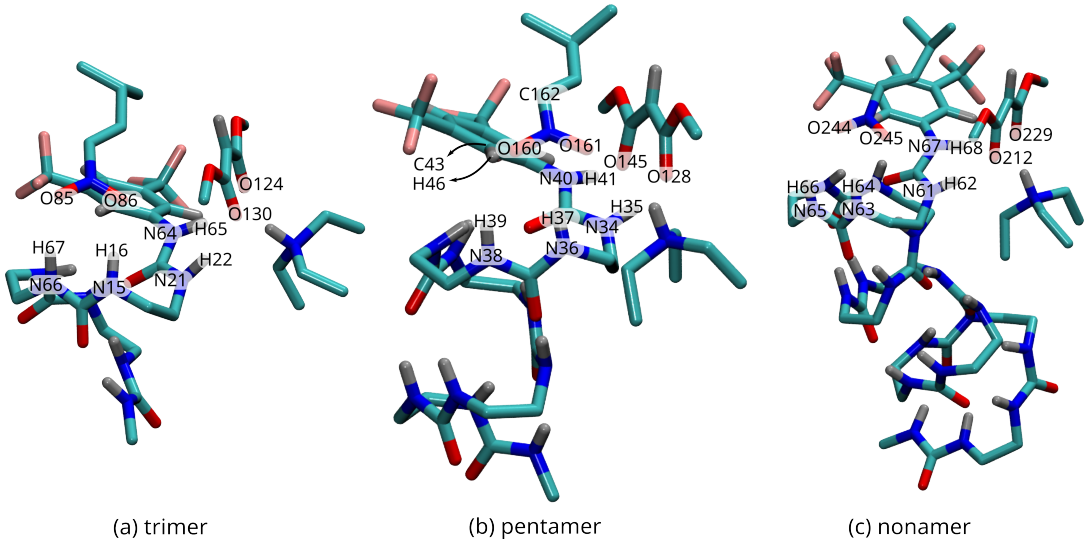


Figure 2.8. Numbering of the atoms involved in the NBO analysis presented in Table 2.2.

Donor (i)	Type	Acceptor (j)	Type	E(2) [$kcal/mol$]
		Trimer		
O85	LP(2)	N66-H67	BD*(1)	3.56
O85	LP(2)	N15-H16	BD*(1)	0.22
O86	LP(2)	N15-H16	BD*(1)	7.14
O124	LP(2)	N64-H65	BD*(1)	4.15
O130	LP(1)	N21-H22	BD*(1)	9.23
		Pentamer		
O160	LP(1)	N38-H39	BD*(1)	0.93
O160	LP(2)	N36-H37	BD*(1)	2.54
O161	LP(2)	N36-H37	BD*(1)	3.06
O145	LP(2)	N40-H41	BD*(1)	1.94
O128	LP(1)	N34-H35	BD*(1)	10.26
		Nonamer		
O244	LP(1)	N65-H66	BD*(1)	0.75
O244	LP(2)	N63-H64	BD*(1)	3.01
O245	LP(2)	N63-H64	BD*(1)	2.85
O229	LP(2)	N67-H68	BD*(1)	1.88
O212	LP(1)	N61-H62	BD*(1)	10.50

Table 2.2. Second Order Perturbation Theory Analysis of Fock Matrix in NBO Basis for the 3-mer, 5-mer, and 9-mer foldamers represented in Figure 2.8. E(2) is the energy of hyper conjugative interaction (stabilization energy), LP stands for Lone Pair, and BD stands for bond (all related to the NBO nomenclature).

2.5 ESP and reactive behavior

The theoretical estimation of the interaction strength between substrates and the foldamer, and the stabilization of charged intermediates arises from the evaluation

of the electrostatic potential and some quantities statistically related to it [211] (see section 1.14 on page 32 for more information). The ESP surfaces of the foldamers as a function of their chain length (monomer, dimer, hexamer and dodecamer) and their unfolded form are shown in Figure 2.9. A set of indexes that illustrates the reactivity of the catalyst with respect to the number of its residues is also reported: (i) the index of polarization (Π) which is the internal charge separation within the molecule and is related to its dipole moment (Figure 2.9g), (ii) the minimum and maximum value of the ESP (Figure 2.9h), (iii) the sigma indexes (σ^2_{tot}) that are used as indicators of the tendency of the molecule to interact by its positive (σ^2_+) or its negative (σ^2_-) region (Figure 2.9i).

From Figure 2.9g, it is clearly seen that the polarization increase with the number of residues when the catalysts is folded, while in the unfolded state it remains nearly constant and weaker. The polarization difference in each case is directly related to the intramolecular H-bonding cooperation between ureas, which is present in the folded catalyst but not in the unfolded form. This highlights the importance of the helical conformation. The Figure 2.9g also reveals three ranges of internal charge separation: (i) from monomer to trimer in which the positive region of the ESP surfaces encloses partially the binding sites `[1]` and `[2]`, (ii) from tetramer to heptamer for which the positive region fully enclose the binding sites and the polarization is slightly higher, (iii) from octamer to dodecamer for which the positive ESP region extends further outside the two reactive sites.

Furthermore the maximum value of the positive charge (see Figure 2.9h) which is centered on site `[1]` increases form the monomer to dodecamer. About 70% of the gain is reached for the tetrameric form and beyond in the series. This shows that the electron acceptor character of the foldamer increases with the number of residues and seems to extend beyond the site `[1]` and `[2]` for oligomers larger than the heptameric form (see the red line in Figure 2.9i).

Before the end of this section, a few words on the conformational analysis of the catalyst are pertinent, because a different conformation could change the electrostatic profile and the reactivity properties. In fact, the 1-mer and the 2-mer foldamers have been observed to be more stable when the second urea is oriented in the opposite direction with respect to the first urea (see Appendix B). This behavior is the consequence of the foldamer's length: with just 3 ureas in the foldamer, the third urea is connected to the first but the second is not connected to any other. Then, the establishment of H-bonds between the carbonyl group of urea 2 and the two NH of urea 1 will make the system considerably more stable. Note that in the case of the 3-mer (tetraurea) this will not be the case, as the second urea will be

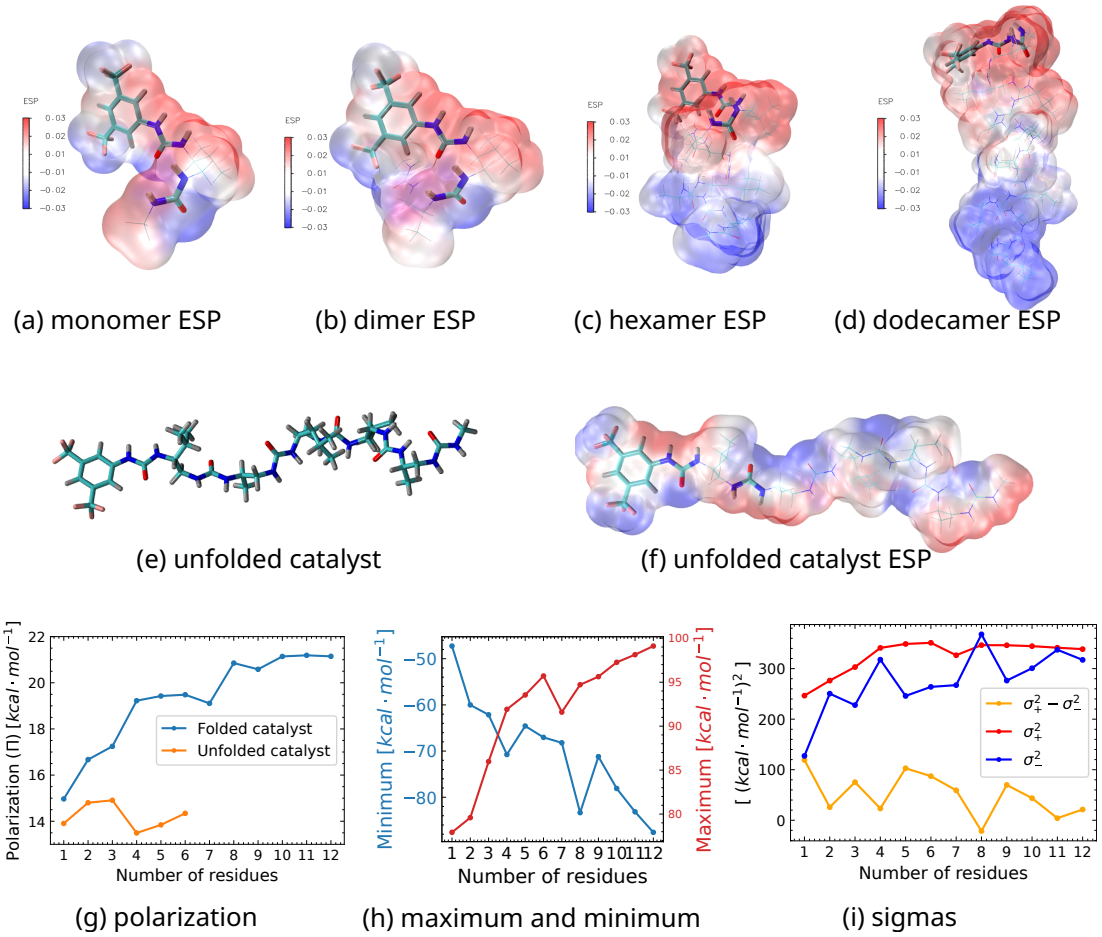


FIGURE 2.9. Catalyst electrostatic potential and related properties. Subfigures a–d represent the ESP surfaces of various folded catalysts with different chain lengths. Subfigures e and f represents the hexaurea **2** unfolded. Finally, subfigures g–i show interesting ESP statistically based quantities for the folded molecule, and for the unfolded oligomer when specified (see section 1.14 on page 32 for more information on this statistically based quantities).

connected to urea 4.

Figure B.3 on page 109 clearly states that site [1] will be occupied, as usual, for the dimethylmalonate, and site [2] by the 1(E)-3-methyl-1-nitrobut-1-ene, however the reaction pathway is considerably more complicated for the interaction, if compared with the geometries analyzed in Figure 2.5. It is possible then to conclude that despite being the geometry with the first two ureas oriented in different directions (↑↓) more stable than the geometry with the two ureas oriented in the same direction (↑↑), the reactivity will also considerably decrease for the ↑↓ system. Finally, the low reactivity and enantiocontrol observed in the experiment for the 1-mer and the 2-mer could probably be related to the orientation of the foldamer's first two ureas.

2.6 HOMO-LUMO gap of the reactants

According to the Frontier Orbital Theory [232], the reactivity between an electrophile and a nucleophile is governed by the HOMO-LUMO gap of the overall system. On that point, the Figure 2.10 illustrates the evolution of the HOMO-LUMO gap with respect to the number of the residues, when the reactants bind to the catalyst.

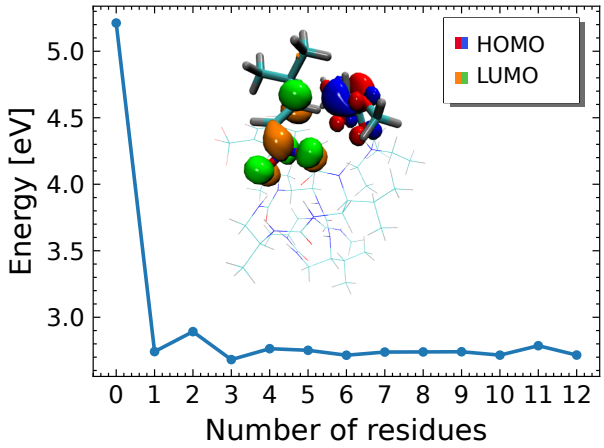


Figure 2.10. HOMO-LUMO gap behavior in the substrates as a function of the catalyst chain length. The HOMO was calculated for the deprotonated malonate ester (nucleophillic species), and the LUMO for the nitroalkene (electrophillic species).

It is clearly seen that the gap significantly decreases with the introduction of the catalyst in the system, from 5.21 eV to 2.74 eV, then it becomes nearly constant from monomer to dodecamer. This suggest two important points (at least), first, that the only presence of the 1-mer catalyst considerably lowers the HOMO-LUMO gap between the substrates, facilitating the reaction. Second, the changes in reactivity and enantioselectivity observed in the experiment [29], which were dependent on the catalyst's size, should not be related to its influence on the frontier orbitals of the substrates positioned this way, as their gap is constant no matter the foldamer's number of residues.

2.7 Conclusions

In this chapter it has been reported the first computational insights into the molecular recognition and catalytic properties of aliphatic N,N'-linked oligoureas (foldamers) which together with achiral Brønsted base catalyze the addition of malonate to nitroolefins at low loading (0.01–0.1 mol % in chiral catalyst). The present

work reveals that the foldamer interacts with and activate simultaneously both reactants and not only the electrophile (nitroalkene) as could have been anticipated when designing this synergistic catalytic system [29]. This computational analysis shows that the nitroalkene is bound to the second site, whereas the malonate ester is located at the first site, the two substrates being engaged in multiple H-bond interactions. It is pointed out that the distance between the two reactive carbons in the structures identified by the Global Optimization procedure were around 3.34Å and 4.23Å, specifically in the most stable arrangement the value was $d_{C_E C_{NU}} = 3.36$Å.

Orbital and electrostatic interaction are of equivalent magnitude but only the latter evolves with the size of the foldamer. The polarization of the foldamer was observed to increase, due to the internal urea-urea cooperation of the foldamers resulting from the helix conformation, which increases the capability to anchor the substrates, unlike its unfolded form. Therefore, it can be concluded that non-covalent interactions are predominant in the systems catalytic properties, similarly to the peptides and enzymes in which these foldamers are inspired on.

Finally, our analysis reveals that the reactivity of the reactants should increase in the presence of the catalyst as the difference between the highest occupied molecular orbital of the nucleophile (malonate) and the lowest unoccupied molecular orbital of the electrophile (nitroalkene) decreases by 2.47 eV when the catalyst is added, and place in a better suited position for the reaction to take place.

Overall, these results open the possibility of further designs and improvements useful to expand the scope of foldamer catalyzed reactions. In this respect, molecular modeling of energetic profile from localisation of the corresponding transition states to estimate the enantiocontrol of C-C bond forming reaction between (1E)-3-methyl-1-nitrobut-1-ene and dimethylmalonate, in presence of the oligourea foldamer catalyst, and as a function of its chain-length, is still needed and will be reported in the next chapter.

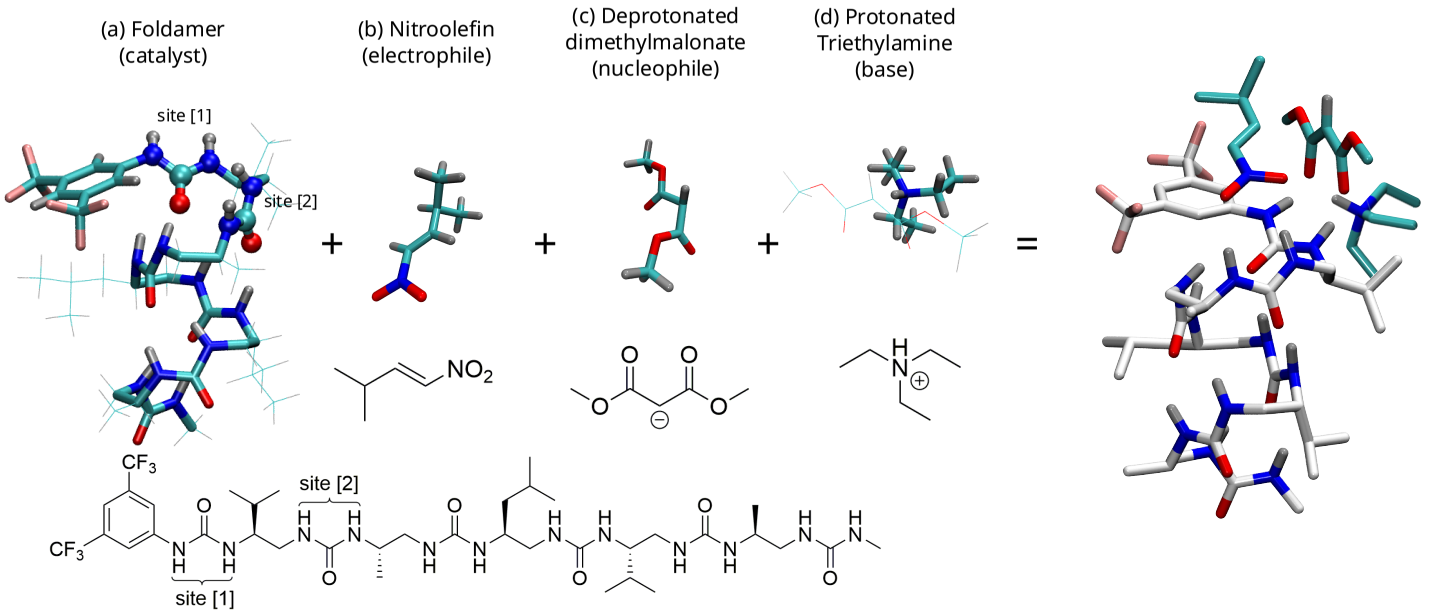


Figure 2.11. Most probable localization sites for the (1*E*)-3-methyl-1-nitrobut-1-ene and the deprotonated dimethylmalonate on the catalyst surface.

Chapter 3

Modelling of the energetic profile

The objective of Chapter 3 is to inquire the structures identified during the molecular recognition processes about the transition states of the systems, therefore obtaining a picture of their energetic profile. Starting from the reactants geometries, enantiomers R and S will be identified, to further compute the products of the reaction (see Figure 3.1). This data allows to obtain the activation energy and the reaction enthalpy, values defining the predominance of one enantiomer over the other, and the stability of the products, respectively. This information is closely related to the catalytic properties of the studied foldamers.

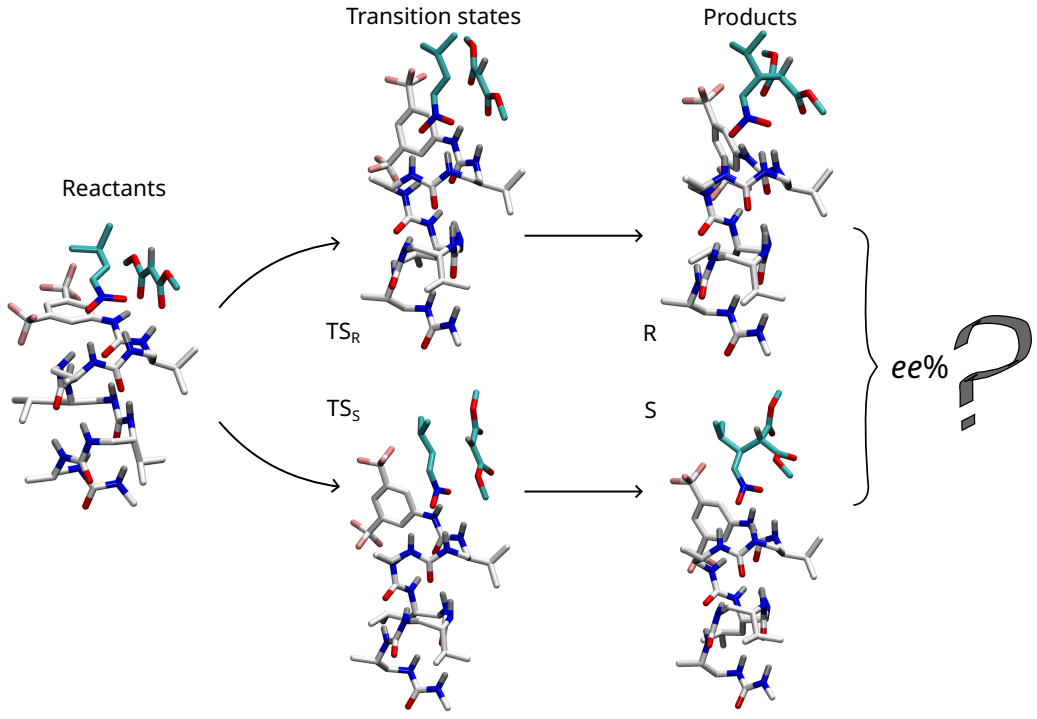


Figure 3.1. Schematic representation of the steps carried out during the modelling of the energetic profiles. In this case the hexaurea foldamer is represented, and the protonated triethylamine achiral base has not being depicted for the sake of clarity.

3.1 Introduction

The efficacy of an asymmetric catalyst in its function as supporter in the formation of one enantiomer over the other is directly related to the system transitions states, and thus to the system non-covalent interactions, which of course are crucial in the area of non-covalent asymmetric organocatalysis [47, 82, 193]. A comprehensive exploration of the system's non-covalent interactions has been presented in Chapter 2, and then the computational analysis of the energetic profiles will be the main topic of the present chapter.

From a computational point of view, the determination of the relative energies of the transition states of the possible reactive pathways that bring the reactant to become products is a needed study when investigating asymmetric catalysis. The activation energy barrier is the measurement computational chemists frequently employ in order to discriminate between different probable reaction pathways proposed by chemical intuition or obtained from previous computations. This barrier determines which enantiomer will be favored in the reaction, therefore, it is closely related to the experimentally obtained enantiomeric excess (*ee* %). However, the identification of the transition state of a reaction is not always a straightforward task.

As discussed at the end of section 2.1 on page 42, normally the most difficult part of determining the transition states of a reaction from a computational point of view is finding its geometry and identifying the path by which the reaction occurs. Nevertheless, this task is considerably facilitated with the determination of the reactants by a Conformational Search and/or a Global Optimization analysis.

Following the description of the foldamer-substrates molecular recognition processes and the identification of the key geometries of the reactants in Chapter 2, the molecular modelling of the energetic profile is considerably smoothed. From the geometry of the reactants, and in some cases also the products, several methods have been implemented in commercially available Computational Chemistry packages to find the transition states from this information. For example, in Gaussian 16 Rev. C.01 [141] it is possible to request an optimization to the transition state of a geometry rather than to the local minimum, using the Berny algorithm [139]. Then, knowing the reactants and with some chemical intuition to predict a close geometry to that of the real transition state will suffice to find it. Also, two different approaches of the STQN method [187, 188] are available, named QST2 and QST3. Further details on the description of these methods will be provided in section 1.12 on page 29.

Finally, to obtain a computational estimation of the relative activation energy

barriers characteristics of enantiomers R and S, and their correspondence with experimental results, the analysis should not be constrained to the hexa- and hepta-ureas, but ideally extended to a larger range of foldamers to identify the foldamer length/catalytic performance relationship observed experimentally. In this study, the foldamers from 1-mer to the 9-mer will be analyzed (see Figures A.1 and A.2), which will consequently provide an estimate of the enantiocontrol of the C-C bond forming reaction between (1*E*)-3-methyl-1-nitrobut-1-ene and dimethylmalonate in presence of the foldamer catalyst/achiral Brønsted base catalytic-binary system as a function of the number of residues of the foldamer.

3.2 The chemical model

As described in the introduction of this chapter, the catalysts from 1-mer to 9-mer will be taken into account in the exploration of the energetic profiles, with the objective to identify how the number of residues in the foldamers influences its catalytic properties. This means that the 9 reactant structures have to be optimized, and additionally for each enantiomer, 9 scan of the characteristic C-C bond coordinate, followed by the transition state search and the optimization to the products has to be carried out. The computational resources consumed by such benchmark of calculations could be considerably high. For this reason, a methodology with similar energies and reduced computational time to that so far employed in Chapter 2 would be desired.

Several models have been tested searching for a better suited methodology to carry out the computations. For example, computing the systems with B3LYP in the reactive region and a semi-empirical (and in other cases MM) method in the remainder residues of the foldamer. The results of this inquiry can be found in Appendix C on page 115. There, it was concluded that the model which provides accurate enough energies (error $\lessapprox 2\ kcal/mol$) and at the same time saves over the 60% of the computational time is that when the reactive region of the system is computed with B3LYP-D3/6-31G“+”(d,p) and the rest of the catalyst is computed with PM6. In Figure 3.2 it is represented with tubes the reactive region of the system computed by DFT, while the part drawn by lines is computed with PM6.

Concerning the identification of the enantiomers produced by the asymmetric catalyst in the studied reaction, the enantiomer R transition state geometry for the C-C bond formation between (1*E*)-3-methyl-1-nitrobut-1-ene and dimethylmalonate is represented in Figure 3.3a, and the enantiomer S is represented in Figure 3.3b. These geometries were obtained following the proceedings discussed in section 1.12:

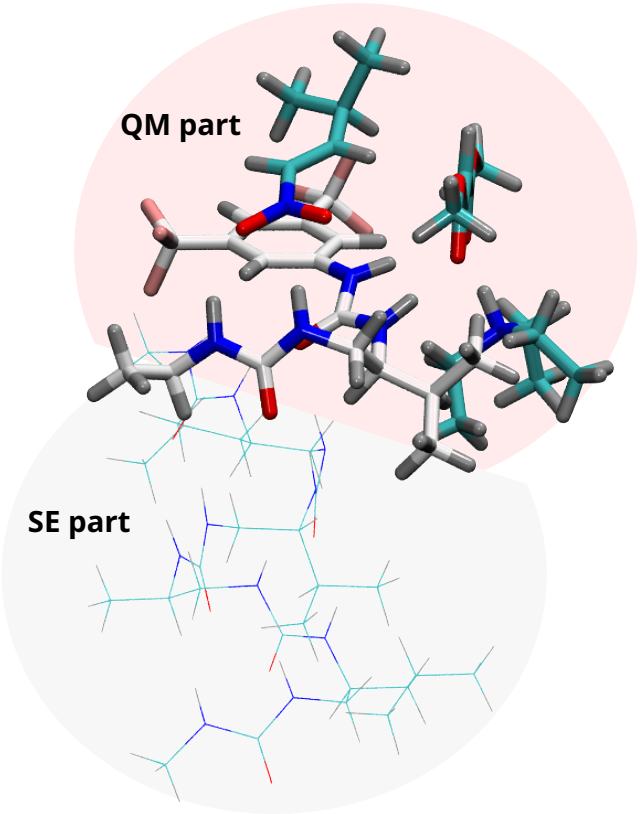
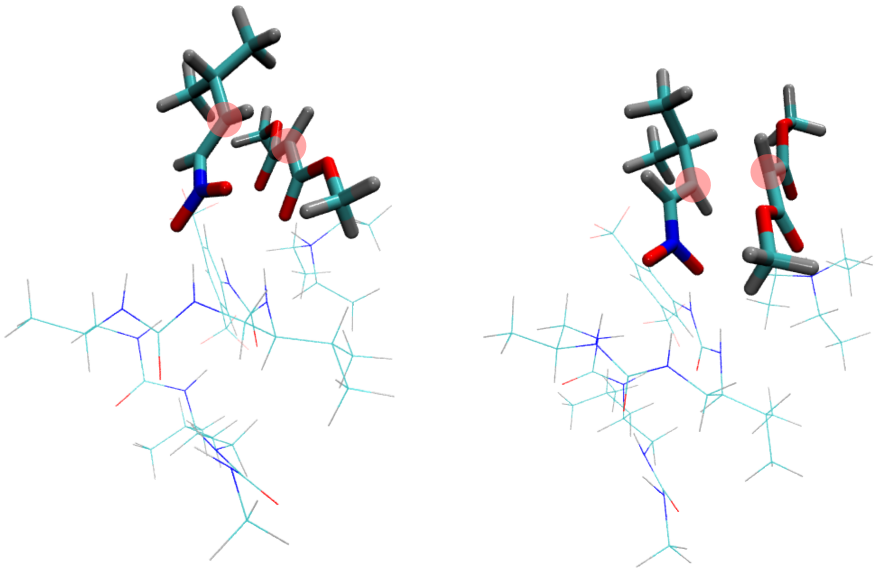


Figure 3.2. Representation of the chemical model used in the investigation of the energetic profiles characteristic of the reaction between substrates (1*E*)-3-methyl-1-nitrobut-1-ene and dimethylmalonate in the presence of the catalyst hexaurea **1** and the triethylamine achiral base.

scaling towards the TS from the reactants geometry using the C-C coordinate scan.

3.2.1 Conformational search

Despite the Global Optimization procedure used in Chapter 2 provided a structure which can be considered the global minimum of the system, it undoubtedly failed in the exploration of the conformational space of the molecules involved, because this goes out of the code's scope. However, when dealing with molecules with such high amounts of single bonded dihedral angles as the foldamers, which has plenty side-chain alkyl functional groups, a conformational analysis on top of the structures identified by the Global Optimization is usually necessary. Nevertheless, in this specific case, the side chain functional groups of the foldamers do not interact with the substrates. Additionally, the transition states and products of the two enantiomers were started in each case from the same reactant geometry, meaning



(a) Enantiomer R (b) Enantiomer S

FIGURE 3.3. Representation of the transition states leading to the enantiomers obtained by the reaction between (1*E*)-3-methyl-1-nitrobut-1-ene nitroolefin and dimethylmalonate in the presence of the triethylamine and the catalyst hexaurea **2**.

that the relative energies will be determined with the same rotamers of the side-chain functional groups for both enantiomers. These two points open the possibility to reduce the conformational search analysis of each foldamer treated, which is a considerably large work to do, to (i) the functional groups of the foldamers which actually interact with the substrates, and (ii) to the backbone of the foldamer.

Rotamers of the phenyl group

In the first place, the functional group which indeed interacts with the substrates and influences its positioning in the foldamer surface is the 3,5-bis(trifluoromethyl)phenyl. To study the different possible rotameric forms of this phenyl group, a scan of the pertinent dihedral angle have been carried out with and without the substrates in section B.4 on page 111 of the Appendix B. It was obtained that for the reactants, the foldamer rotamer where the phenyl group is perpendicular to its helix axis is the most stable (see Figure B.7a), and thus it has been selected in all computations hereafter carried out for the reactants. For the transition states, on the other hand, the foldamer rotamer with the phenyl group parallel to its helix axis was found to be the most stable (see Figure B.7b), and therefore has been selected in the computations of transition states in this chapter.

Conformers of the foldamer's backbone

Concerning the different conformations of the backbone of the foldamers, they have been carefully studied in Appendix B. It was found that in the case of the 1-mer (section B.2) and the 2-mer (section B.3), the most stable structures were obtained when the first and second ureas of the foldamers were oriented in opposite directions (↑↓), in such a way that the carbonyl group of the second urea was connected by means of H-bonds to the NH groups of the first urea (see Figures B.2 and B.4 for the 1-mer and the 2-mer, respectively). Of course, this is a consequence of the fact that the second urea does not establish H-bonds with any other urea when its NH groups are oriented towards the same direction of the NH groups of the first urea (↑↑). Therefore, in the ↑↓ conformation two new H-bonds are established, which will considerably decrease the energy of the systems.

Furthermore, it was also identified that in the case of the 3-mer, the ↑↓ conformer will no longer be the most stable, as in that case the foldamer would be four ureas long, enough to establish H-bonds between ureas 1 and 3, and 2 and 4, leaving no urea unconnected in the helix.

Additionally, at the end of section 2.5 on page 52 in Chapter 2, it was proposed that probably the low reactivity found in the 1-mer and 2-mer foldamers in the experiment could be related to the difficult approximation pathway of the reactants in the ↑↓ conformation (see Figures B.3 and B.5 for the 1-mer and 2-mer, respectively). Then, finding the transition states would be extremely difficult for the 1-mer and 2-mer in the conformation ↑↓, and it will certainly be different than the transition state found for the 3-mer, 4-mer, 5-mer, etc., which posses an ↑↑ conformation.

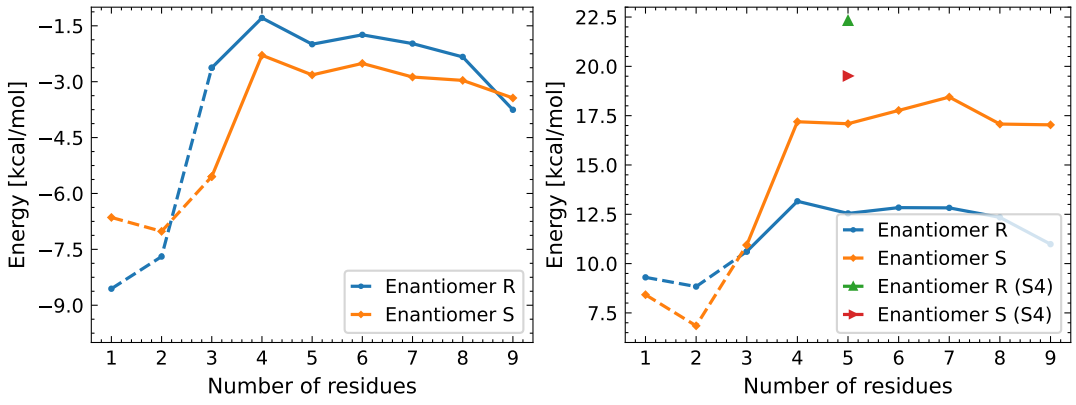
Also, the conformation ↑↑ was around 8 $kcal/mol$ (1-mer) and 12 $kcal/mol$ (2-mer) higher in energy than the ↑↓ one. However, with the addition of the substrates these energy discrepancies considerably decreased, becoming 3.29 $kcal/mol$ and 0.27 $kcal/mol$ for the 1-mer and the 2-mer, respectively. Based on these low energy differences, it was then decided to include in the analysis of the transition states the 1-mer and 2-mer of the conformation ↑↑.

3.3 Energetic profile of the catalyzed reaction

The reaction here studied (see Figure 2.1 on page 39) are exothermic, as can be clearly seen in the enthalpy of reaction in Figure 3.4a. In general, the enantiomer S has lower energy (~ 1 $kcal/mol$) than the enantiomer R throughout the series of foldamers, with a slight increase of the gap in the 3-mer which is around 3 $kcal/mol$

lower.

Concerning the activation energy (E_a) —and therefore the enantiocontrol—, it was obtained that foldamers larger than the tetraurea will present high enantiocontrol, whereas from 1-mer to 3-mer the enantiocontrol should be considerably reduced (see Figure 3.4b). These results are deduced from the difference in E_a between enantiomers R and S: the enantiomer R has lower E_a barrier than the S, therefore being favored. Furthermore, observe that in the case of the 1-mer and 2-mer the activation energy is lower for the enantiomer S, which represents a discrepancy with the experiment [29] because the enantiomer R has always been measured to be predominant in the mixture. However, it has to be remembered that the 1-mer and 2-mer foldamers here reported are those with the ↑↑ conformation (see discussion at the end of section 3.2.1 on the previous page). To avoid misunderstandings, their estimation has been represented with dashed lines hereafter.



(a) Reaction enthalpy (b) Activation energy

Figure 3.4. Reaction enthalpy and activation energy of the studied systems as a function of the foldamer's number of residues. The `S4` energy values correspond to the top-right configuration in Figure 2.5.

Moreover, a single point computation for each B3LYP-D3 activation energy reported in Figure 2.5 was carried out with the M06-2X functional, keeping unchanged all other variables. Then, an activation energy confidence region was built with both results (see Figure 3.5). The M06-2X functional provided lower values of activation energy, but with exactly the same trend. This filled picture further highlights the exceptional catalytic properties of these foldamers when formed up by 5 or more ureas, and the low enantiocontrol obtained with shorter backbones.

With the values of activation energy barriers, the next logical step is to estimate form a theoretical point of view the enantiomeric excess (ee) values and compare them with those experimentally obtained. To that end, one can use equations (3.1)

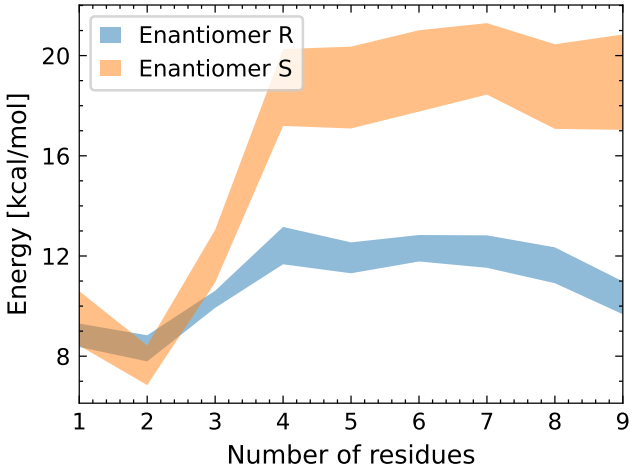


FIGURE 3.5. Region between the values of activation energy obtained using B3LYP-D3/6-31"+"G(d,p) and a single point from those optimized structures with M06-2X/6-31"+"G(d,p).

and (3.2), as described by Wåhlander et al. [109]

$$\%ee = \frac{\left(\frac{k_R}{k_S} - 1\right)}{\left(\frac{k_R}{k_S} + 1\right)} \tag{3.1}$$

$$k = \frac{k_b T}{h} \exp\left(-\frac{\Delta E_a}{RT}\right) \tag{3.2}$$

being k_b the Boltzman constant, R the molar gas constant,[1] h the Plank constant, and T the temperature of the system under investigation.

However, when E_a differences larger than 1 or 2 $kcal/mol$ are observed between the enantiomers, the ee values will always be $> 99\%$ following equation (3.1). These large differences in E_a are due to the fact that the accuracy of relative energies continues to be a challenge for computational chemistry [233]. The application of DFT to molecular systems of practical interest usually raises questions on the appropriateness of the functional and the size of the basis set, with other topics, such as solvation models or free energy corrections often being considered. Despite in this work all the possible error sources mentioned in the previous sentence have been carefully treated, activation energies close to 5 kcal/mol were obtained for foldamers larger than the pentaurea, which will result in $ee = 100\%$. Note that, although the results here presented are not perfect, they do provide a pertinent description of the enantiocontrol resulting from employing these foldamers as catalysts in the studied

[1]Note that the R in the parameter k_R refers to the enantiomer R and not to the molar gas constant.

asymmetric reactions.

Additionally, the transition states of the configuration `S4` depicted in Figure 3.6, which also presented a suitable reaction pathway, were investigated with the hexaurea as catalyst (see points labeled as `S4` in Figure 3.4b). It was obtained that not only the activation energy barrier is higher for both enantiomers in this case, but also the enantiomer S is favored in this situation, which is contrary to the experimental experience. These results support the structures obtained by applying the Global Optimization procedure in section 2.3 on page 43, specially those reported in Figure 2.5 on page 45. Therefore, it is possible to conclude that the global minimum found for the foldamer-substrates-achiral base catalytic system will probably be the geometry adopted in the real chemical system.

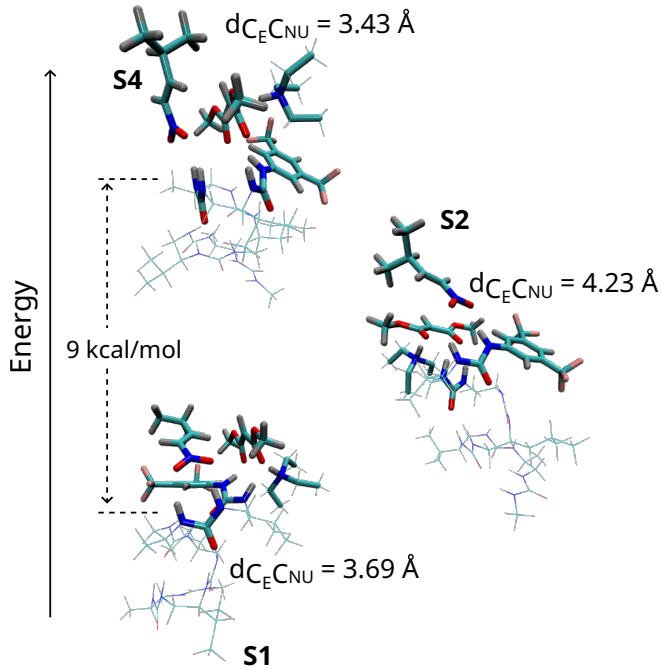
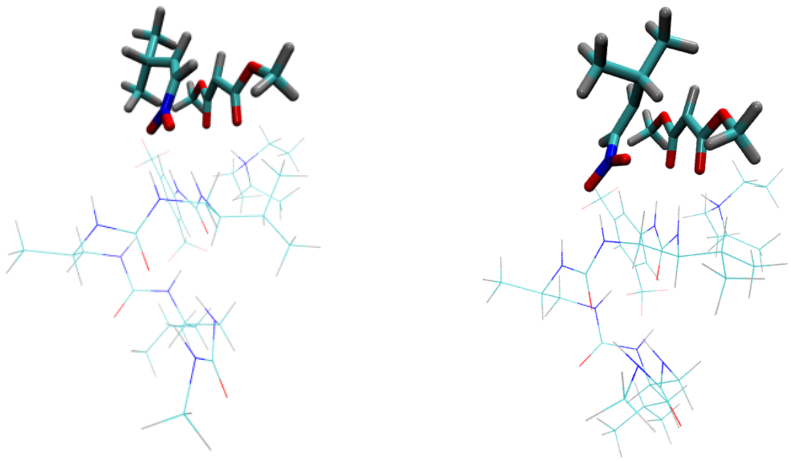


Figure 3.6. Summary of the possible structures identified for the system under investigation (see complete information in Figure 2.5 on page 45 in Chapter 2).

Finally, it has to be pointed out that accordingly to the values of activation energy presented in Figure 3.5, large foldamers as the 8-mer and the 9-mer should also perform correctly in the catalysis of the studied reactions. However, unpublished results obtained experimentally by the group of Gilles Guichard at the University of Bordeaux show that the 9-mer presents enantiocontrol and reactivity as low as the 2-mer and 3-mer. This behavior could be related to the appearance of new sites in the electrostatic potential surface others than sites `[1]` and `[2]`, which would attract one of the substrates, thus lowering the catalytic performance of the foldamer. Observe how in Figure 2.9(d) the ESP positive region encloses a larger portion of the foldamer outside sites `[1]` and `[2]`.

3.4 Cis conformation of the nitroalkene

We next investigated the Cis conformation of the nitroalkene studied so far, despite this conformer was never used in the experiments. Although the analysis of this conformation is not the general objective of this manuscript, it could be interesting to inspect how a change in the conformation of the nitroalkene could influence the outcome of the catalyzed reaction. Finally note that the Cis conformation of the nitroalkene will be treated exclusively within the limits of this section.



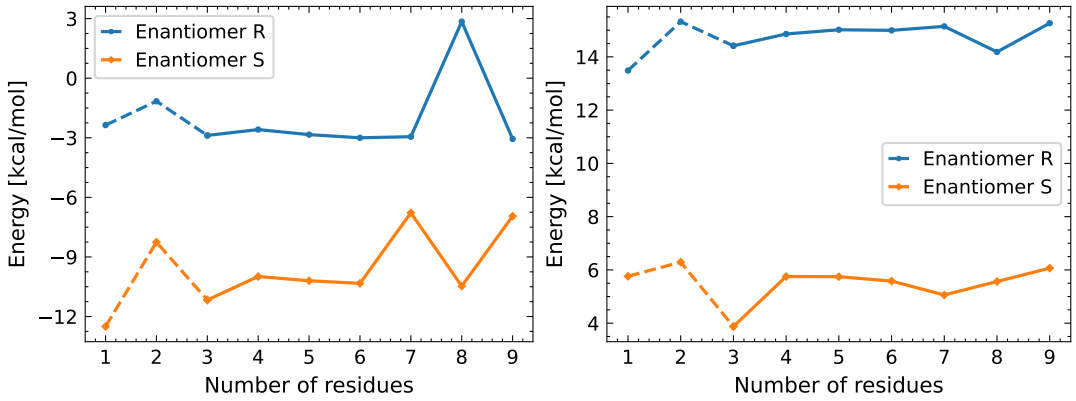
(a) (Cis) Enantiomer R (b) (Cis) Enantiomer S

Figure 3.7. Representation of the enantiomers obtained by the Cis conformation of the nitroalkene nitroolefin when reacting with the dimethylmalonate in the presence of the triethylamine and the catalyst hexaurea **2**. Only the substrates are highlighted, but just to gain in clarity.

Figure 3.7 shows the geometry of the cis conformation of the nitroalkene, and the disposition of both reactants in the catalyst surface in the transition states geometries leading to enantiomers R and S.

It was found that the enantiomer S is considerably favored over enantiomer R in both kinetics and thermodynamics. The stability of the products is around 7 $kcal/mol$ higher in the enantiomer S than in the R, as can be observed in Figure 3.8a. Besides, the activation energy difference between enantiomers R and S is ~ 8 $kcal/mol$, favoring the formation of S (see Figure 3.8b). Finally, it is interesting to note that when the conformer Cis of the nitroalkene is employed in the reaction, the enantiocontrol provided by the foldamers could be elevated, and the enantiomer S will be probably favored no matter the chain length of the foldamer.

These results are quite interesting, because if the enantiomer S of this compound were the target in some application, and the Cis configuration of the nitroalkene could be purchased and used, probably the employment of the catalytic binary



(a) Reaction enthalpy (b) Activation energy

FIGURE 3.8. Reaction enthalpy and activation energy of the Cis conformation of the nitroalkene when reacting with the dimethylmalonate in the presence of the triethylamine and the catalysts. In this case the Cis conformation of the (1*E*)-3-methyl-1-nitrobut-1-enewas taken into account.

system of the foldamers and the triethylamine as achiral base would lead to an excellent enantiocontrol.

3.5 Analysis of the system geometry

Despite the gap between the activation energies obtained theoretically being in agreement with the enantiocontrol obtained experimentally (see section 3.3 on page 61), the identification of the factors making these foldamers exceptionally enantioselective is missing from the analysis at this point. The purpose of the present section is to identify and carefully analyze key geometrical parameters of the foldamers and the substrates. Parameters describing the (i) interaction between the reactants, (ii) the interaction of the substrates with the catalyst, and (iii) the disposition of the residues of the foldamers in the helix will be measured.

3.5.1 Distance between the two reactive carbons

When the distance between the two reactive carbons is measured as a function of the number of residues in the transition states geometries, in the enantiomer R it is observed a decrease in value which is coincident with the increase in enantiocontrol of the foldamers (see Figure 3.9). Thus, a clear correlation is present between the activation energy and the C-C distance in the enantiomer R. This happens because the closer the transition state C-C distance to that in the product, the easier the reaction will proceed. For the enantiomer S, on the other hand, the C-C distance

increases for the 4-mer and larger foldamers, suggesting that the reaction pathway could be more difficult than in the case of the enantiomer R.

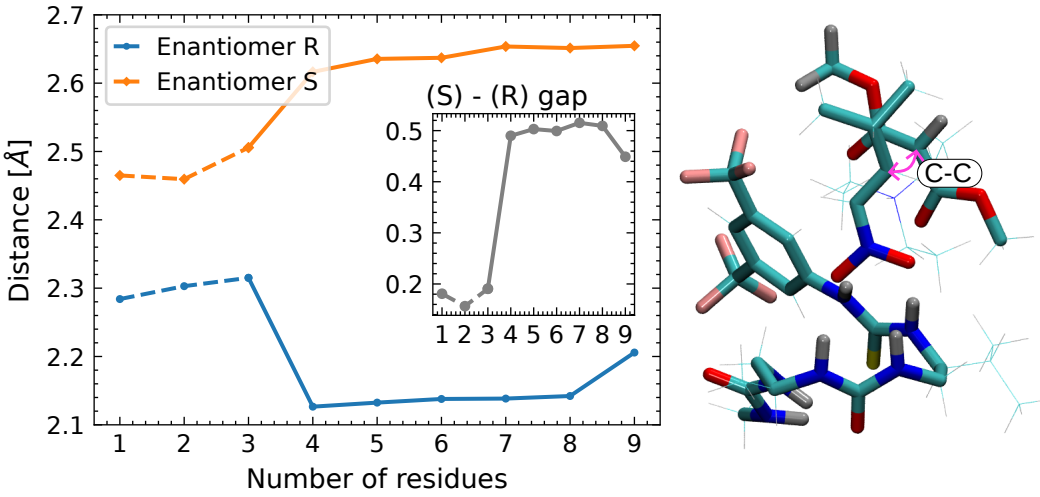


Figure 3.9. Distance evolution as a function of the number of residues of the two reacting carbons in the TS for the enantiomers R and S. In the right, a schematic representation of the reported value.

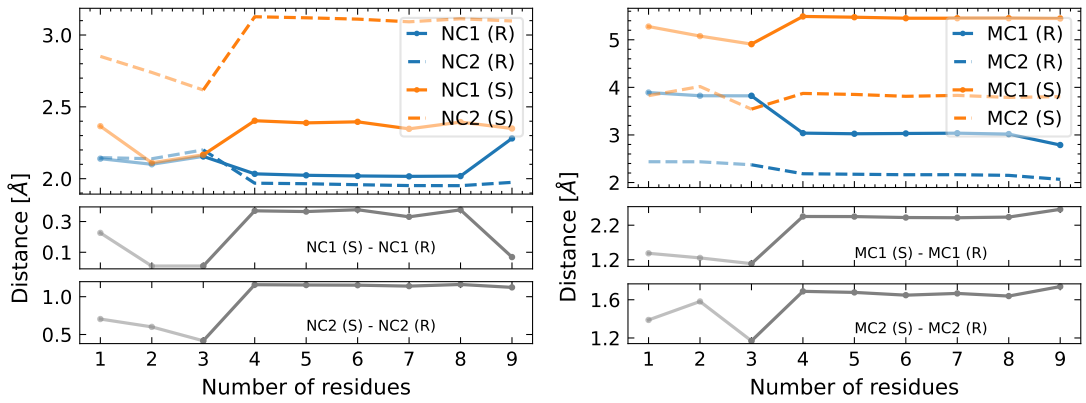
Furthermore, if the difference between the C-C distance of the enantiomer S and the enantiomer R is taken (reported in gray in Figure 3.9), a profile coincident with the enantiomeric excess (*ee*) reported by Bécart et al. [29] and the enantioselectivity theoretically reported in section 3.3 on page 61 is obtained. Although these results do not explain the basic mechanisms leading to the enantiocontrol of the foldamer, they clearly show that the enantiomer R transition state will be favored by 0.5Å in the C-C distance over the enantiomer S, which will probably influence the foldamer's ability to produce one or the other enantiomer.

3.5.2 Distances between the substrates and the foldamer

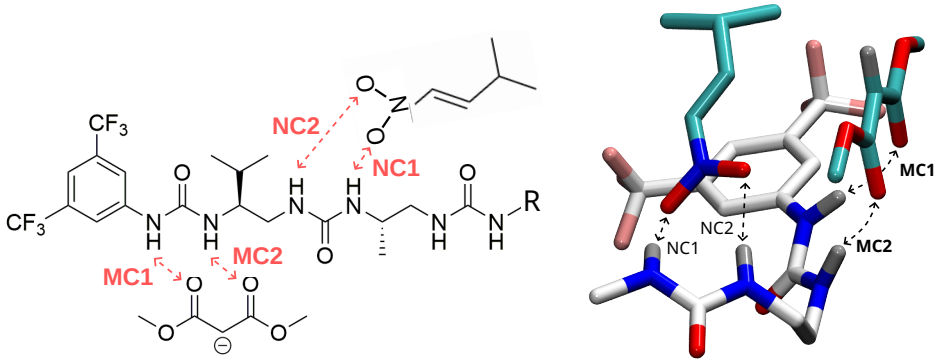
We now investigate the relation of the interactions substrates/foldamers with the enantiocontrol. To that end, the stabilizing H-bonds established between the catalyst and the substrates were characterized by measuring their corresponding distances as a function of the foldamer's length (see Figure 3.10). Note that the name `MC` makes reference to the interactions between the malonate (M), and the catalyst (C), the same being applied for the interactions (NC) between the nitroalkene (N) and the catalyst.

It was obtained that whether in the interactions involving the nitroalkene or those of the malonate, the transition state of the enantiomer R present smaller H-bonds distances, therefore stronger stabilization.

The stabilization resulting from shorter MC and/or NC distances increases for foldamers larger than the trimer, and it is reflected by the gray curves in the bottom of each figure. These gray curves are obtained from the subtraction of the distances of enantiomer S with its corresponding distance in the enantiomer R. The greater the value of the gray curves, the more favored the enantiomer R should be. Thus, one more time, it is obtained that the foldamers from 1-mer to 3-mer are disfavored in transition states stability if compared to the larger foldamers, which seems to be correlated with the enantiocontrol.



(a) NC distances (b) MC distances



(c) Distances label (d) Distances label

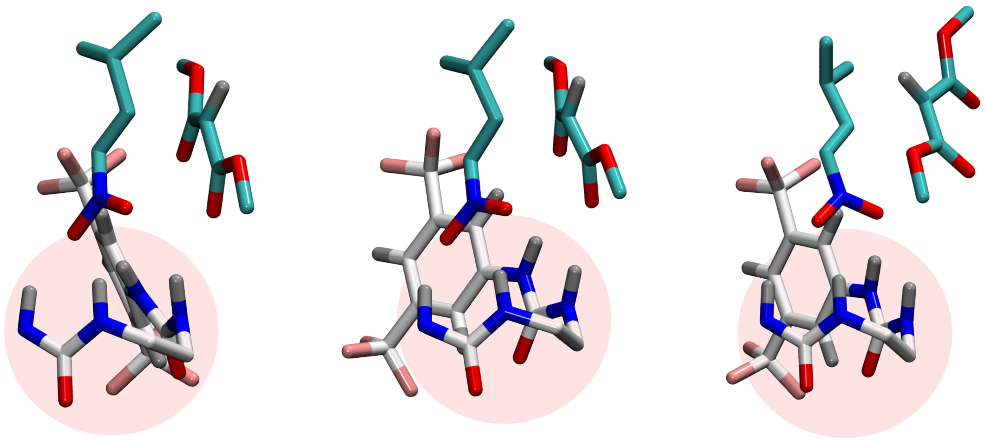
FIGURE 3.10. Measurement of interesting distances in the foldamers as a function of the number of residues, in the transition state geometry.

3.5.3 Geometrical analysis of the foldamers

Finally, we investigate how could the interactions between the substrates and the catalyst be determined by the geometry of the foldamers alone. In Figures 3.11a, 3.11b, and 3.11c it is represented the transition states geometry of the tetraurea resulting in the enantiomer R, the heptaurea resulting in the enantiomer R, and the heptaurea resulting in the enantiomer S, respectively.

First, let's compare the 3-mer (R) with the 6-mer (R). In the 3-mer (R), the NH groups of the first two ureas are closer, and it seems the consequence of a slight counterclockwise rotation of the first urea. As result, the foldamer has less space on site `[1]` to accommodate the malonate without generating hindrance with the nitroalkene. Additionally, as a result of the previous point, the malonate is farther from the foldamer, which could decrease the interaction strength between them, and thus the stabilization of the system. For those reasons, the transition state leading to the enantiomer R will be probably favored to a greater extent over the enantiomer S in the 6-mer than in the 3-mer.

Now let's compare the 6-mer (R) with the 6-mer (S). The NH groups of the two ureas are closer one to the other in the case of the 6-mer (S), despite it is not as clear as in the case of the 3-mer (R). However, the remarkable difference is the position of the substrates on the catalyst surface. In the 6-mer (S), the nitroalkene's acceptors has moved towards the first urea, and the malonate has completely dissociated from the catalyst (see distance `MC1 (S)` in Figure 3.10b). This behavior will inevitably result in a less stable transition state structure for the case of the 6-mer (S), and thus the increase in activation energy.



(a) 3-mer (R) (b) 6-mer (R) (c) 6-mer (S)

Figure 3.11. Geometrical analysis of different catalysts and substrates (TS). The foldamers have been represented with only two ureas for the sake of clarity in the figures.

Overall, the most favored geometry of the transition states seems to be that of the 6-mer (R), which is reflected in the energetic profiles presented in Figure 3.4 by being superior in enantiocontrol over the two others. This aspect can be correlated with the strength of the H-bonds established between the foldamer and the dimethylmalonate: the greater the ability of the catalysts to strongly bind the dimethylmalonate on site `[1]`, the better the enantiocontrol should be. Note that the

strength of the H-bonds in site [1] can be measured experimentally and estimated theoretically by means of the pK_a, which will be reviewed in detail in Chapter 4.

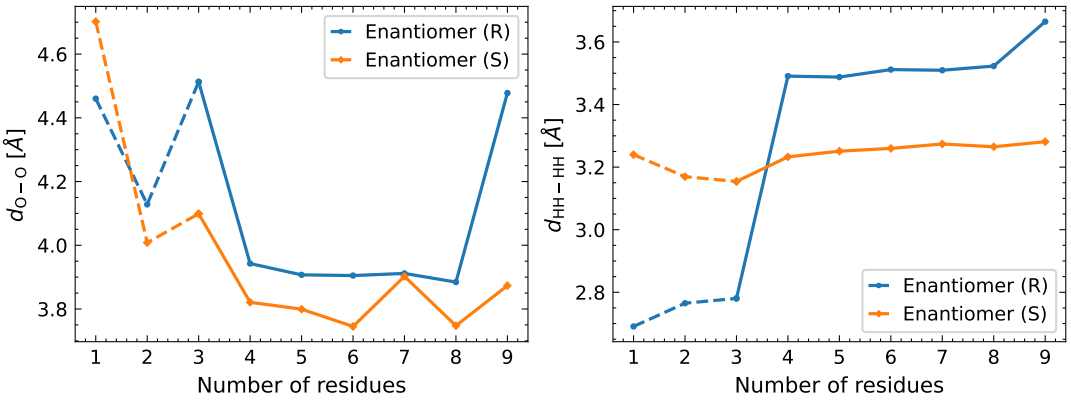
Distance between the first two ureas

Nevertheless, it is possible to quantitatively measure the qualitative observations made around Figure 3.11. For example, (i) the distances between the midpoint of the NH groups in the first two ureas ($d_{\mathrm{HH-HH}}$) will describe how much closer is one to the other, i.e. how much is the first urea rotated counter- or clockwise, and (ii) the distance between the two oxygen atoms of the carbonyl groups in the first and second ureas ($d_{\mathrm{O-O}}$), which should behave inversely to $d_{\mathrm{HH-HH}}$.

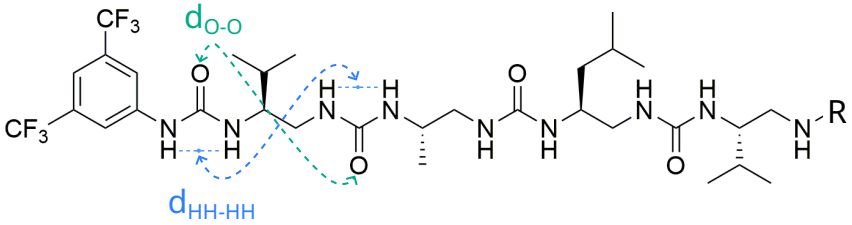
As previously discussed, it is probable that the closer the NH groups of the first two ureas are, the less favored the transition states should be due to the reduction of the space to accommodate the substrates on the foldamer (small $d_{\mathrm{HH-HH}}$), and the consequent loss of interactions between the malonate and the foldamer. Figure 3.12b reflects that from 1-mer to 3-mer the enantiomer R will posses a smaller — and thus unfavorable— distance between the ureas, whereas from 4-mer to 9-mer the game completely changes and the ureas split apart, facilitating the accommodation of both substrates. Therefore, the transition states structures should be more stable in the later range, and higher the enantiocontrol.

The urea-urea distance in the enantiomer S, on the other hand, remains nearly constant. This, in combination with a careful study of Figure 3.11c, suggests that in the case of the transition states geometry of the enantiomer S the distance between ureas does not influences the enantiocontrol of the foldamer. Of course, as the malonate has completely dissociated from the foldamer, no matter how much space can sites [1] and [2] grant, it will not be able to use the NH donors to establish stabilizing H-bonds.

Finally, note how the distance $d_{\mathrm{O-O}}$ in the transition states geometry of both enantiomers has a tendency to decrease in value as the number of residues of the foldamers increases, meaning that more space is created in the reactive region by larger catalysts (see Figure 3.12a). However, it is interesting to highlight how both distances $d_{\mathrm{HH-HH}}$ and $d_{\mathrm{O-O}}$ in the enantiomer S are smaller than in enantiomer R, suggesting that both, the Oxygen atoms of the carbonyl groups of the first two ureas of the enantiomer S, and also the NH groups are closer one to each other. This means that not only the dihedral angle defining the counter- or clockwise rotation of the first urea is involved in the process, but also the angle defining the distance between them.



(a) d_{O-O} (b) d_{HH-HH}



(c) Schematic representation of the distances

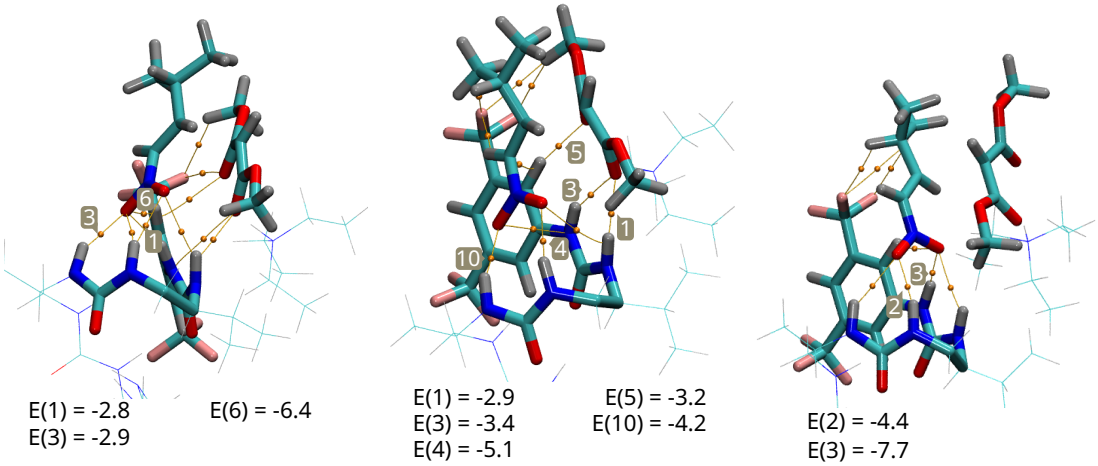
FIGURE 3.12. Distances between the mid points of the NH groups of the first and second ureas, and their Oxygen atoms, as a function of the foldamer's number of residues, in the transition states geometries.

3.6 Analysis of the system NCI

In addition to the geometrical analysis developed in the previous section, an investigation of the stabilizing non-covalent interactions between the substrates and the foldamers could lead to interesting and complementary findings. This is the objective of the next few paragraphs.

A QTAIM study has been carried out in the structures reported in Figure 3.11. In agreement with the geometry and distances analysis made with them, it is the 6-mer (R) the transition state which posses more stabilizing interactions between the substrates and the foldamer (see Figure 3.13). The other two structures are considerably disfavored with only 2 and 3 H-bonds, all of which are between the nitroalkene and the catalyst.

If the H-bonding interaction energies reported in Figure 3.13 were summed up, it could be a measure of stabilization between the substrates and the foldamers. However, it would not be a rigorous value as many other interactions would be missing. On the other hand, the IGMPlot code [197, 234–236] proposes the Grid Integration Score (GIS) quantity as a measure of stabilization energy between two fragments of a system. The greater the GIS value, the stronger the stabilization



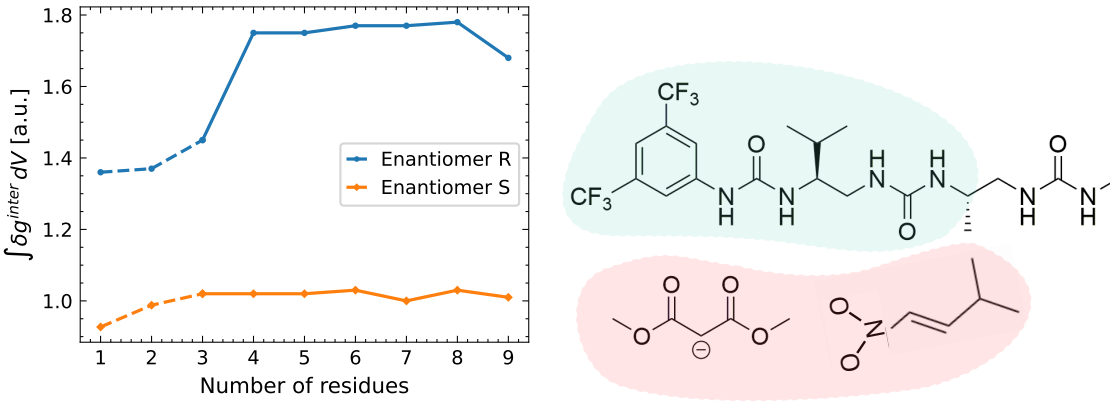
(a) 3-mer (R) (b) 6-mer (R) (c) 6-mer (S)

FIGURE 3.13. QTAIM analysis and related H-bond binding energies E deduced from the electron density at the Bond Critical Point (BCP) following the method proposed by Emamian et al. [219] and discussed in section 1.14 on page 32. Values of E are in $kcal/mol$.

resulting from non-covalent interactions.

The descriptor δg represents locally the difference between a virtual upper limit of the electron density gradient (∇_ρ^{IGM}) representing a non-interacting system, and the true electron density gradient of the system ∇_ρ (see [197] for more theoretical details on the Integrated Gradient Model electron density reference model). With δg, the GIS can be defined by the following equation:

$$\text{GIS} = \int \delta g^{inter}\, dV \tag{3.3}$$



(a) Grid Integration Score (b) NCI defined regions

FIGURE 3.14. Grid integration score as a measure of stabilizing energy coming from non-covalent interactions between two fragments of a system. It has been computed with the IGMPlot code.

Figure 3.14a shows the GIS values as a function of the number of residues in the foldamer for enantiomers R and S, with the defined interacting fragments drawn in Figure 3.14b (substrates interacting with the reactive region of the catalyst). As depicted for the enantiomer R, the systems in the 4-mer to 9-mer range present high values of GIS, meaning they have a strong stabilization coming from non-covalent interactions. The shorter catalysts will produce less stable configurations. This trend has been repetitive throughout the analysis of the activation energies, the geometrical parameters, and now the non-covalent stabilizing interactions studied in this chapter, being highly correlated with the experimental enantiocontrol. For the enantiomer S, on the other hand, the stabilizing interactions posses lower —and almost constant— value than those of the enantiomer R.

3.7 Conclusions

The chapter here concluding revealed for the first time a theoretical estimation of the enantiocontrol provided by **1**-like catalysts in the reaction between (1*E*)-3-methyl-1-nitrobut-1-ene and deprotonated dimethylmalonate. The pentaurea was identified to be the turning point: shorter foldamers will present low enantiocontrol, and larger foldamers will provide high enantiocontrol. These results are in agreement with the experimental measurements, and support the structures found by Global Optimization procedures in Chapter 2.

Several methodologies were tested to carry out the computations, and it was observed that the combination of B3LYP-D3/6-31G“+”(d,p) with PM6 is viable to considerably save time and computational resources, while obtaining deviation of around 2 $kcal/mol$ with respect to computations carried out entirely with B3LYP-D3/6-31G“+”(d,p).

The results discussed around the energetic profile were coincident with the experiment. Besides, several analysis of key geometrical parameters were carried out, and a correlation between them and the enantiocontrol was identified. The distance between the two reactive carbons is favored (shorter) for the enantiomer R, and disfavored (larger) for the enantiomer S. This supports the possibility that enantiomer R will be predominant in the mixture over enantiomer S, as the reaction path should be easier in the former.

Furthermore, when looking at the interactions of the substrates with the foldamers in the transition states geometries, it was identified that the catalysts which have the first and second ureas more separate one from the other also posses higher enantiocontrol. This behavior may be linked to a greater space created be-

tween sites [1] and [2], which would decrease the steric hindrance between the substrates, and therefore increase their stabilizing interactions with the foldamer. Note how the space created in sites [1] and [2] is correlated to the strength of the interactions between site [1] and the dimethylmalonate, which is ultimately the substrate affected by a reduced space in the reactive region.

Moreover, it can be also concluded that the cooperation between ureas internally in the helix seems to emerge as a quite important parameter to look at when studying enantiocontrol in these foldamers. As observed in the case of the first urea for short foldamers, it tends to rotate counterclockwise, reducing the space in the reactive site to accommodate the substrates. Then, the stronger the attachment of each urea, but mainly the first two, to the internal ureas cooperation in the helix, the more difficult it should be for it to rotate clock- or counterclockwise. For example, in the tetraurea only urea #4 is connected to urea #1, and only urea #3 is connected to urea #2; extending the analysis to the pentaurea, two residues are connected to urea #1, and only one residue is connected to urea #2. This suggests that maybe a certain number of residues are needed to strongly enough attach the first and second ureas in a suitable position to receive the substrates, leading one more time the analysis to a relation between enantiocontrol and the strength of the H-bonds that the catalyst would be capable to establish, mainly with the dimethylmalonate.

As direct consequence of the geometrical parameters, which are considerably favored for the enantiomer R, the stabilizing non-covalent interactions established between the substrates and the foldamers are higher for enantiomer R in the pentaurea to decaurea foldamers.

These results in combination with the Molecular Recognition between substrates and catalyst provide deep insights about the basic processes ruling these catalytic systems. Many of the ideas developed in the experimental work have been supported by our computations, and at the same time many of the observation discussed in our computations find foundation on the experimental results.

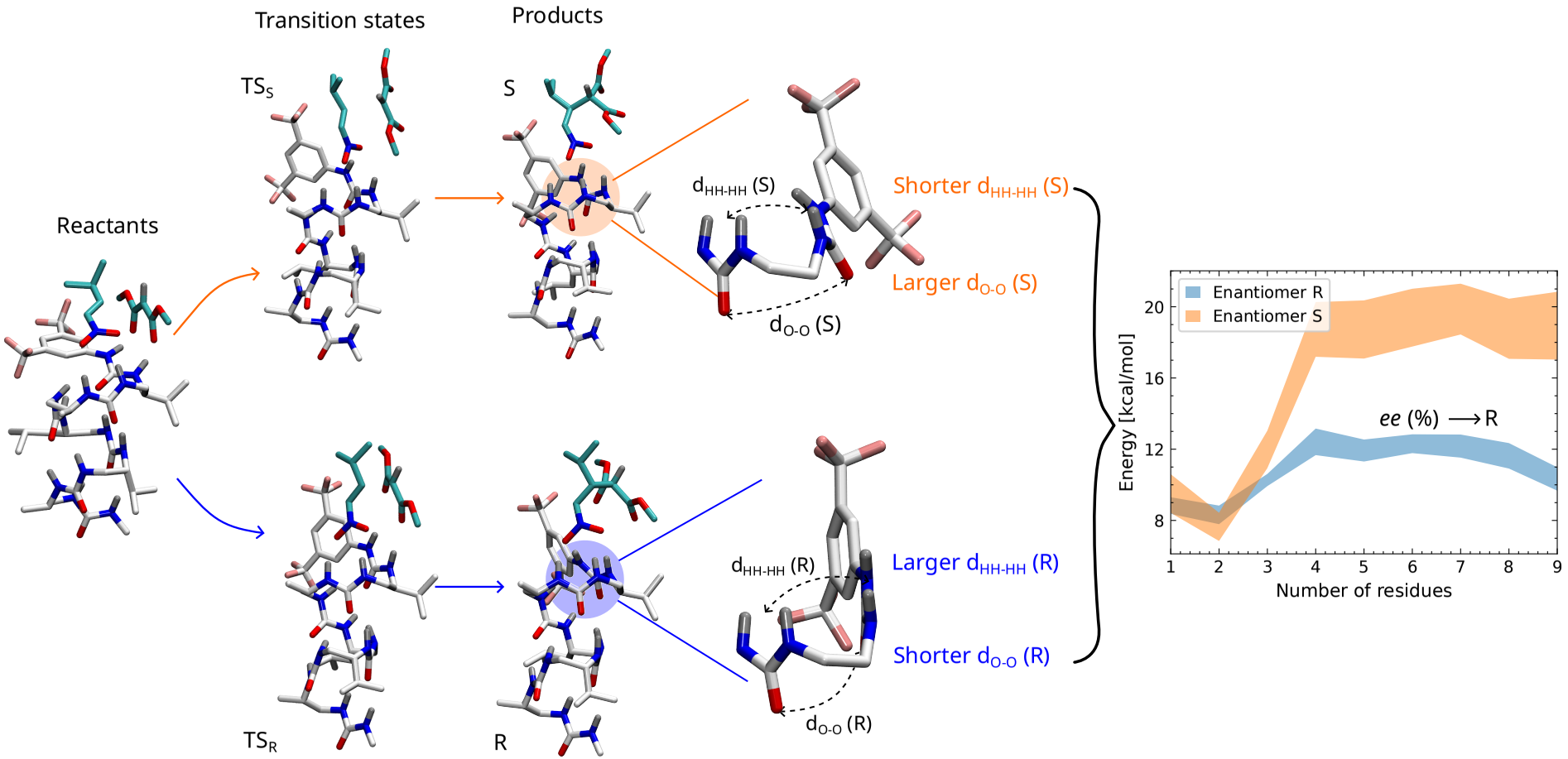


Figure 3.15. Enantiocontrol should be related to key geometrical parameters, which in turn define the H-bond strength between the dimethylmalonate and the first site of the catalysts. The triethylamine base has not being depicted for the sake of clarity.

Chapter 4

Introducing changes to the foldamer

The insights gained into the catalytic properties of aliphatic N,N'-linked oligoureas like **1** in Chapters 2 and 3 allow a better understanding at the atomistic level of their capabilities and a rationalization of their performance. It is then possible, by modifying the nature of the atoms and the functional groups at punctual locations of the foldamers, to tune the key parameters influencing the enantiocontrol (see Figure 4.1). The objective of the present chapter is to move a step further, employing the tools of computational chemistry in the field to go from rationalization to molecular design.

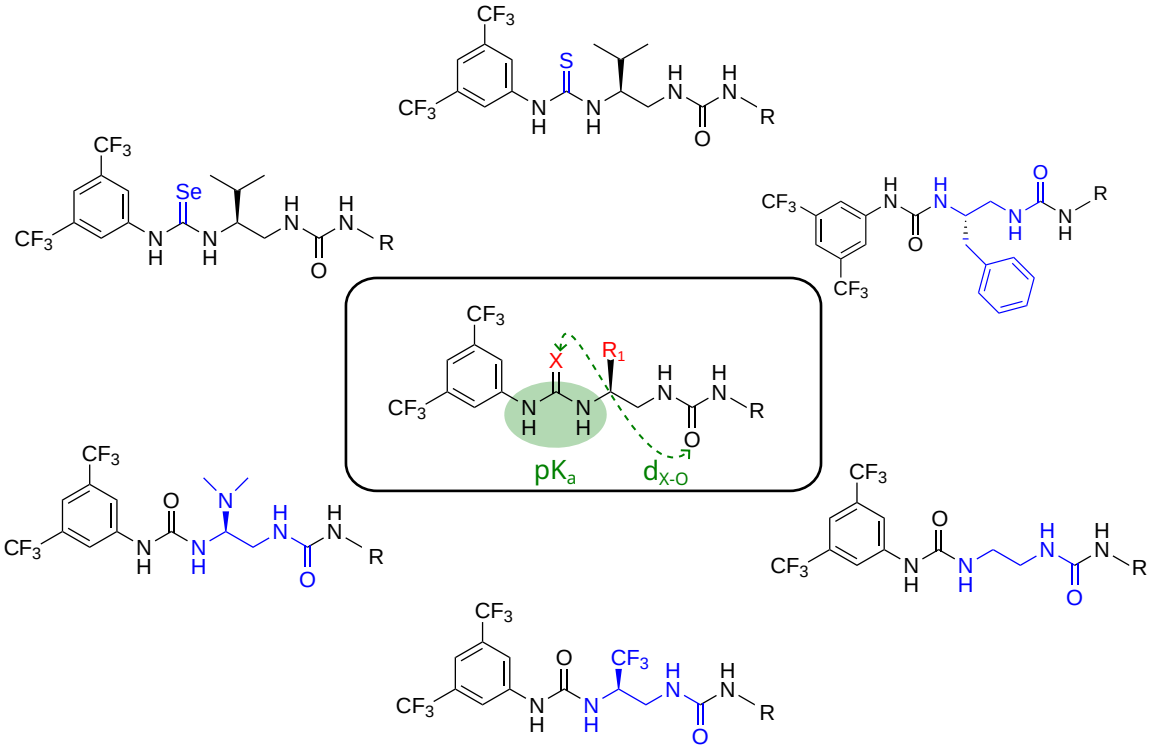


Figure 4.1. Modifications made to the foldamer **1** with the objective to vary the acidity of site `[1]`, and consequently investigate the influence of its catalytic properties.

4.1 Introduction

The investigation of the catalytic properties as a function of small structural modifications is a common practice in the development of new and improved catalysts [237, 238]. More specifically, in the systems investigated by Bécart et al. [29] a rationalization study was carried out by modifying the sites `[1]` and `[2]`, and the related side-chain functional group in the foldamers, obtaining many interesting hypotheses to work with in their studies. Just to mention some, the second urea of the foldamer was blocked, and as consequence the enantiocontrol and yield considerably decreased, thus the site `[2]` was identified to posses a great importance in the catalytic performance of these oligomers. Precisely the role and importance of the second urea in the system has been extensively discussed in this manuscript, throughout a comprehensive description of the foldamer-substrates interactions (see Chapter 2).

Another interesting modification was the substitution of the Oxygen atom in the first urea of the foldamer by a Sulfur atom (see foldamers **3** and **3a** in Figure 4.2). This replacement of a urea by a thiourea was supposed to improve the enantiocontrol and reactivity of the foldamers, due to the acidity increase in the reactive region and the consequent increment of the donor character of site `[1]`. However, despite foldamer **3a** was observed to perform well, it was overall less effective than foldamer **1**.

Furthermore, the use of isothioureas, isoureas, and isoselenoureas has been widely explored by Smith and coworkers in asymmetric catalysis (see the review Bitai et al. [239], for example). In the enantioselective conjugate addition of simple nitroalkenes to α,β-unsaturated aryl ester Michael acceptors, they have observed that the isoureas gave a racemic conjugate addition product. Isothioureas and isoselenoureas, on the other hand, performed well with *ee* above 95% and yields $\sim$ 55%. Nevertheless, the isoselenoureas presented a drastic increase in reaction rate ($t_{1/2} = 8\,min$) relative to the same reaction catalysed by an isothiourea ($t_{1/2} = 145\,min$). Based on these results, the substitution of the Oxygen atom in foldamer **2** by a Selenium atom will also be taken into account in our investigation, despite the catalysts analyzed by Smith and coworkers are quite different than those here studied (see foldamer **4** in Figure 4.2).

These modifications are expected to influence the acidity of the first site of the foldamers, and thus its donor character towards the binding of dimethylmalonate. Nevertheless, recall from Chapter 3 that it was the malonate which largely suffered from a reduced space in the reactive region of the foldamers, therefore, the rear-

rangement of the helix to accommodate bigger atoms than Oxygen could lead not only to variations of the site's pK_a, but also the distances between the first and second ureas ($d_{\mathrm{O-O}}$ and $d_{\mathrm{HH-HH}}$).

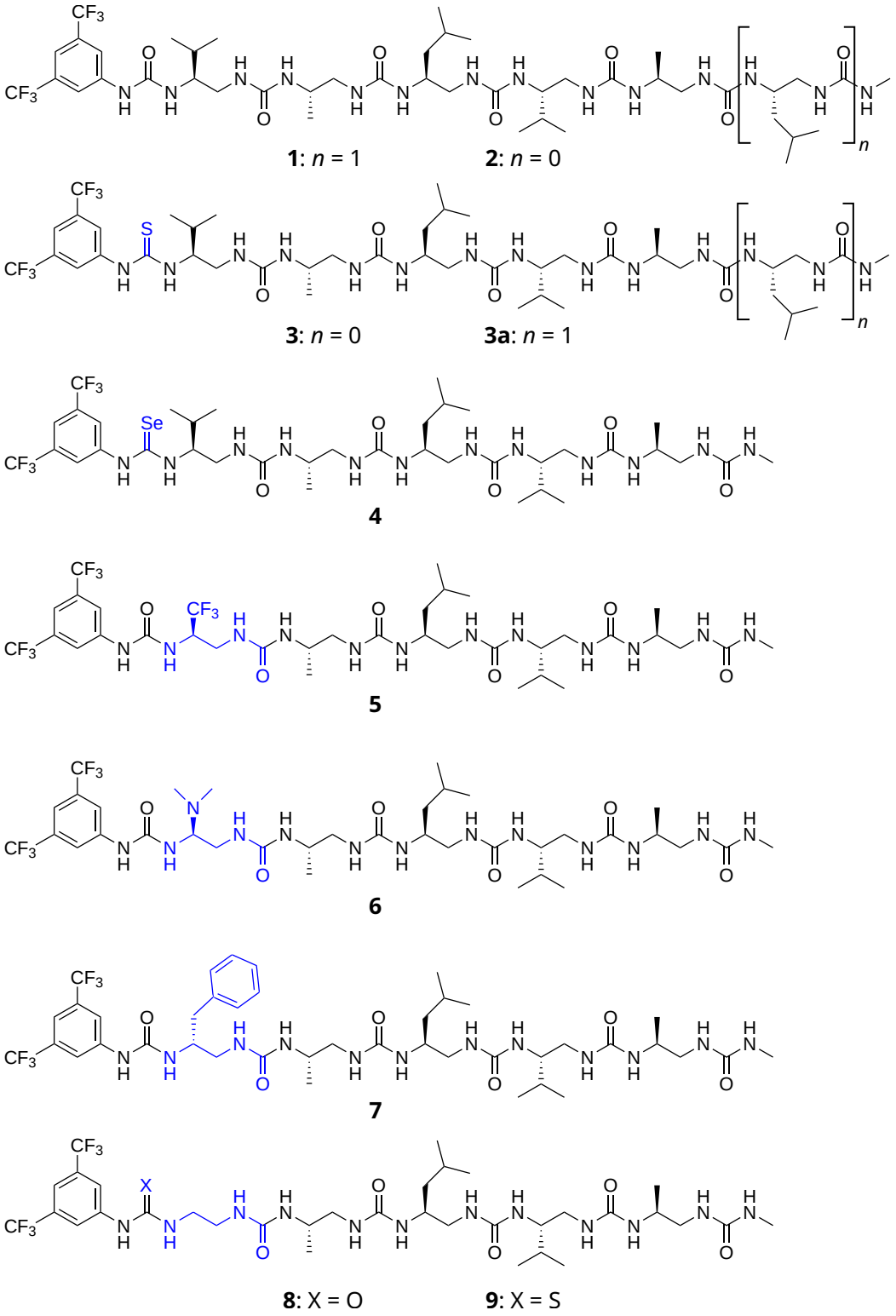


Figure 4.2. Modifications made to foldamer **1** in order to investigate their influence on the enantiocontrol and reactivity.

With that in mind, it was also design the study on how changes not just in the first urea group, but also on the first side-chain functional group, could influence the catalytic properties of the foldamers. The first side-chain group was substituted by others with larger capacity to attract electrons to it: (i) a CF_3 group, (ii) a $N(CH_3)_2$ group, and a benzyl group. In all cases, it is expected that the substitution is going to induce a displacement of electrons from the first urea to the substituent group, and thus modify the acidic character of the first urea. An increase in the H-bond donor character of the first urea could strongly bind the deprotonated dimethylmalonate, which would facilitate the formation of enantiomer R. Note that enantiomer S should not benefit from this changes because its transition state geometry presents poor interaction between the first urea and the deprotonated dimethylmalonate.

The following sections present the results obtained for each modification proposed in the above paragraphs and some others represented in Figure 4.2.

4.2 Replacing Oxygen by Sulfur and Selenium

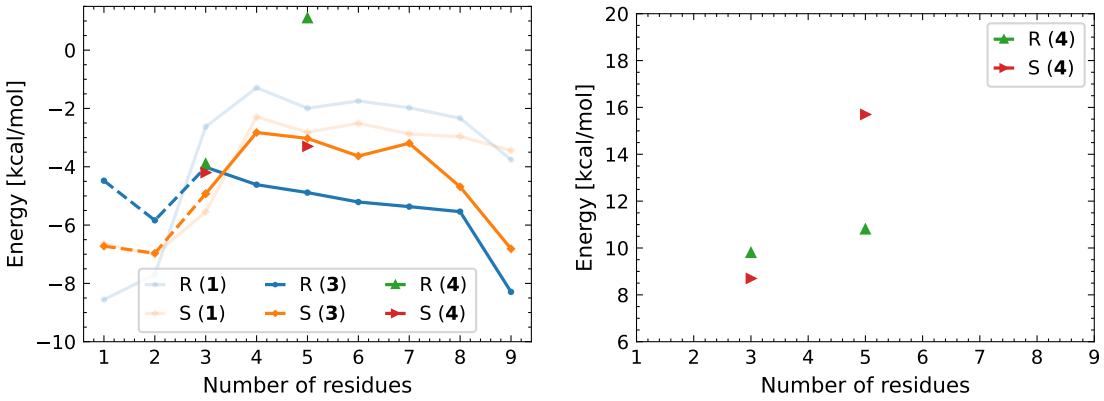
As Sulfur and Selenium are bigger atoms than Oxygen, some geometrical modifications are expected in sites `[1]` and `[2]` of foldamers like **3** and **4** when compared to **1**-like.[1] In fact, a slight decrease in catalytic performance was observed in the experiment [29] when they substituted **1** by **3**, and it was attributed precisely to a possible structural change.

Energetic profile

It was observed in our computations that with the introduction of the Sulfur atom, from a thermodynamic point of view, both enantiomer R and S are exothermic (see Figure 4.3a). Also, note that the reaction enthalpy is around 2 $kcal/mol$ lower for **3**-like foldamers (thiourea in site `[1]`) than for **1**-like (urea in site `[1]`).

The activation energy barrier for foldamers like **3**, on the other hand, posses a profile more irregular than foldamers like **1**, the latter characterized by a larger and continuous gap between activation energies of enantiomers R and S which is not present in the former (see Figure 4.3b). Nevertheless, the enantiomer R appears to be favored over the enantiomer S also for foldamers like **3**, and thus the enantiocontrol obtained from their use should be satisfactory, but lower than in foldamers like **1**

[1] Hereafter, the phrases "foldamers like **n**" or "**n**-like foldamers" refer to variations in the length of foldamer **n**, being **n** one of the foldamers presented in Figures 2.2 and 4.2. For example, a foldamer like **3** is the foldamer **3a** (and vice versa), because the only change is the addition (removal) of a residue at the end, thus changing from 5-mer (6-mer) to 6-mer (5-mer).



(a) Reaction enthalpy (b) Activation energy

FIGURE 4.3. Reaction enthalpy and activation energy of the series of **3**-like and **4**-like foldamers as a function of the number of residues. To compare, in transparent line are represented the results obtained for **1**-like foldamers, which were reported in Figure 3.4 on page 62.

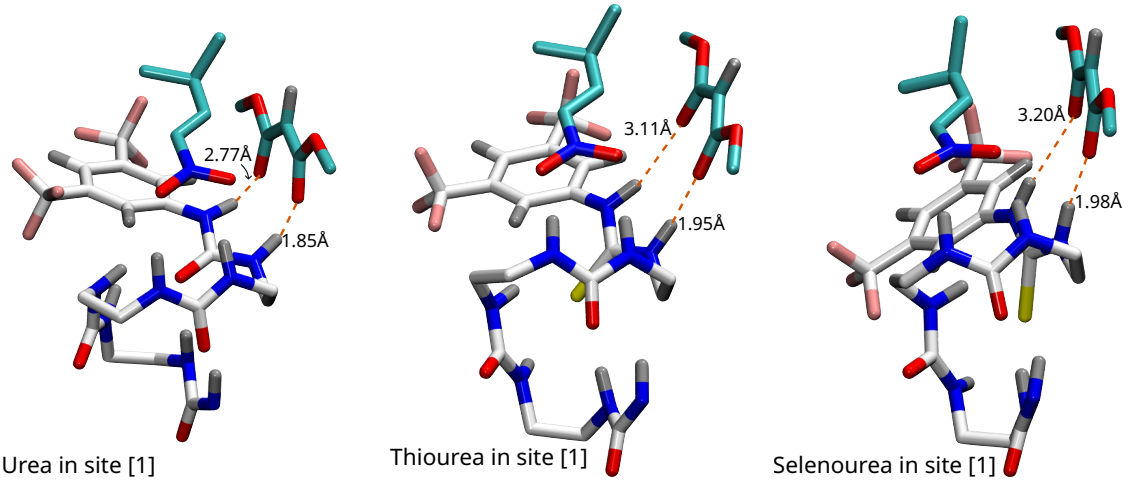
—in line with the experimental measurements—.

In the case of the substitution of Oxygen by Selenium, the activation energy in the 3-mer suggests that the formation of the enantiomer S should be slightly favored, however, for the 5-mer it is the enantiomer R which has lower E_a value, and thus the enantiocontrol characteristic of these foldamers should be regained.

Analysis of the system geometry

We next wondered if, as observed for **1**-like oligoureas in Chapter 3, the enantiocontrol provided by these foldamers could be related to their key geometrical parameters, and ultimately to the strength of interaction between site `[1]` and the dimethylmalonate. To that end, a comparison between foldamers **2**, **3** and **4** is presented in Figure 4.4.

It is observed that the introduction of the Sulfur atom replacing the Oxygen in site `[1]` has increased the distance of the H-bonds it establishes with the third urea. As a result, the thiourea in foldamer **3** is slightly more rotated counterclockwise than the urea in foldamer **2**, which results in a smaller space in the reactive region of the catalyst to accommodate the substrates. Furthermore, it is interesting to note that despite the increase in acidity gained by site `[1]` with the introduction of the Sulfur atom, the distances between dimethylmalonate and the foldamer have increased. This suggests that indeed, the "negative effect" induced by structural changes after the introduction of Sulfur should outweigh the positive ones induced by the increase in acidity.



(a) Reactants geometry of **2** (b) Reactants geometry of **3** (c) Reactants geometry of **4**

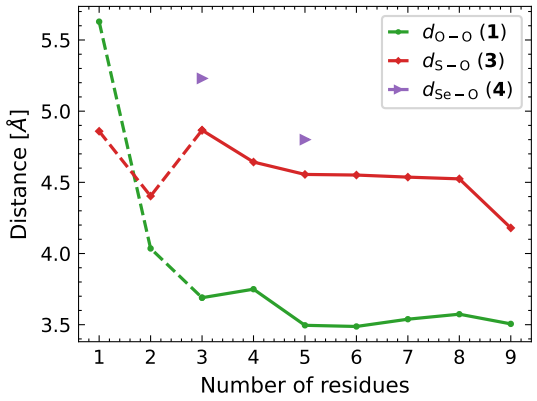
FIGURE 4.4. Geometry of the reactants in the foldamers **2**, **3** and **4**, represented side by side to identify structural changes which are consequence of the replacement of the Oxygen atom by a Sulfur and Selenium. Only 4 ureas of the foldamer (hexaurea) have been drawn, to gain in clarity.

Furthermore, the substitution with the Selenium atom (foldamer **4** and alike) shows similar structural changes than with Sulfur, but in this case the distances between site `[1]` and the dimethylmalonate are even longer, and the selenourea has a greater counterclockwise rotation.

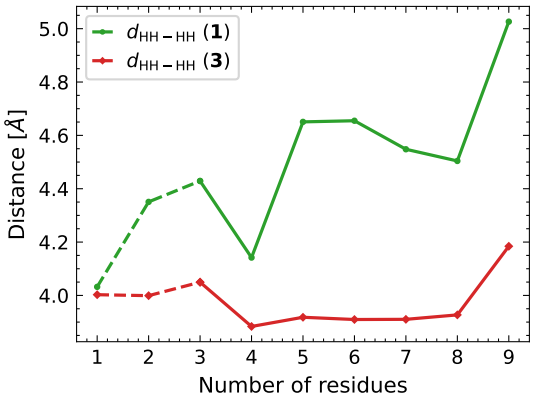
Therefore, it is possible to conclude that, generally, increasing the acidity of site `[1]` by replacing the Oxygen atom in **1**-like foldamers will probably introduce a "negative structural effect" which could outweigh the increase in interaction strength between the dimethylmalonate and the first urea, being detrimental to the catalytic performance.

This "negative structural effect" is evaluated by the space created between sites `[1]` and `[2]` in the reactants geometry. Furthermore, it can be characterized by (i) the distance $d_{\mathrm{X-O}}$ (X = S, Se) between the Sulfur/Selenium atom of the first site and the Oxygen atom of the second urea; and by (ii) the distance $d_{\mathrm{HH-HH}}$ between the mid point of the two NH groups in the first and second ureas (see Figure 4.5c on the following page). Similarly to the analysis of Figure 3.12 on page 71, the shorter (larger) the $d_{\mathrm{X-O}}$ ($d_{\mathrm{HH-HH}}$) distance, the greater the space created in the catalyst reactive region, and the better the catalytic performance of the foldamers should be.

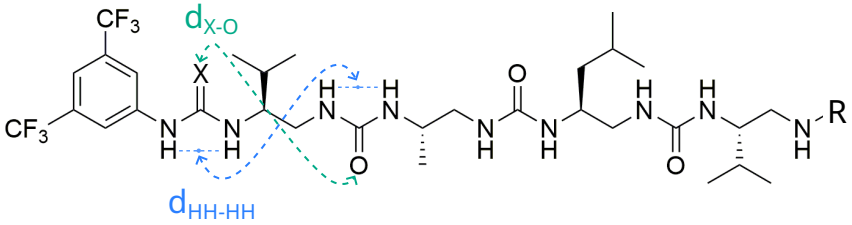
The measurements revealed that the distances $d_{\mathrm{S-O}}$ and $d_{\mathrm{Se-O}}$ decrease when the number of residues of the foldamers increases, as is the case of $d_{\mathrm{O-O}}$ in foldamers like **1** (see Figure 4.5a). Therefore, it can be seen from the figure that foldamers



(a) d_{O-O}



(b) d_{HH-HH}



(c) Schematic representation of the distances

Figure 4.5. Key distances characterizing the space created in the reactive region of the catalysts for foldamers like **1**, **3** and **4**, as a function of their number of residues. These values correspond to the geometry of the reactants.

from 5-mer to 7-mer should have more space available in its reactive region to accommodate the substrates, and consequently greater catalytic performances they should offer over shorter versions. Also, the fact that $d_{O-O} < d_{S-O} < d_{Se-O}$ reflects that the greater the atoms substituting the Oxygen in **1**-like foldamers, the larger the structural changes.

Concerning the distances d_{HH-HH}, in the case of foldamers like **1** they increase as d_{O-O} decrease (see Figure 4.5b). However, in the case of foldamers like **3**, d_{HH-HH} slightly decreases while d_{S-O} also decreases. This suggests that, as previously observed for enantiomers R and S of **1**-like foldamers in the last paragraph of section 3.5.3 on page 70, not only a counter- or clockwise dihedral rotation of site `[1]` is involved in the generation of space in the reactive region, but also the angle that moves the first two ureas closer or farther apart one from the other. Anyhow, a simplification of the structural movements identified in the reactive region of the catalysts is depicted in Figure 4.6.

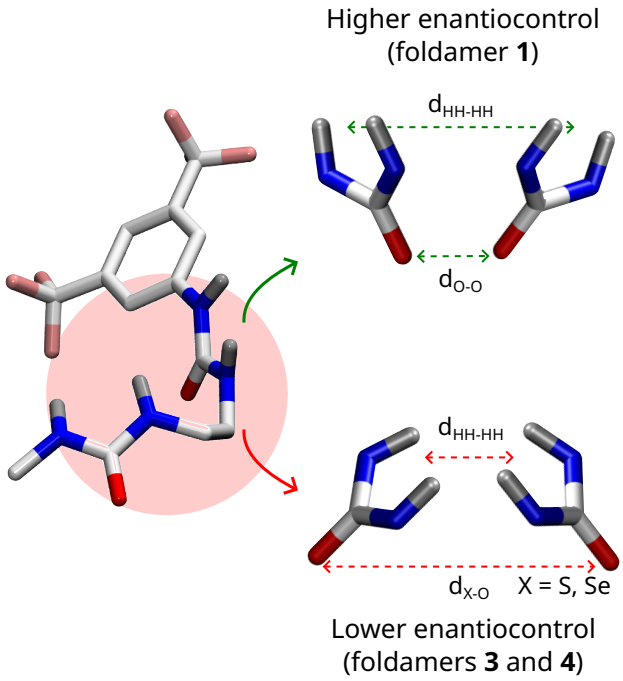


Figure 4.6. A simplification of the motion identified in the first two sites to be correlated with enantiocontrol in the foldamers.

Interaction strength between site [1] and dimethylmalonate

Ultimately, the increase or decrease of space in the reactive region of the foldamers will affect its interactions with the dimethylmalonate to a much greater extent than with the nitroalkene, because it is the first urea which is largely affected by modifications to the carbonyl group of site [1]. For this reason, we now investigate the interaction strength between catalysts and substrates with an emphasis on those maintained between the malonate and site [1].

The Grid Integration Score (GIS), defined by equation (3.3) on page 72, provides a measure of the stabilizing non-covalent interactions strength between two regions of a system, and therefore it is capable to embrace structural and electronic effects at the same time. Then, defining one interacting fragment as the catalyst reactive region and the other as the substrates (see Figure 3.14b on page 72), we could quantitatively measure if the introduction of the Sulfur and Selenium atoms will increase (or not) the interactions strength with the substrates. Recall that the GIS has already been used for the analysis of the transition states of the oligoureas in Figure 3.14a.

It was found that the non-covalent interaction strength increases with the number of residues in foldamers like **1**, **3**, and **4** (see Figure 4.7a). Also, the GIS index in foldamers like **1** (urea in the first site) is larger than in foldamers like **3** (thiourea in the first site) and **4** (selenourea in the first site). This is in agreement with

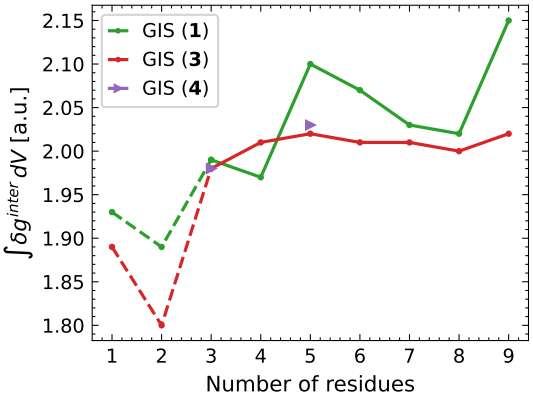
the geometrical parameters analyzed around Figure 4.5 on page 82, and suggests that despite the increase in acidity provided by the thiourea and the selenourea in site `[1]`, the interactions between the substrates and the catalysts are weaker than when the first site was a urea. This should be related to the detrimental changes introduced in the foldamers structure by such modifications.

Furthermore, its is also possible to isolate only those components coming from the H-bonds and thus estimate their contribution to the overall substrates/catalyst interactions. This is possible by employing the Intrinsic Bond Strength Index (IBSI) proposed by Klein et al. [236]. Based on the Integrated Gradient —electron density reference— Model (as the GIS index), the IBSI is a score designed to internally probe the strength of a given atom pair: the greater its value, the strongest the pair interaction. It is also implemented in the IGMPlot code [197, 234–236], which will be the choice in this investigation.

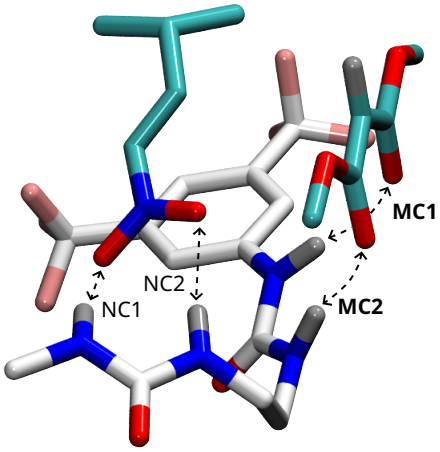
Note that the IBSI does not belongs to the conventional class of bond order (like Mulliken, Wiberg or Mayer, giving the number of electron pairs shared between two atoms), but rather assesses the intrinsic bond strength.

Figures 4.7c and 4.7d show the IBSI values obtained for the H-bonds between the substrates and the catalyst in the case of the foldamers with the urea (**1**) and the thiourea (**3**) in the first site, respectively. It is observed that the strongest interaction is established between the catalysts and the dimethylmalonate (MC2), and whereas its IBSI value remains nearly constant in the foldamers like **3**, it has a peak around the 5-mer and the 6-mer in foldamers like **1**. Also note that, overall, the MC1 and MC2 interactions are stronger in the foldamer with the urea than in the foldamer with the thiourea, suggesting one more time stronger H-bonds between the dimethylmalonate and the catalysts like **1**.

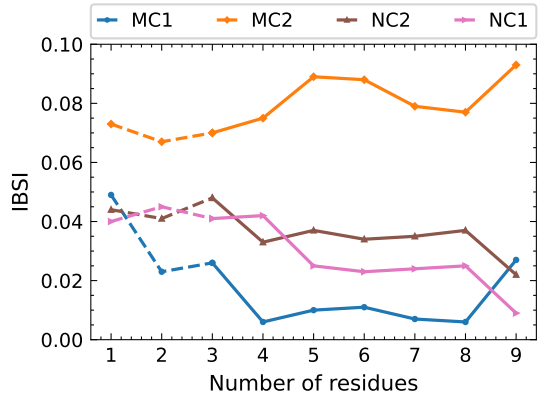
Finally, we computed the pK_a of the foldamers like **1** and **3** following the methodology discussed in Appendix D (see Figures 4.7e and 4.7f). As expected, (i) the `Hydrogen 1` is more acid than `Hydrogen 2`, because it is more close to the 3,5-bis(trifluoromethyl)phenyl group, and (ii) the **3**-like foldamers have lower pK_a values than **1**-like. However, the most interesting point to highlight here is that the pK_a value decreases also in correspondence with the number of residues of the foldamers, therefore, it could be related to the internal ureas cooperation in the helix.



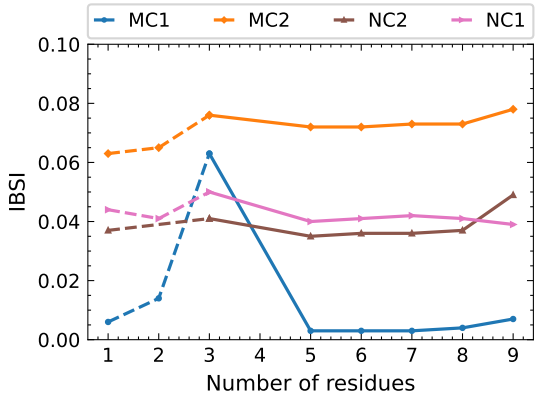
(a) GIS index



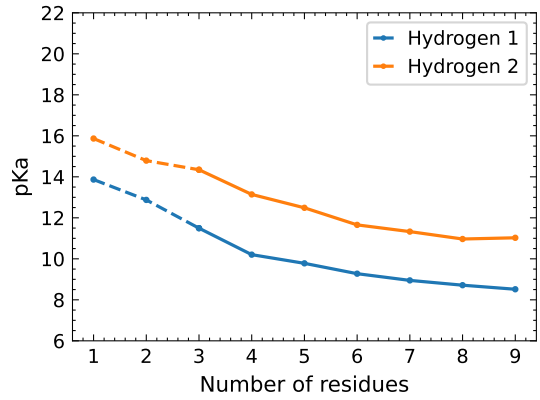
(b) Representation of interactions



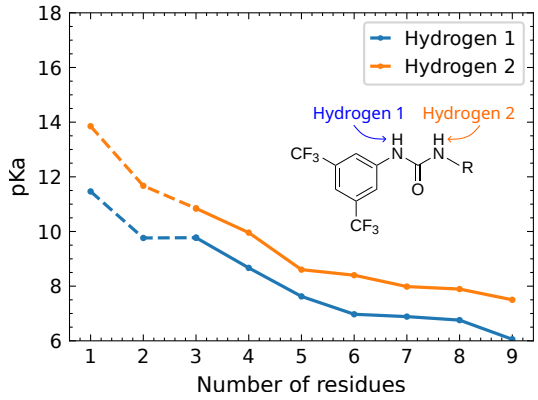
(c) IBSI index **1**-like



(d) IBSI index **3**-like



(e) Estimated pK_a for **1**-like



(f) Estimated pK_a for **3**-like

Figure 4.7. Parameters related to the interaction strength between the substrates and the catalysts. The Grid Integration Score, the Intrinsic Bond Strength Index, and the estimated theoretical pK_a values are reported.

4.3 Partial conclusions

From the energetic profiles, geometrical parameters, and characterization of the interactions between the catalysts and the substrates discussed in the previous section,

it is possible to conclude that there seems to exist a direct relationship between the catalysts performance and the interaction strength it establishes with the dimethylmalonate species.

The stronger are the non-covalent interactions between site [1] and the dimethylmalonate, and specially the MC2 H-bond, the more favored the enantiomer R will be,[2] probably resulting in better enantiocontrol. Indeed, one possibility to strengthen the MC1 and MC2 H-bonds is to increase the acidity of the NH groups of the first site. On the one hand, this practice reveals satisfactory when increasing the number of residues of the same foldamer as a mean to increase acidity. On the other hand, modifications made to the first site were shown to also increase its acidity, but introduced at the same time structural changes which seemed to counteract whatever positive effect the acidity could achieve.

These observations suggest that not only the acidity of the foldamers is determinant in their catalytic performance, but also this space generated by the first two ureas to accommodate the substrates, which is represented in our analysis by the distances $d_{\mathrm{O-O}}$ and $d_{\mathrm{HH-HH}}$. If one manages to have a foldamer with the needed characteristics to strongly bind the dimethylmalonate in site [1], but in fact the dimethylmalonate molecule cannot easily access the site because of repulsive interactions with the rest of the system, then a lower performance than expected could be obtained.

According to our studies, keeping the Oxygen atom in the first urea carbonyl group guarantees a correct helix conformation and space in the reactive site of the foldamers. Furthermore increasing the number of residues to at least 5 should reach the needed acidity of site [1] to strongly enough bind the malonates and conduct the catalytic process successfully.

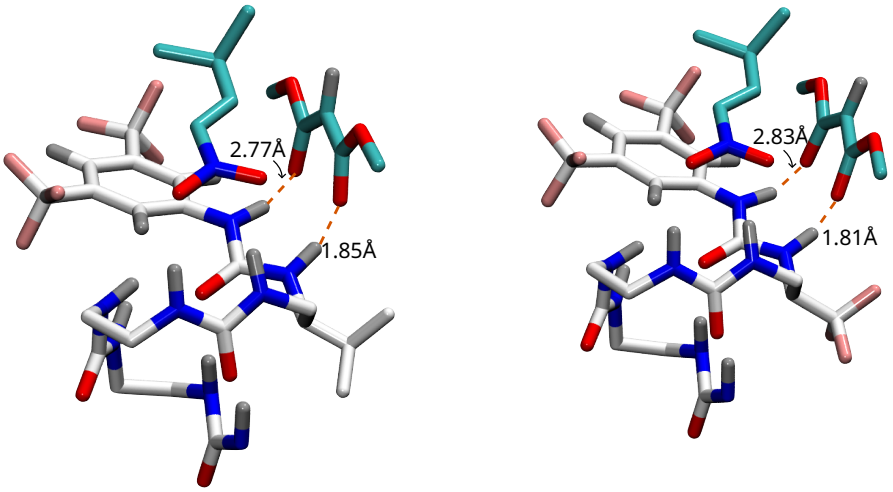
In the next sections it will studied how changes to the first side-chain functional group influences the acidity and structure of the foldamers.

4.4 Replacing first side-chain isopropyl by CF_3

The replacement of the first side-chain isopropyl group in **1**-like foldamers by the CF_3 functional group (see foldamer **5** in Figure 4.2) seems not to introduce considerable structural changes to the helix as foldamers **3** and **4** did. In fact, the distance $d_{\mathrm{O-O}}$ was slightly lower for foldamer **5** than for **2** (see Figure 4.9b). Furthermore, similar distances than in **2** were measured describing the interaction between the

[2] Recall that in enantiomer R has larger interactions with site [1] than enantiomer S, therefore it will be favored by strengthening them (see their geometries in Figures 3.11b and 3.11c on page 69).

deprotonated dimethylmalonate and site [1] in the geometry of the reactants (see Figure 4.8).



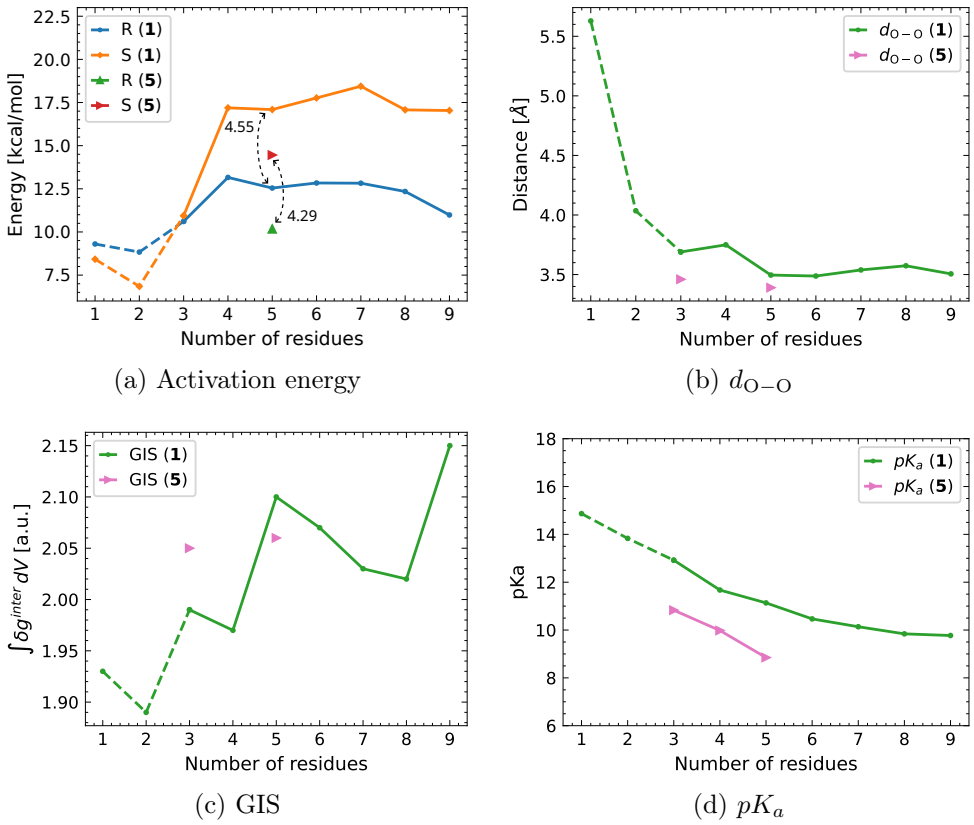
(a) Reactants geometry of **2** (b) Reactants geometry of **5**

FIGURE 4.8. Geometry of the foldamers **2** and **5**. Only 4 ureas of the foldamer (hexaurea) and the pertinent functional groups have been drawn, to gain in clarity.

This shorter MC2 distance should be related to the increase in acidity of site [1] induced by the CF_3 group, which has larger capabilities to attract electrons than isopropyl (see Figure 4.9d). However, despite the —slightly— shorter d_{O-O} values in foldamer **5**, and the larger acidity of the foldamer, the non-covalent interactions between foldamer and substrates are lower than in foldamer **2** (see Figure 4.9c).

Finally, note that despite the decrease observed in pK_a and d_{O-O} values for foldamer **5**, the enantiocontrol was very similar to foldamer **2**, because a difference of 0.26 $kcal/mol$ is insignificant. This behavior should be related to the fact that the foldamer **2** is already extremely efficient in its function, therefore, some increase in enantiocontrol by any modification made to it will probably be trivial. Detrimental changes like those observed in the case of foldamers **3** and **4**, on the other hand, should be easily identified, and also their reduction of catalytic performance.

Overall, these parameters measured for the foldamer **5** seems to be very similar than those of foldamer **2**, which suggests that changes to the first side-chain functional group like the one here presented should have a small positive effect on the structural parameters of the helix, and the interactions between catalyst and substrates.



(a) Activation energy (b) d_{O-O}

(c) GIS (d) pK_a

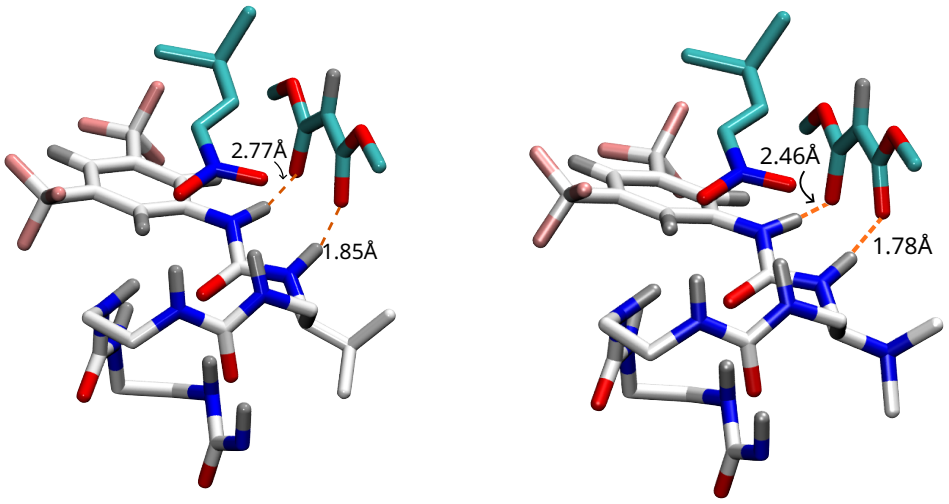
FIGURE 4.9. Representation of the key parameters related to the enantiocontrol of foldamers like **5**, compared with those of foldamers like **1**.

4.5 Replacing first side-chain isopropyl by $N(CH_3)_2$

Additionally, another substitution to the first side-chain functional group of foldamers like **1** was made. In this case, the isopropyl was changed by a $N(CH_3)$ group (see foldamer **6** in Figure 4.2), which should, similarly to the CF_3, attract electrons and acidify the site `[1]`.

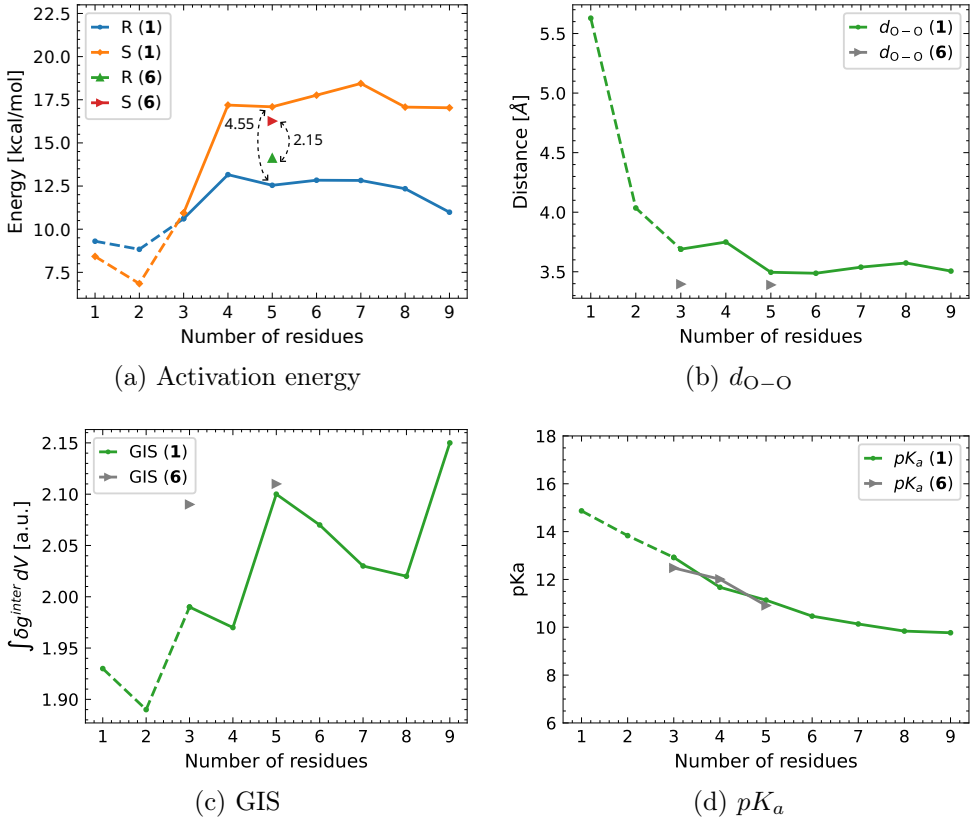
This substitution strengthened the interactions between the malonate and the first urea, as can be observed from the smaller `MC2` H-bond distance (see Figure 4.10b). Furthermore, also the d_{O-O} distances are slightly lower in foldamer **6** than in **2** (see Figure 4.11b), and as consequence, the stabilizing non-covalent interactions established between the catalyst and the substrates are larger for **6** (see Figure 4.11c). Concerning the pK_a values, their estimated values for foldamer **6** were very similar to those found in foldamer **2**, then, any variation in catalytic performance should come almost entirely from structural changes of the reactive site, more precisely the d_{O-O}. This supports the hypothesis that not only the acidity of the reactive site is important but also its geometry.

Nevertheless, the activation energy gap between enantiomers R and S was smaller



(a) Reactants geometry of **2** (b) Reactants geometry of **6**

FIGURE 4.10. Geometry of the foldamers **2** and **6**. Only 4 ureas of the foldamer (hexaurea) and the pertinent functional groups have been drawn, to gain in clarity.



(a) Activation energy (b) d_{O-O}

(c) GIS (d) pK_a

FIGURE 4.11. Representation of the key parameters related to the enantiocontrol of foldamers like **6**, compared with those of foldamers like **1**.

for enantiomer **6** than for **2**. Therefore, it is possible to conclude that this new modification to the isopropyl functional group introduced small positive changes in the geometry of the helix, and in the interactions malonate/catalyst, which ultimately resulted in slightly lower values of the computed enantiocontrol. Note that even with the obtained lower activation energy gap, the enantiomer R is considerably favored over enantiomer S in foldamer **6**.

4.6 Other replacements made in the experiment

In the experimental work by Bécart et al. [29], they made several other modifications to the foldamers in order to identify a structure/catalytic activity correlation (see Figure 4 and Table 4 in their publication). Three of these modification are in line with those so far investigated in our work, i.e are located at the first side-chain functional group (see foldamers **17**, **18**, and **19** in their publication). For this reason, these modified foldamers (see Figure 4.12) were selected to compare the experimentally obtained data from them with the computational parameters identified in this work to be related with the enantiocontrol.

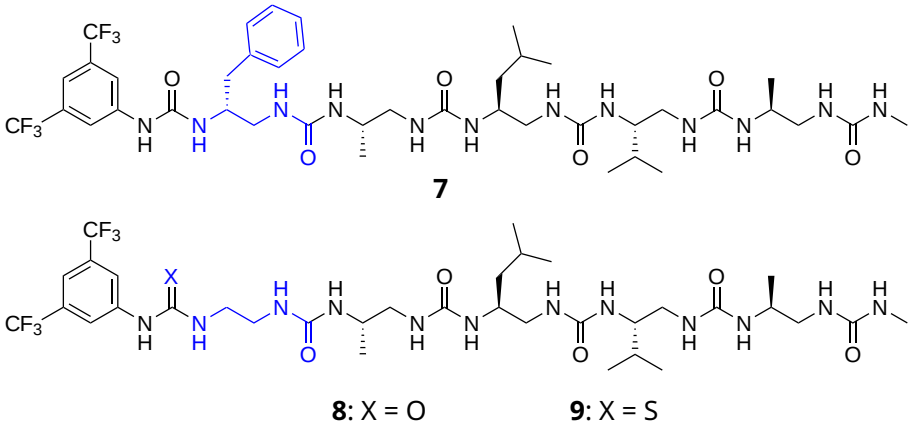


Figure 4.12. Modifications made by Bécart et al. [29] to the first side-chain functional group of foldamer **1** in order to investigate their influence on the enantiocontrol and reactivity.

Table 4.1 report a compilation of computed structures/estimated pK_a/experimental catalytic activity related parameters for a set of foldamers (including those mentioned in the above paragraph). Note that the experimental *ee* values were measured in all cases on a system set to $-20\,°C$ utilizing β-nitrostyrene as nitroalkene in the C-C bond forming reaction, and our computations were carried out at room temperature and employing $(1E)$-3-methyl-1-nitrobut-1-ene as nitroalkene. For such reason, the analysis is generalized here, focusing on the overall

catalytic performance of the foldamers, which includes not only enantiocontrol, but also conversion, yield, catalyst loading, and reaction time.

		Computational			Experimental
Entry	Foldamer	Estimated pK_a	d_{X-O} [Å]	X	*ee* (%)
1	**1**	10.47	3.49	O	95
2	**2**	11.13	3.50	O	$90 < ee < 95$
3	**3**	8.11	4.56	S	—
4	**3a**	7.69	4.55	S	93
5	**7**	10.87	3.40	O	88
6	**8**	10.72	3.46	O	83
7	**9**	7.02	4.66	S	35

Table 4.1. Relationship between the pK_a, the space created in the reactive region of the catalysts, and their experimentally obtained enantiomeric excess. The d_{X-O} values were measured in the computationally optimized geometry of the reactants, and the *ee* was obtained from experiments carried out using Et_3N as achiral base and toluene as solvent.

It is observed that trying to increase the H-bond donor character of site `[1]` by replacing the Oxygen atom of its carbonyl group by a Sulfur, overall resulted in a lost of catalytic properties (compare entries 1 and 4, for example). If it is true that such modifications considerably lowered the pK_a of the foldamers, it increased at the same time the distance d_{X-O}, reducing the capabilities of the malonate to utilize the stronger H-bond donor character of the site.

In general, foldamers **1**, **2**, **7**, and **8**, which have a urea in site `[1]`, the pK_a values are larger than for those with a thiourea, but the space in the reactive region is wider (lower d_{O-O} values), and therefore better catalytic performances are measured. Note how in the foldamers with a urea in site `[1]` the acidity is gained by the addition of residues to the foldamers, coming from the internal ureas cooperation established in their helix.

Finally, a closer look to the entries 4 and 7 reveals that with similar pK_a values and d_{S-O} distances, the *ee* in **9** is considerably lower. This suggests that there could exists other(s) parameters, in addition to the pK_a and the distance between ureas, which could influence the catalytic performance of the foldamers, despite those two showed a correct consistency during all previous analyses. Nevertheless, it has to be taken into account that there exists important differences between the computational and experimental conditions. Additionally, **9** was the only foldamer which suffered modification at both the first urea Oxygen and the first side-chain functional group. For these reasons, further investigations specific to this system where new experiments and computations are closely related could be needed.

4.7 Conclusions

A set of modified foldamers derived from **1** were studied in this chapter with the objective to identify how a variation in acidity of site [1] (pK_a) could change the interactions between the catalysts and the substrates, and its influence on structural parameters as the distance between the first two sites (d_{O-O} and/or d_{HH-HH}) and the helix conformation. Two modifications were located in the carbonyl group of the first urea, and the remaining in the first side-chain isopropyl functional group.

It was identified that ultimately, the catalytic performance of the foldamers should be determined by the strength of the H-bond interactions it establishes with the deprotonated malonate. To larger stabilizing interactions, greater enantiocontrol should be obtained from the system.

Two general ways to increase the acidity of site [1] were tested. First, the substitution of the Oxygen atoms in the carbonyl group of the first urea by a Sulfur and Selenium. Despite the acidity of site [1] was estimated to indeed increase, the structural changes the helix suffered to adjust to bigger atoms seemed to counteract the positive effect of the acidity.

The second set of variations, on the other hand, was focused on the first side-chain functional group. It was observed that the acidity of the first site also increased as result of introducing electron attractive functional groups at the mentioned position, while at the same time the structural characteristics the modification induced on the helix was in fact supportive to sustain an excellent enantiocontrol. Therefore, the increase of the donor character of site [1] in future tunings of these foldamers should be focused on the first side-chain functional group, and not in the carbonyl group of the first urea, to avoid detrimental structural changes to their helix geometry. Foldamers **5**, **6**, and **7** made clear that modifying the first side-chain with electron attractive functional groups have small influence on the internal ureas cooperation and configuration of the helix, while improving the properties related to catalytic performance. Modifications to the carbonyl group of the first urea, on the other hand, always lead to a rearrangement of the helix with structural modification in the reactive region, and the consequent weakening of the interactions malonate-catalyst.

Finally, it must be pointed out that to improve the catalytic performance of foldamers like **1** and **2** is no easy endeavor, because they already perform exceptionally. Nevertheless, valuables insights were provided by this investigation in the understanding of the nature of the key parameters related to such performances.

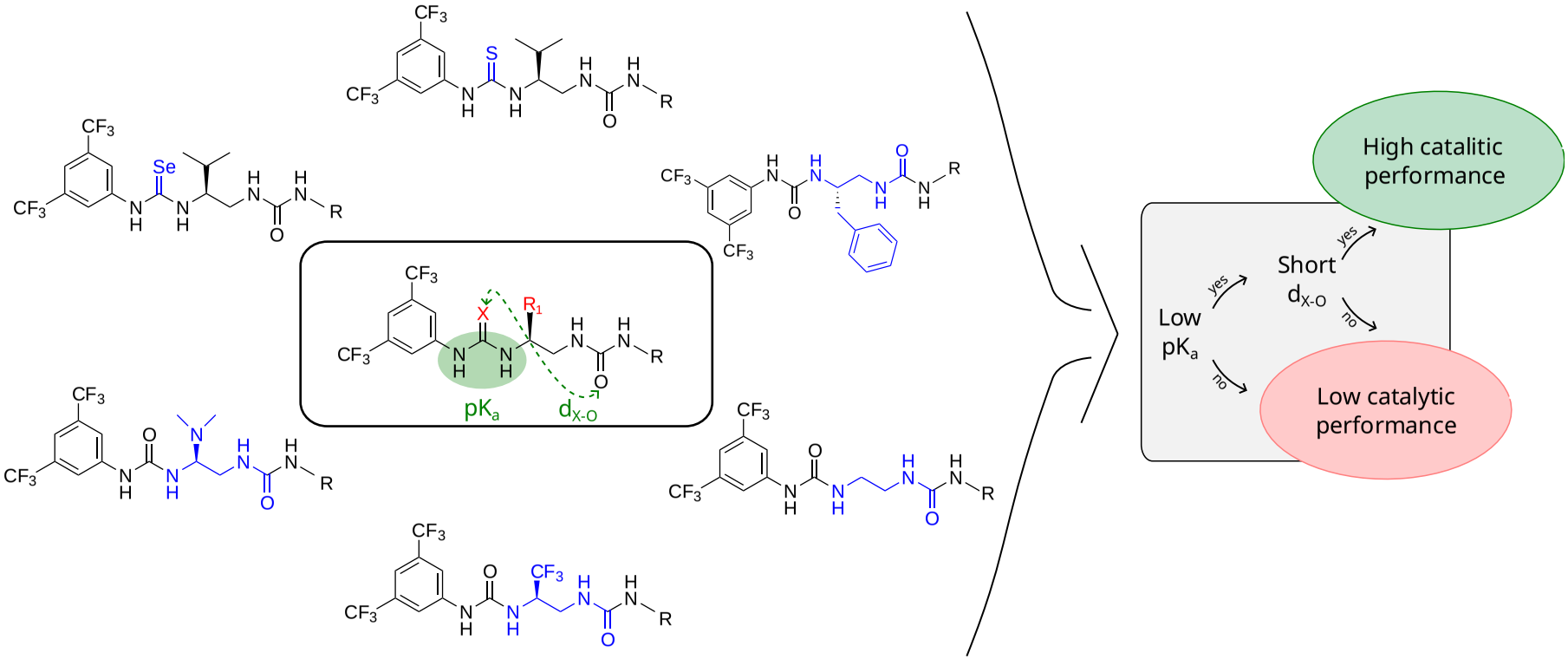


Figure 4.13. Simplified representation of the relation between low pK_a values of site [1] and structural parameters as the space crated in the reactive region of the catalysts, and catalytic performance of the studied foldamers.

General conclusions

The studies presented in this manuscript constitute the first theoretical investigation directly tailored to the description and rationalization of the basic phenomena taking place in the conjugate addition of 1,3-dicarbonyl substrates to nitroalkenes, when catalyzed by chiral aliphatic N,N'-linked oligoureas.

A DFT based Global Optimization procedure was first employed to identify the global minimum in the Potential Energy Surface of the system composed by the catalyst, the achiral Brønsted base, and the substrates. The outcome of this specific study is of great importance, because not only it revealed the positioning of the molecules in the systems, but also reciprocally supported experimental data from $^1\mathrm{H}-\mathrm{NMR}$ measurement regarding the substrates-catalysts interactions. The $(1E)$-3-methyl-1-nitrobut-1-ene was identified to utilize the NH groups of the second site urea as H-bond donors, while the dimethylmalonate was located in the first site urea throughout similar interactions. With this configuration, the reaction pathway between the two carbons is shorter and easier than other arrangements. This information made possible the identification of a bifunctional activation mechanism facilitated by the foldamer's first two ureas in the catalytic process, which is linked to the accommodation of the substrates in a perfectly suited position, one with respect to the other, for the reaction to occur.

It was observed that a positive charge concentration is located around sites 1 and 2 of the foldamers, which value and area increase with the number of residues. This is result of the internal ureas cooperation in the helix, creating a macrodipole. Following this line, it was checked that despite the electrostatic potential maximum value seems to indefinitely augment with the number of residues of the foldamer, its polarization (Π) and tendency to interact with negative poles of other molecules (σ_+^2) posses an asymptotic behavior. This suggests that at some point the increase in the number of residues will not considerably increase the electrostatic properties which influences the catalyst-substrates interactions. Additionally, the positively charged region extends beyond the first two sites in foldamers larger than the 8-mer, and this could eventually lead to the attraction of the substrates for sites other than the

two reactive ones. Then, the indefinitely increase of the number of residues should lead to a lower system reactivity.

Based on the previously obtained information about the catalytic system, the computations of the energetic profiles for a set of size-varying foldamers was carried out. It was found that, indeed, the enantiocontrol provided by these foldamers in the specific reaction here studied starts to be acceptable with 4, 5 and 6 residues long oligomers. Shorter molecules were identified to have unfavorable structural and electronic characteristics, as for example the stabilizing non-covalent interactions between the catalyst and the substrates, and the distance between the two reactive carbons in the transition states geometries.

At this point, a strong correlation between the catalyst performance and the space created in its reactive site from the orientation of the ureas was found. If the NH groups of the first two ureas were pointing one towards the other, then the space created for the accommodation of the substrates was smaller than in the case where they were oriented in a more open way, and thus the stronger the repulsive interactions as consequence of the competition for space. This detrimental effect seems to even counteract the positive impact that was expected from an increase in acidity of the first site. It is possible that ultimately, a relation between the mentioned space in the reactive site and the donor character of the first site are the two parameters defining the catalytic performance. If one manages to increase the H-bonds interaction strength of site `[1]` with the dimethylmalonate, while maintaining enough space for it to effectively utilize this acidification, then a correct catalytic performances should be expected from these foldamers.

Finally, several modifications made to the foldamer with the intention to study each one's influence on the properties related to the catalyst performance revealed that substitutions made to the first side-chain isopropyl group will have less influence on the helix, and thus should have minor (or none) detrimental consequences to the catalyst geometry and performance. On the other hand, changes to the first urea does considerably affects the helix of the foldamers and induces an adverse effect on the performance by increasing the repulsive interactions of the rest of the system with the malonate and destabilizing the geometries.

In addition to the interesting applications of the product of the reaction here studied as building blocks for pharmaceutical developments [240], the investigated catalysts have been also proven capables to catalyze certain polimerization processes in the production of polylactides [241]. If done efficiently, they could be an interesting alternative due to current state of the art catalysts in the production of biosourced and degradable materials as replacement for conventional petroleum-

based plastics.

Zaky [241] in the associated group of Prof. Daniel Taton at the University of Bordeaux have recently found that oligoureas similar to those investigated here performed slightly worst than mono(thio)ureas like Takemoto's [30], however, with further improvements they could perform better and reduce the loading of catalysts needed to fulfill their function, just as they do for the conjugate addition of 1,3-dicarbonyl substrates to nitroalkenes. It is in our hope that the information here provided concerning the catalytic behavior and characteristics of these chiral aliphatic N,N'-linked oligoureas serve to develop further improvement, and the consequent application in technologies capables to reduce the waste of plastic based materials to the environment, utilizing more friendly alternatives as the polylactides.

Appendix A

Molecular recognition supporting data

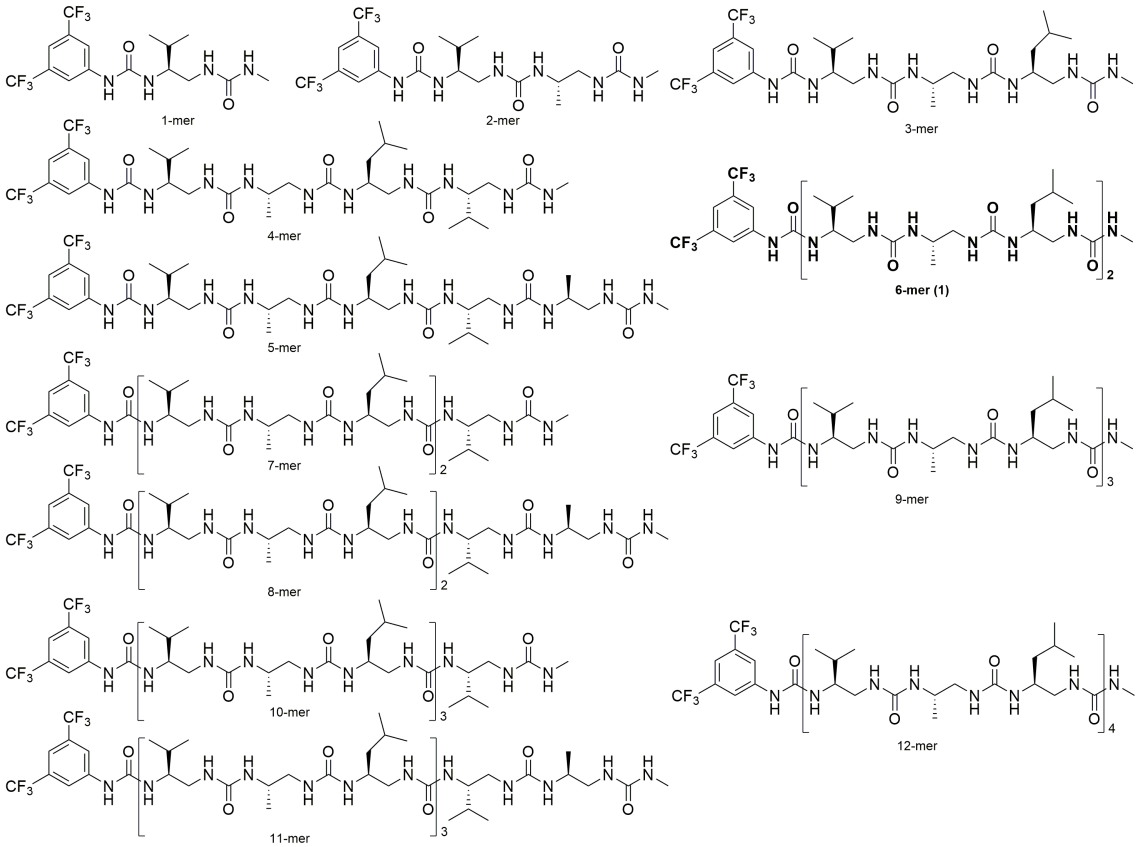
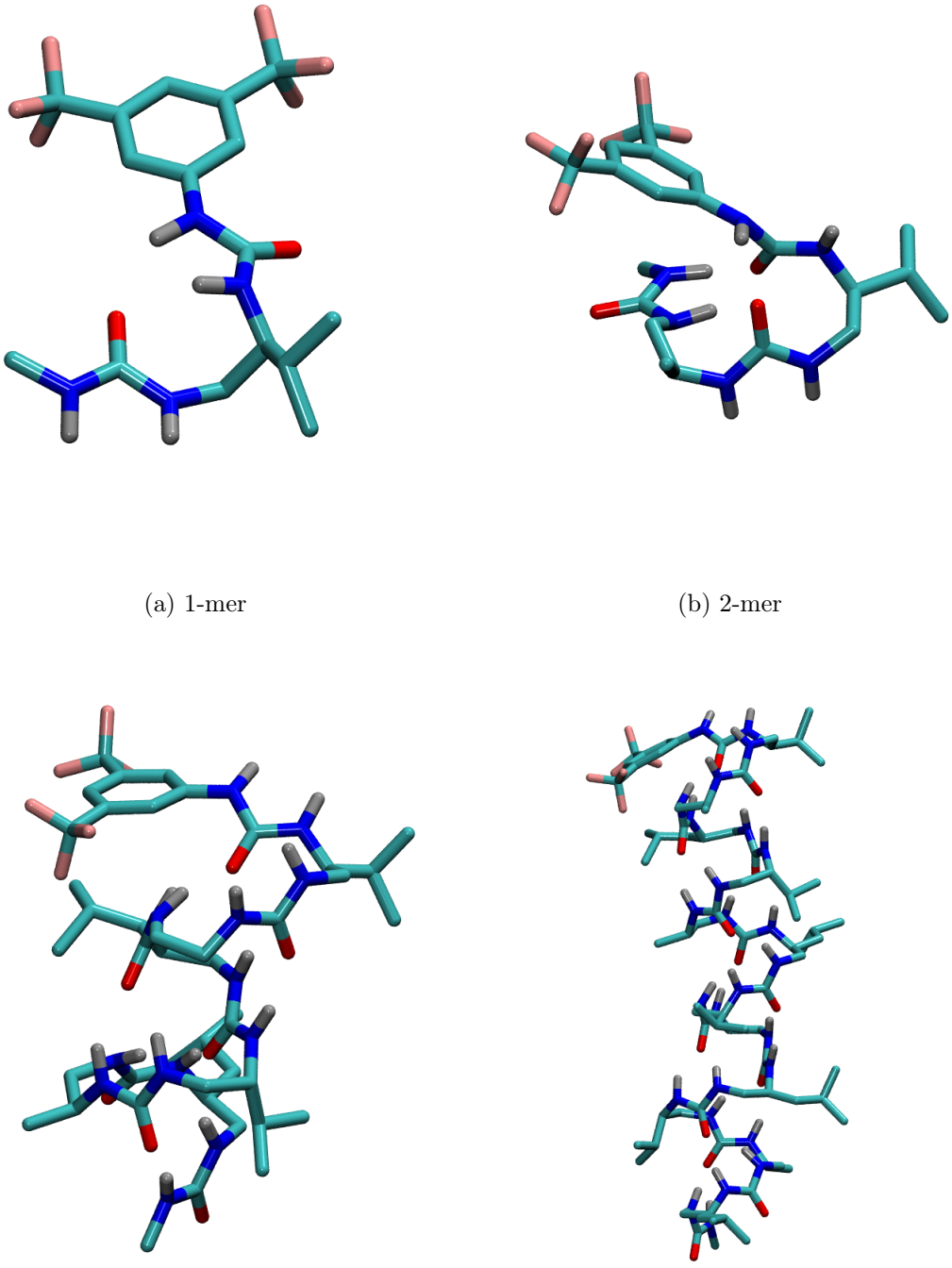


Figure A.1. Specification of the foldamers used in this investigation, from monomer to dodecamer.



(a) 1-mer

(b) 2-mer

(c) 6-mer

(d) 12-mer

Figure A.2. A selection of the foldamers represented in Figure A.1 showed in a 3 dimensional view.

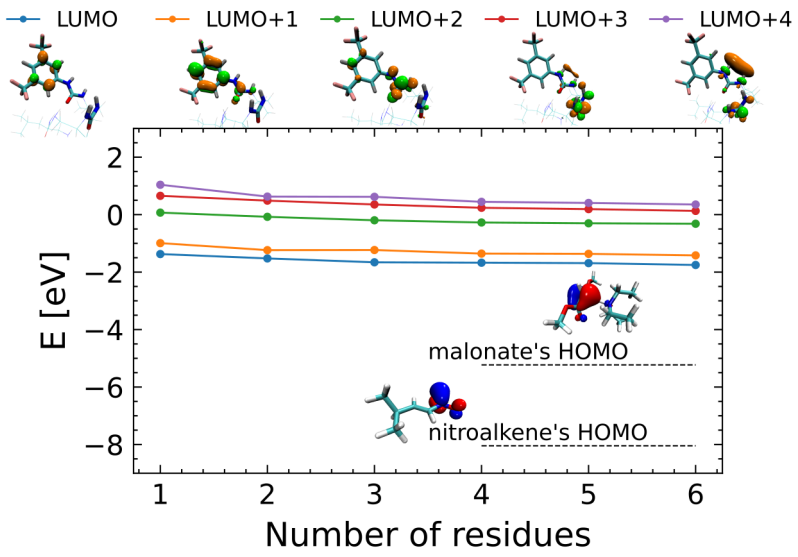


Figure A.3. HOMO(substrates) - LUMO(catalyst) gap with respect to the length of the catalyst.

Substrates-catalyst interactions as observed by NMR

Information

All the information presented in the following sections concerning the experimental measurements was obtained by Bécart et al. at the University of Bordeaux (author of the paper [29]), and was provided to us by Gilles Guichard as part of the internal collaboration of the ANR project this thesis is part of.

Oligourea **1** was fully characterized by NMR in a mixture acetonitrile-[d_3]/DMSO-[d_6] (88/12 v/v) and all the NH and NH' signals were unequivocally attributed in this solvent system (Figures A.4 and A.5). Three titration experiments were performed: (i) Addition of nitrostyrene to the NMR tube containing the catalyst to monitor the interaction between the nitrostyrene and the catalyst (Figure A.6) (ii) Addition of solution of malonate and triethylamine (molar ratio 1:1) (Figure A.7), and (iii) Successive addition of all the reactants to more precisely mimic the model reaction (Figure A.8).

Evaluation of the interaction between the catalyst and nitrostyrene

No visible change was observed in the $^1H-$ NMR spectra of **1** even when more than 100 added nitrostyrene equivalents were added in acetonitrile-[d_3]/DMSO-[d_6] (88/12 v/v). This tends to suggest that there is no strong interaction between nitrostyrene and the catalyst observable by NMR under these solvent conditions (Figure A.6). One must keep in mind that DMSO which is necessary to solubilize the oligomer also competes for binding to ureas.

Evaluation of the interaction between the catalyst and deprotonated malonate

Using the same experimental settings, an equimolar stock solution of malonate:Et_3N 1:1 was progressively added to an NMR tube containing catalyst **1**. Two signals were found to progressively shift: the most significant chemical shift difference cor-

responds to the (most acidic) proton NH' between the aromatic ring and the first catalytic site, the second one, still very clear, corresponds to the NH proton of the first residue Val^{uI} (Figure A.7).

Competition NMR experiments by successive addition of malonate, nitrostyrene and triethylamine to the catalyst

Here, we observed and variations of urea proton signals while adding the different substrates separately and successively to the catalyst **1**. Our experiment consisted of the successive additions of malonate, triethylamine and nitrostyrene, in this order (Figure A.8). The addition of malonate induced a very light downfield shift: only the NH' (indicated in light blue) in α-position of the aromatic group presented a notable shift of 0.0155 *ppm*. The addition of triethylamine induced a more pronounced downfield-shift of the NH' (indicated in light blue) in α-position of the aromatic group (0.0776 *ppm* from the reference spectra). The final addition of nitrostyrene impacted again the NMR spectra with a downfield shift: 0.1308 *ppm* for NH' signal from the reference catalyst spectra (indicated in light blue) and 0.0552 *ppm* for $NH^{Val\ I}$ (indicated in red).

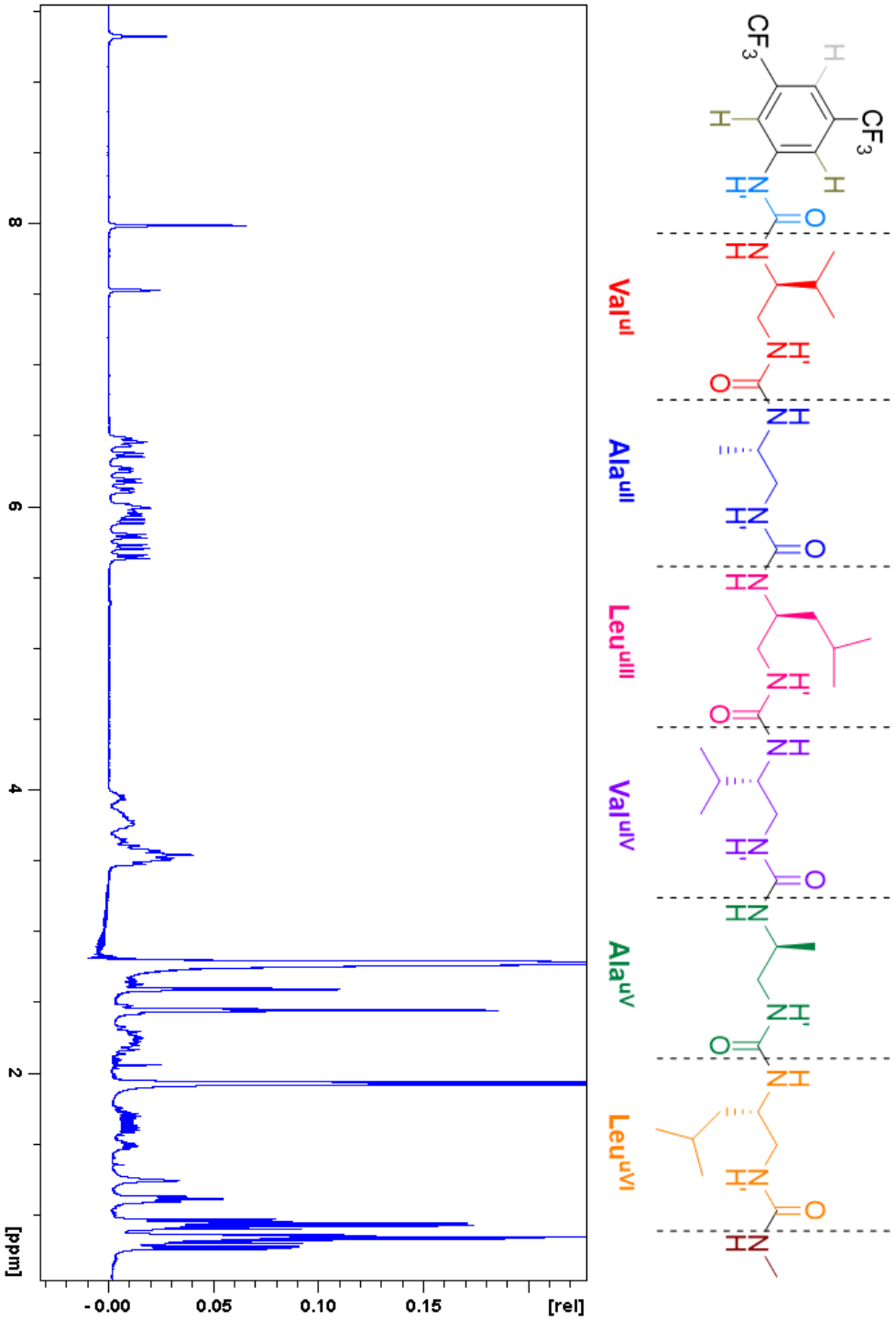


FIGURE A.4. $^1H - NMR$ spectrum of **1** in acetonitrile-[d_3]/DMSO-[d_6] (88/12 v/v).

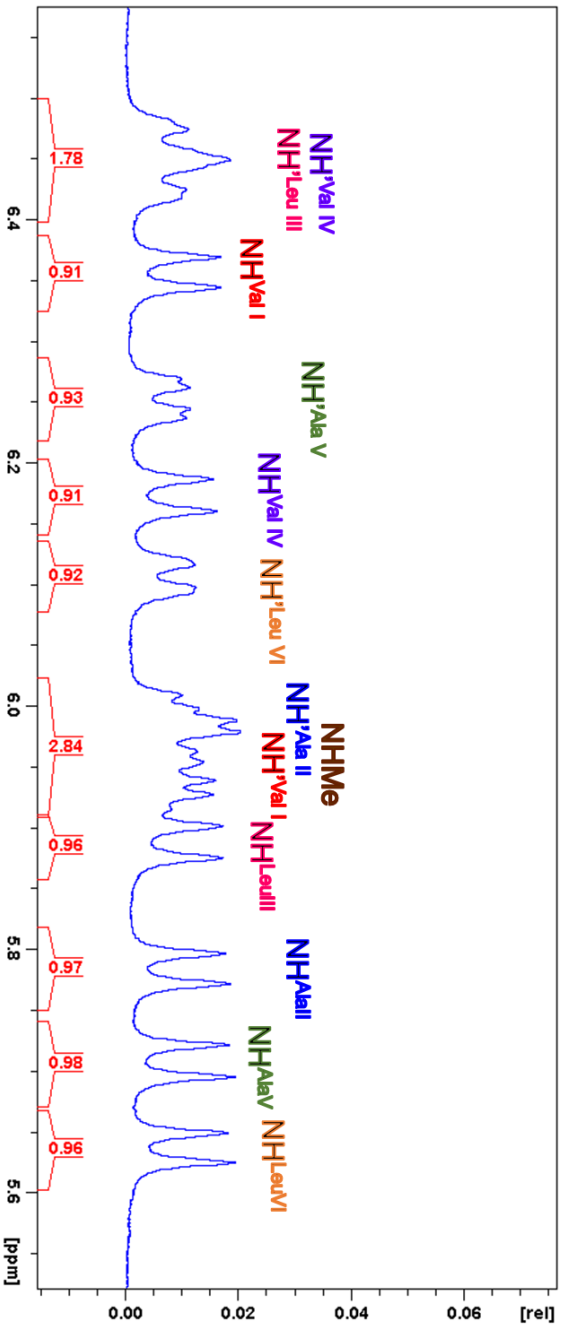


FIGURE A.5. Full attribution of NH and NH' region in ^1H-NMR spectrum of **1**.

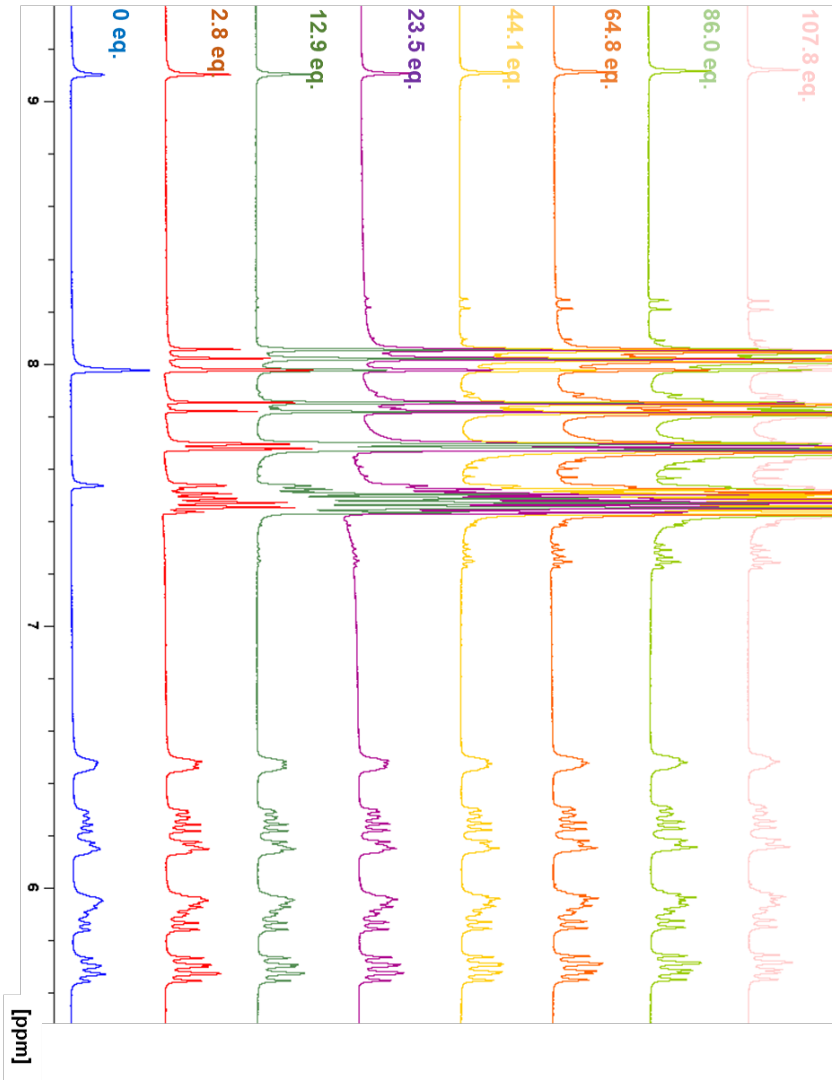


Figure A.6. Titration of **1** with nitrostyrene.

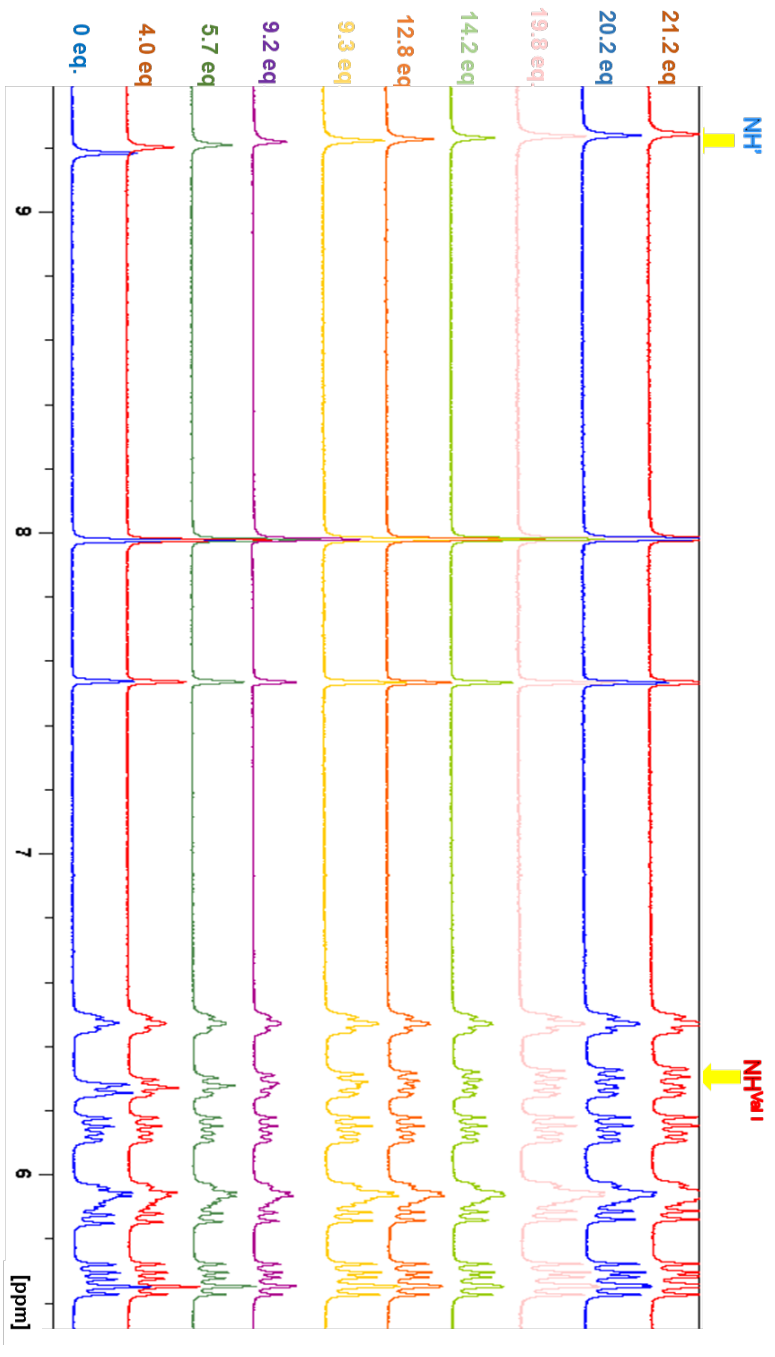


FIGURE A.7. Titration of **1** with deprotonated malonate.

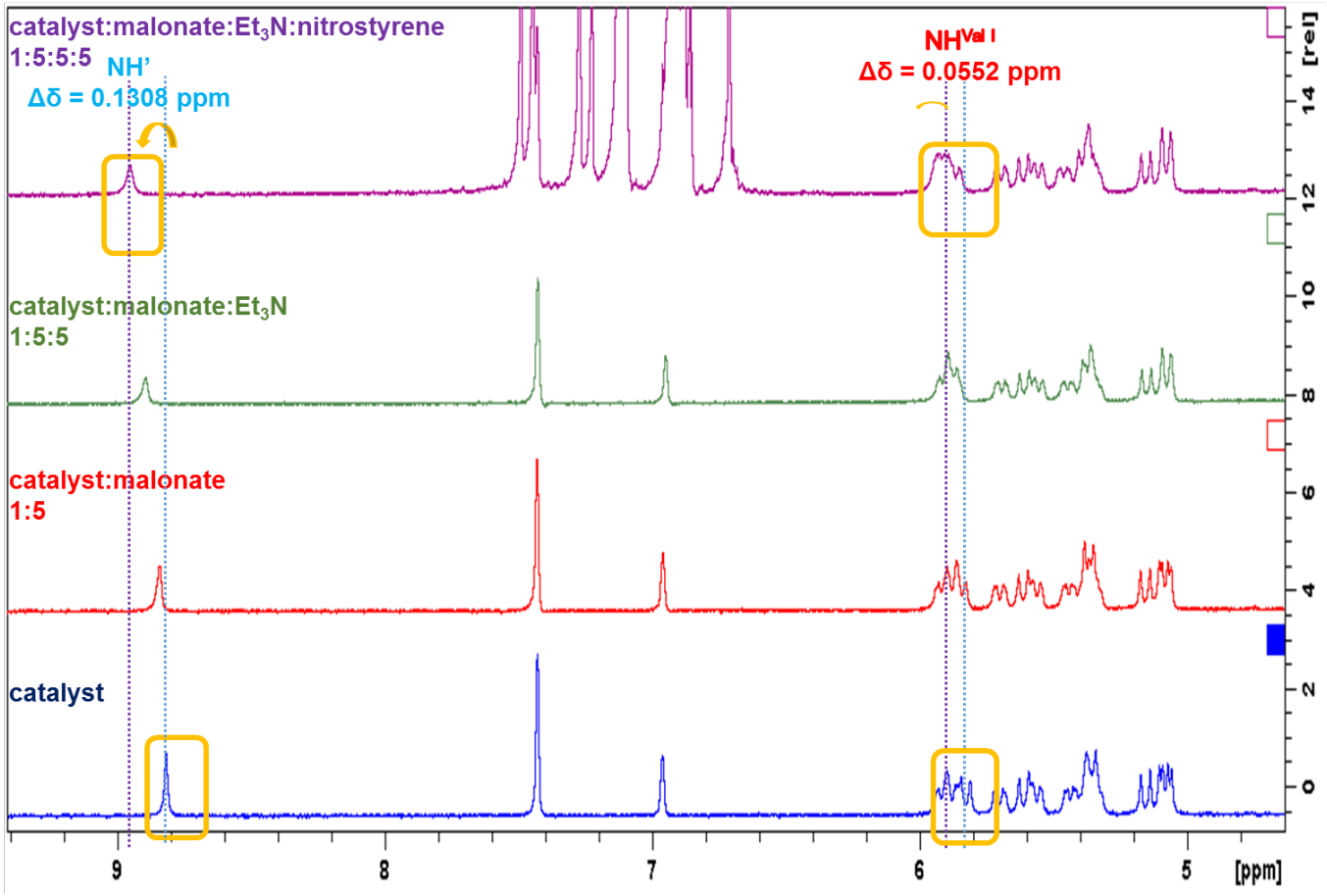


Figure A.8. $^{1}H-NMR$ analysis of interactions between catalyst and malonate, triethylamine, nitrostyrene. Experiments conducted on 300 MHz spectrometer, in DMSO-d_6:CD_3CN (12% v:v), at 298 K.

Appendix B

Rotameric analysis of the foldamer

B.1 Introduction

Frequently, when working with molecules which have a considerable number of single bonded dihedral angles, an investigation of the conformational space is required to successfully describe the processes they take part on. This is because a dihedral angle could have a value which is not the most stable for it, but the system's energy is not enough to surpass the energetic barrier which separate one conformer from the other (see Figure B.1).

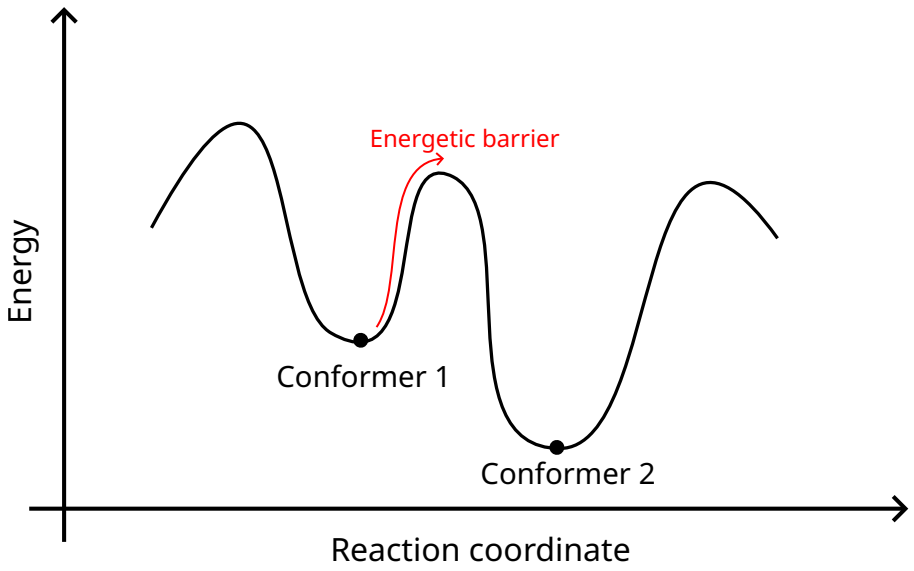


Figure B.1. Illustrative example of the energetic barrier separating two conformers.

A conformational search is a procedure in which most dihedral angles of a system are varied artificially and then the resulting structures are optimized. In this way, most of the conformational space is explored and the global minimum can be found.

As a consequence, the computational resources consumed for a conformational search are considerable, and a Molecular Mechanics method is usually employed to reduce time and requirements.

In some cases, a complete exploration of the conformational space of a system is not needed, because the involved dihedral angles will have a small impact on the system's function. In these situations, a simple scan of the specific coordinate which does influence the system's performance and function should be enough.

B.2 Monomer rotameric forms

It has been observed in the experiment [242, 243] that the two ureas in the monomeric form of the foldamers are oriented in opposite directions. To theoretically investigate this observation, a scan of the coordinate influencing the orientation of the ureas has been developed (see Figure B.2). It has been obtained that when the ureas are oriented in the same direction (structure 1), the monomeric foldamer is $\sim 8\ kcal/mol$ higher in energy than when the two ureas are pointing in opposite directions (structure 2). This behavior is due to the presence of two strong H-Bonds established in structure 2 in Figure B.2 which are not present in structure 1.

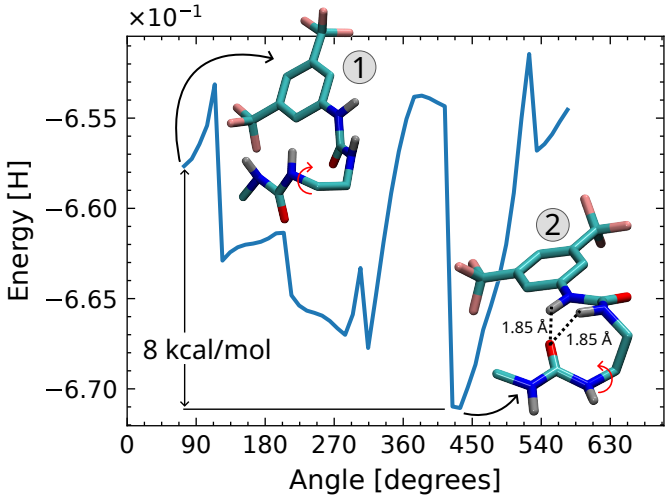


Figure B.2. Energetic profile of the coordinate describing the orientation of one urea with respect to the other in the 1-mer. The red arrows show the rotation studied.

Despite structure 2 being most stable, it is still possible that with the incorporation of the substrates it becomes disfavored. To also check this possibility, a Global Optimization procedure with the GSAM code was carried out taking the substrates and structures 1 and 2 as starting geometries (see Figure B.3).

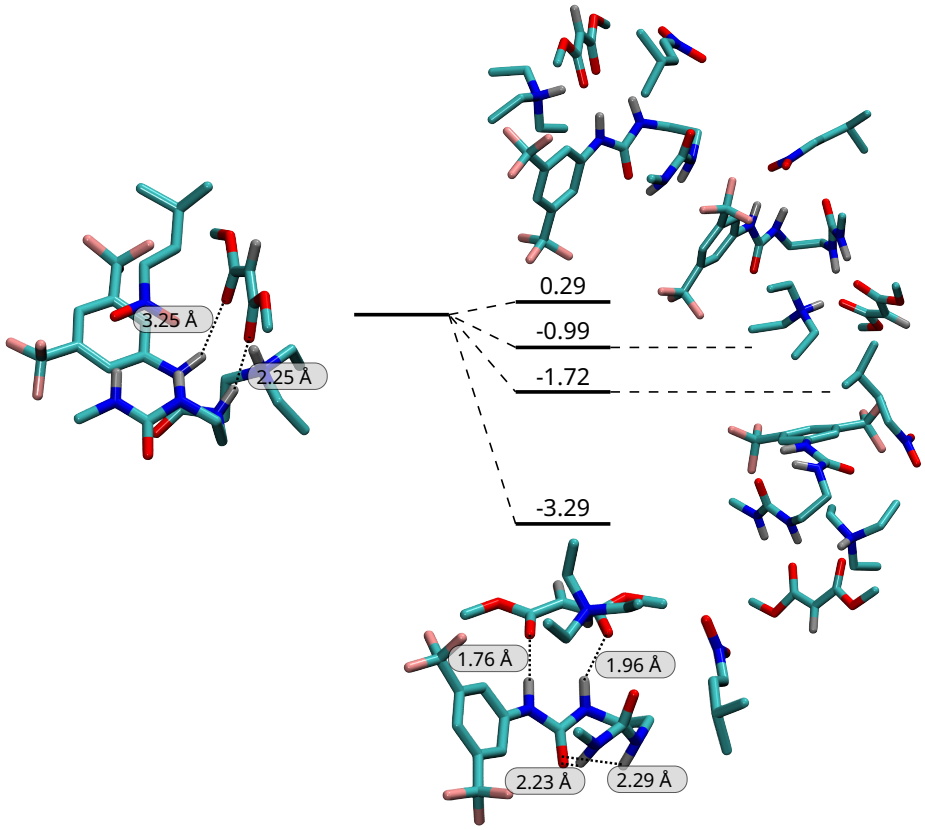


Figure B.3. Energetic difference between the most stable configurations found for structures 1 and 2 interacting with the substrates in the 1-mer. The energy values are represented in $kcal/mol$.

First, it should be noticed that the cooperation through H-Bonds between ureas 1 and 2 has been modified as a result of the substrate's arrival in structure 2, and in some cases even lost. For example, in the most stable configuration in Figure B.3, the ureas H-Bonds is established between the oxygen of the first and NH of the second, whilst in structure 2 of Figure B.2 the inverse happens. Furthermore, the internal urea-urea H-Bond strength has been decreased as a result of the substrate's presence. On the other hand, the structural changes in structure 1 are smaller when the substrates are added to the system.

In any case, a stabilization of 3.29 $kcal/mol$ is observed between structures 1 and 2. These results clearly show that the monomeric form of the foldamer is more stable when the two ureas are oriented in opposite directions, thus corroborating the experimental observations described at the beginning of this section. However, it is interesting to note that from the point of view of reactivity, the structure 1 should be better than structure 2 due to the positioning of the substrates one with respect to the other.

If any, the structure which should present more reactivity could be the top-right

geometry, where the substrates are close enough one to the other and the pathway is clean to their approach. But note that it is 3.68 $kcal/mol$ higher in energy than the most stable found.

B.3 Dimer rotameric forms

Similarly to the monomeric form of the foldamer (see section B.2 on this appendix), the 2-mer has been observed in the experiments not to adopt a folded form [242, 243]. The same analysis developed for the 1-mer is adequate for the exploration of the most stable form of the 2-mer. For that reason, a scan of the coordinate defining the helicoidalicity of the 2-mer has been carried out (see Figure B.4).

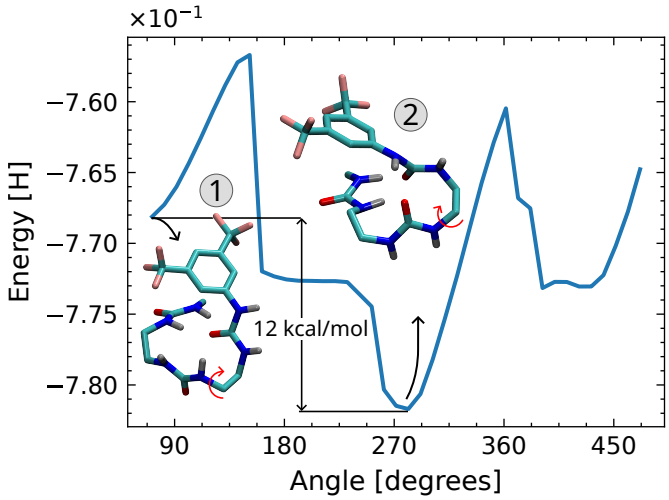


Figure B.4. Energetic profile of the coordinate describing the orientation of the first urea with respect to the second in the 2-mer. The red arrows show the rotation studied.

It is clearly observed that the structure 2 is more stable than structure 1, and it is consequence of the H-Bond interaction now established between the second urea oxygen and the first urea NH. This second configuration leads to a stabilization of 12 $kcal/mol$ with respect to the structure 1. However, this structural change could significantly influence the performance of the catalyst, as the positive accumulation of charges observed in structure 1 will be certainly lost in structure 2. Furthermore, the accommodation of the substrates could also be disfavored.

To corroborate these hypotheses, a Global Optimization with the structure 2 represented in Figure B.4 has been developed, and the obtained energy values have been then sorted with respect to the energy of structure 1 interacting with the substrates (see Figure B.5).

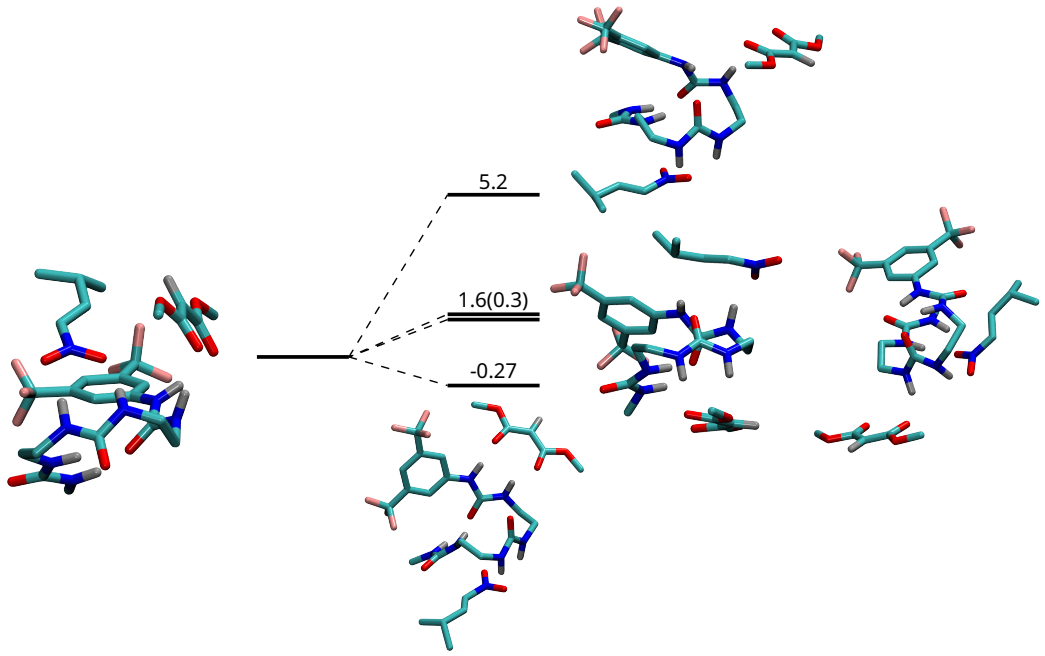


Figure B.5. Energetic difference between the most stable configurations found for structures 1 and 2 interacting with the substrates in the 2-mer. The energy values are represented in $kcal/mol$.

The stabilization produced by the use of the structure 2 in the Global Optimization is just 0.27 $kcal/mol$. This value is considerably low, and both configurations could perfectly coexist in a chemical system judging from their energies. However, it needs to be taken into account the fact that the catalyst and the substrates are separate substances which are added to form a mixture, and thus, the catalyst should be almost entirely populated by configurations like structure 2, which would inevitably lead to the most stable configuration showed in Figure B.5. Then, despite the leftmost configuration in Figure B.5 could be present in a real chemical system by its energy, it is unlikely to happen from a chemical point of view.

B.4 Foldamer rotameric forms of the phenyl group

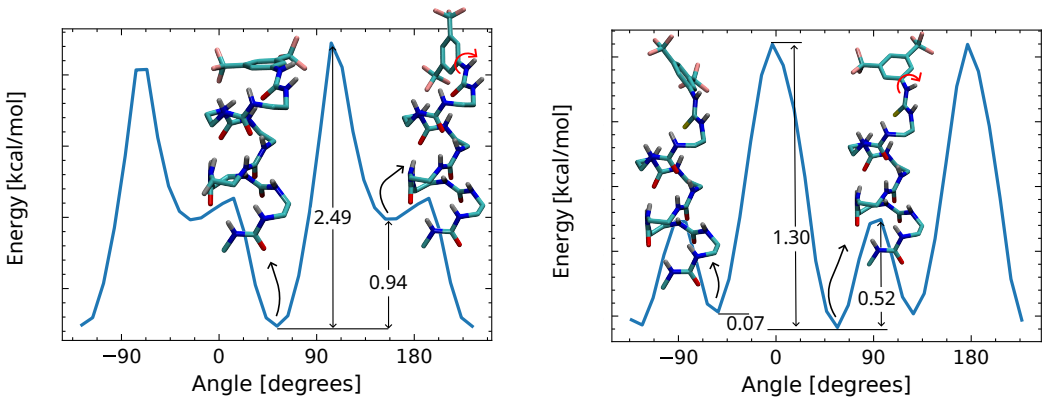
In section C.3 on page 118, a shift in activation energies was identified to correspond to the rotation of the phenyl group of the foldamers. Thus, an exploration of the most stable rotameric form of the foldamer's phenyl group is imperative to determine which is the rotamer we have to work with.

The investigation of the foldamer without the influence of any substrate was first selected (see Figure B.6 on the next page). It was found for the oligoureas that when the phenyl group is perpendicular to the foldamer's helix axis, a global minimum in energy is reached (see Figure B.6a). Furthermore, when the phenyl

group is parallel to the foldamer's helix axis, a local minimum 0.94 $kcal/mol$ higher in energy is found. The energetic barrier separating this local minimum from the global minimum is very low, but enough to avoid the evolution from one rotamer to the other.

A further investigation on a modified foldamer was also carried out, and it is represented in Figure B.6b. The modification introduced to the foldamer was the substitution of the first urea's Oxygen by a Sulfur atom (oligothiourea). In this case, the energy difference between the global minimum (phenyl group perpendicular to the foldamer's helix axis) and the local minimum (phenyl group parallel to the foldamer's helix axis) is just 0.07 $kcal/mol$, which is a negligible quantity. Thus, both rotamers will have similar energy. However, the energetic barrier to pass from one rotameric form to the other is $\sim$ 0.52 $kcal/mol$, higher than the case of the oligoureas and consequently more difficult to surpass.

It can then be concluded that in both the oligourea and the oligothiourea, the rotameric form where the phenyl group is perpendicular to the foldamer's helix axis should be used for the computations involving this molecule.



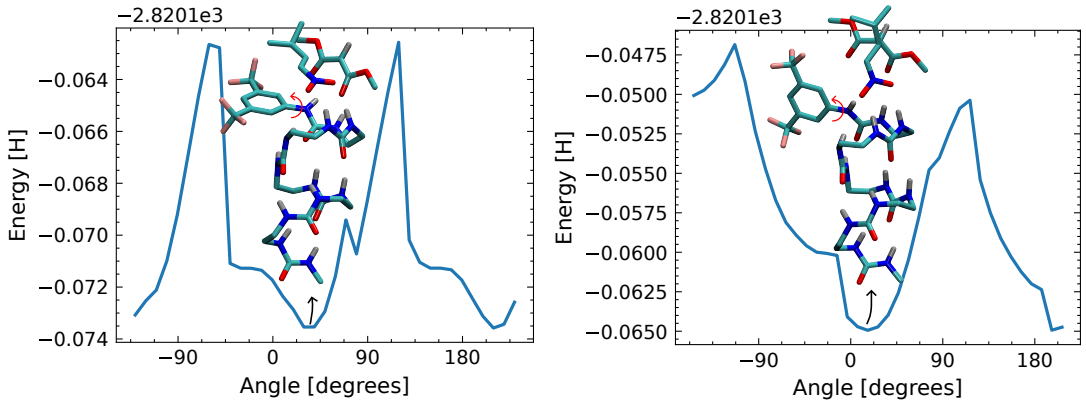
(a) Oligourea (b) Oligothiourea

Figure B.6. Rotameric analysis of the phenyl group in the oligo(thio)ureas to determine the most stable rotamer. The heptaurea has been taken as model system.

However, these circumstances could drastically change with the addition of the substrates. Furthermore, the global minimum may be different in the reactants and the transition states configurations. For that reason, a rotameric analysis having those situations into account has been carried out with the pentameric form of the oligourea, and the results are presented in Figure B.7.

In the case of the reactants (Figure B.7a), the addition of the substrates did not change the global minimum configuration, and the most stable rotameric form of the phenyl group remains that where it is perpendicular to the foldamer's helix axis. On

the other hand, it is interesting to note that for the transition states configuration the global minimum is reached when the phenyl group is parallel to the foldamer's helix axis (see Figure B.7b).



(a) Reactants (b) Transition state

Figure B.7. Rotameric analysis of the hexaurea reactants and transition states to determine the most stable rotameric positioning of the phenyl group.

B.5 Conclusions

From the results presented in sections B.2 on page 108 and B.3 on page 110, which are related to the exploration of the helicoidalicity of the di- and tri-urea, it can be concluded that:

1. The minimum energy is reached when the second urea NH are oriented in opposite direction to the NH of the first urea.

2. The opposite orientation of the first two ureas makes possible in the di- and tri-urea foldamers the establishment of two H-Bonds between ureas one and two which would otherwise not exist, thus decreasing the energy of the system. If this were not the case, the urea two would not cooperate to the H-Bonds network established internally in these oligomers.

3. The inversion of the second urea is a behavior which will certainly disappear in the tetraurea and larger foldamers, because 4 ureas is the minimum amount required to have them all interconnected by H-Bonds, and allow the connection of the second urea to the internal cooperation network without breaking the helicoidalicity of the foldamers.

4. From a reactivity point of view, the configuration with the first two ureas NH towards the same direction in the di- and tri-urea should be superior, as the reactants are better positioned to facilitate the reaction pathway.

These results can then be linked to the observations of unfoldability made in the experiment for the 1-mer and the 2-mer. In addition, the low values of reactivity measured in these short-length foldamers could be easily correlated to the positioning of the substrates one with respect to the other in both situations: (i) when the first two ureas NH are oriented towards the same direction, an easier pathway for the reaction is observed, an thus higher reactivity could be expected; and (ii), when the first two ureas NH are oriented in opposite directions, a more complex pathway is observed.

On the other hand, from the results obtained in section B.4 on page 111, it can be concluded that:

1. The most stable rotamer of the phenyl group in the foldamers is when it is perpendicular to the foldamer's helix axis.

2. When the substrates are taken into account, the rotamer with the phenyl group perpendicular to the foldamer's helix axis remains the global minimum.

3. Differently, in the transition state configuration, the rotamer where the phenyl group is parallel to the foldamer's helix axis becomes the most stable configuration.

Appendix C

Search for an efficient computational method

C.1 Introduction

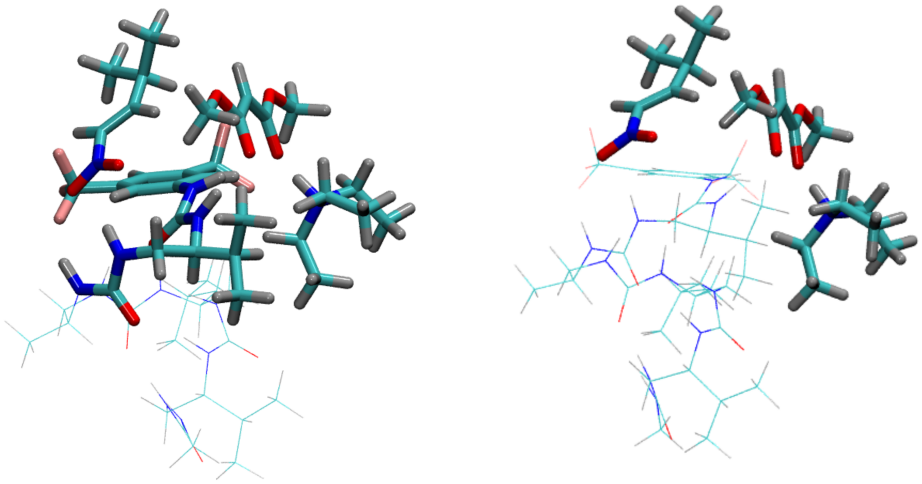
Computationally finding a transition state in relative complex systems can be a very challenging task. Additionally, to successfully describe the energetic profiles and its dependence with the number of residues of foldamers like **1** and **3a** (see Figures 2.2 and 4.2, respectively), several optimizations, reaction coordinates scans and transition states searched need to be carried out. A quick analysis of these numbers prompted us to undertake a search for methods that were capable to describe the systems under analysis with high accuracy, and at the same time less demanding computational resources.

A series of methodologies were compared with the so far used[1] in the Molecular Recognition study developed in Chapter 2. In this comparison, Molecular Mechanics (MM) and Semi-empirical (SE) methods were combined with DFT (QM). Two different QM regions, and consequently two different MM (or SE) regions, were envisaged, one which included the substrates and the catalyst's reactive region, and the other which only had the substrates in the QM part. To structurally represent both cases, the QM region has been highlighted in figure C.1 with tubes, while the rest of the foldamer was computed with the SE or MM method (drawn with lines), as specified.

The Gaussian 16 Rev. C.01 [141] software was employed for the computations, consequently with the line followed throughout all the investigations presented in this work. First, the PM6 [72] and PM7 [245] SE methods were used with the intention

[1]The methodology used in the Molecular Recognition study is DFT [65–68] with B3LYP-D3 [84–86, 99–101] and the 6-31"+"G(d,p) basis set [244]. See section 2.2 on page 42.

to test the QM/SE results accuracy, while the Universal Force Field (UFF) [63] was used to test the QM/MM method. The AMBER Force Field [58] was also used here but, as it is not parameterized for this kind of systems, the obtained results (after a user-defined parametrization) were not corrects.



(a) Non-starred computations (b) Starred computations

Figure C.1. The highlighted region (represented with tubes) was computed with QM, while the region represented with lines was computed with SE or MM (as specified in each case).

C.2 Computational time and energy

Despite the time saved —as a measurement of the computational resources saved— with the application of each methodology is the main objective of this analysis, it is not sufficient to decide which one should be used in our computations. The accuracy of the results provided by the designed methodologies and models compared to those obtained by B3LYP-D3/6-31"+"G(d,p) is an even more important parameter to measure. As different methodologies are being tested here, the use of ground energy is not correct for a direct comparison. For this reason, the Binding Energy (E_B) was the selected parameter.

$$\mathrm{E_B} = \mathrm{E_{full\ system}} - \mathrm{E_{catalyst}} - \mathrm{E_{nitroalkene}} - \mathrm{E_{malonate\ and\ Et_3N}} \tag{C.1}$$

Figure C.2 shows the binding energies obtained. The non-starred QM/PM6 and QM/PM7 methodologies are the two closest curves to the full QM method computed in Guassian (the former in greater agreement than the later). The maximum deviation in energy obtained is $\sim 4\ kcal/mol$ between the QM and QM/PM7 methods in

the heptaurea, but QM/PM6 presents deviations in the order of 2 $kcal/mol$, which is totally acceptable. On the other hand, the starred version of QM/SE computations significantly deviates from the reference value, being the QM/PM7* computations extremely high in binding energy. This could be mainly related to the omission in the QM part of the stabilizing forces between the substrates and the catalysts, which play a fundamental role as demonstrated in Chapter 2.

Also note that the QM/UFF results are relatively close to the full DFT, but still differ in nearly 10 $kcal/mol$.

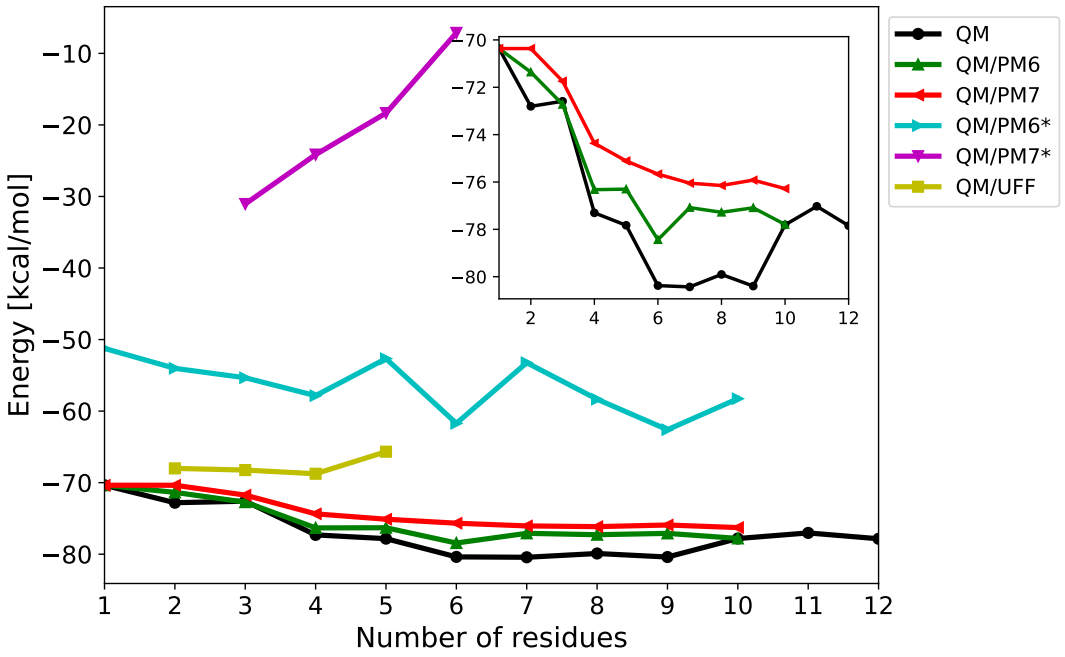


Figure C.2. Binding energy for different methodologies. The starred methods are those where only the reactants were included in the QM region (see figure C.1b).

Then, it was measured the time required for selected methodologies to find the ground state energy (see Figure C.3). The pentaurea and the hexaurea foldamers were selected as representative samples, with the conjugated addition of deprotonated dimethylmalonate to the (1E)-3-methyl-1-nitrobut-1-ene nitroolefin in presence of the protonated triethylamine as achiral base. All computations were started from the same geometry.

The QM/PM6* model data suggests that, despite it will not be theoretically speaking faster than QM/MM, the advantages of its use speeding the computations are very attractive. However, the binding energy values were too deviated from the full DFT computations, making its use energetically unjustified. Nevertheless, note that the non-starred QM/PM6 is not the most efficient method, but when it was used the computing time descended from 5 days to just 2 in the case of the

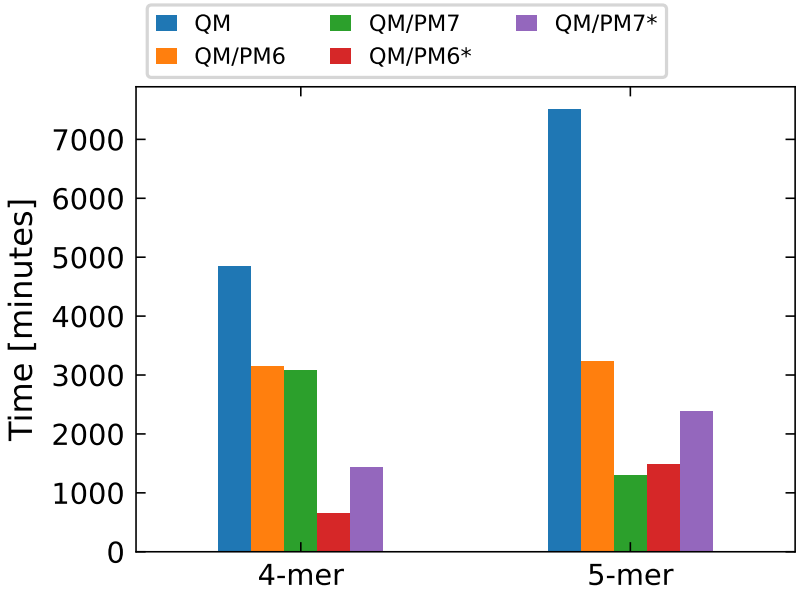


Figure C.3. Time required for different methodologies to optimize the conjugated addition of dimethylmalonate to (1E)-3-methyl-1-nitrobut-1-ene in the presence of the penta- and hexa-ureas, with triethylamine as achiral base. For non-starred and starred computations see figures C.1a and C.1b, respectively.

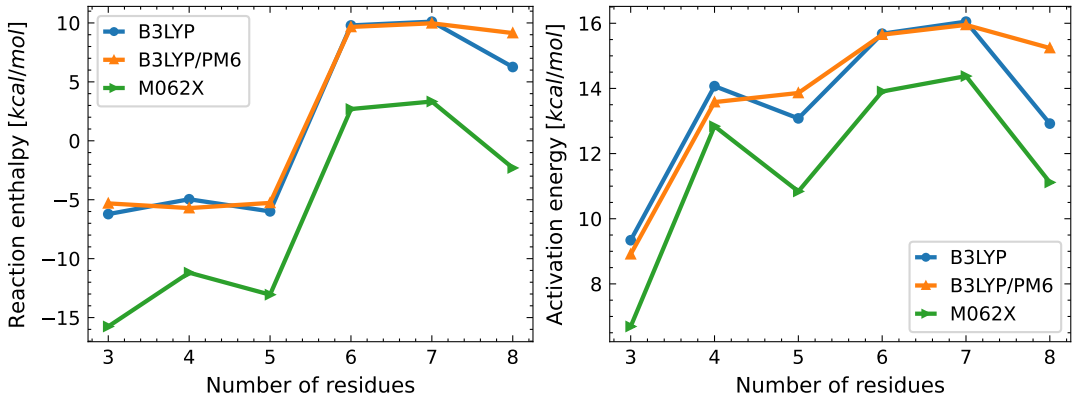
pentamer, while maintaining binding energies close to the full DFT results.

Finally, it can be strange how, for example, the QM/PM7* method required more time then its QM/PM7 counterpart, when its precisely the former which has less atoms in the QM region. This behavior is due to fact that the energy minimization with one methodology could take more optimization steps then others, which will considerably increase the computational time required to successfully complete it.

At this point, the QM/PM6 methodology seems to be the perfect choice. It provides an equilibrium between efficiency and accuracy that is reasonable to us. Despite it may be argued that QM/PM7 is also a promising candidate, we found some difficulties in the optimization process in this case that were not present on its similar.

C.3 Thermodynamic and kinetic properties

Furthermore, a study comparing the activation energy and the reaction enthalpy of the reaction when computed with B3LYP, B3LYP/PM6 and M062X functionals was carried out. Note that B3LYP is in fact what has been named QM through this appendix, and B3LYP/PM6 is equivalent to QM/PM6. The use of M062X is included here because it is a methodology that correctly describes the noncovalent interactions present in these kind of systems, due to its heavy parametrization on



(a) Reaction enthalpy (b) Activation energy

Figure C.4. Activation energy and reaction enthalpy computed with B3LYP-D3/6-31“+”G(d,p), B3LYP-D3/6-31“+”G(d,p)/PM6 and M062X/6-31“+”G(d,p) functionals. The reaction is showed in Figure 2.2 with the substrates represented in Figure 2.1.

this regard. It is important to notice that dispersion corrections are applied in all cases, to cover the DFT deficiencies on this subject.

The trimer to octamer foldamers were selected as representative to all the data. It was obtained that the deviations between QM and QM/PM6 are negligible, with only the octamer having a deviation of $\sim 2\ kcal/mol$ (see figure C.4). On the other hand, the M062X methodology is approximately 5 $kcal/mol$ lower in energy than the so far used B3LYP, but the trend is exactly the same, which will result in equal enantioselectivity of the enantiomers no matter what method is used.

As a final remark, in figure C.4a there are two different levels of energy, the first including trimer to pentamer which has negative values, and the other ranges from hexamer to octamer with positive values of enthalpy. This behavior is due to different rotameric species of the foldamer with the rotation located at the phenly group, which in one case is parallel to the helix axis, and in the other perpendicular. This is the result from a single bond rotation, and has to be taken into account in the computation of transitions states of these systems.

C.4 Conclusions

From the results presented here it is possible to conclude that:

1. The computations of the studied systems in this thesis can be done with the QM/PM6 methodology, being certain that the results will be equivalent to fully QM computed results.

2. The deviation in energy between QM/PM6 and fully QM methods are 2 $kcal/mol$, which is accurate enough for our purposes.

3. With the use of QM/PM6 method we will be saving around the 60% of computational resources and time.

4. The QM/PM6 methodology not only reproduces the QM binding energies, but also the thermodynamic and kinetic profiles.

5. The use of M062X showed lower activation energies and reaction enthalpy than B3LYP-D3, and consequently QM/PM6, but the trend was exactly the same, which is the important aspect to the analysis of enantioselectivity.

6. A conformational analysis is needed to search the stablest conformer of the foldamer's phenyl group, and work with it in the investigations.

Appendix D

Theoretical computation of pKa

D.1 Introduction

Acid dissociation constants (also known as pK_a) are essential for understanding many fundamental reactions in chemistry, because it is a physical quantity indicating protonation states of molecules in solution at a given pH. Therefore, the pKa value shows how a molecule donates or accepts the proton from a counter part molecule, and thus is a very important index for nucleophilicity in both organic chemistry and biochemistry. Following the generic reaction presented in equation (D.1)

$$\mathrm{HA_{(aq)}} \rightarrow \mathrm{A^-_{(aq)}} + \mathrm{H^+_{(aq)}} \tag{D.1}$$

the pKa can be defined as

$$pK_a = -logK_a \tag{D.2}$$

$$K_a = \frac{[\mathrm{A^-_{(aq)}}][\mathrm{H^+_{(aq)}}]}{[\mathrm{HA_{(aq)}}]} \tag{D.3}$$

Oftentimes the pK_a value for a molecule can be measured experimentally relatively easy, however there are some situation where the task gets extremely difficult. For example, (i) in molecules that have not been synthesized; (ii) in larger molecules where the local environment changes the usual pKa values, such as for certain amino acids that are part of a larger polypeptide chain; and (iii) for individual functional groups of complex molecules. For those reasons, there is great interest in using theoretical methods to calculate the pK_a values for many different types of molecules.

However, chemical accuracy in pKa calculations is difficult to achieve from a theoretical point of view, because an error of 1.36 $kcal/mol$ in the change of free energy for deprotonation in solvent results in an error of 1 pKa unit [246–248]. In

addition, the determination of the absolute aqueous solvation free energy of the proton (H^+ in equation (D.1)) cannot be directly determined from the experiment, although several approximate values has been proposed along the years (see Kelly et al. [249], for example).

To overcome the difficulties encountered in the determination of the pK_a theoretically, several methods have been developed throughout the years. For example, the thermodynamic cycles proposed by Liptak et al. [250], where pK_a value depend upon the free energies of the systems represented in equation (D.1) in different states. However, most thermodynamic cycles to compute pK_a make use of the free energy of the proton, which as previously discussed, has not exact value and thus introduces errors to the final result.

Other methods different than thermodynamic cycles have also been tested, as for example those relating pK_a with bond valence methods and bond lengths [251], pK_a correlations with highest occupied molecular orbital (HOMO) energies and frontier molecular orbitals [252], and and artificial neural networks [253] utilized to determine the pK_a. However, it seems that thermodynamic cycles, manly those not utilizing the free energy of the proton, provide more consistent results.

Nevertheless, another approach to avoid the use of the free energy of the proton in solution is to evaluate the free energy difference between HA and A^- in solution with any quantum chemical method with Polarizable Continuum Model to describe the solvent effects. Then, the linear regression between the free energy differences and the experimental pK_a is made for molecules with specific chemical groups similar to those one can be interested in obtaining their pK_a [254]. This method has shown errors no larger than 0.5 pK_a units in heterocycles [255], and amino acids and small proteins [256].

Basically, the pK_a value of the compound HA is represented by the Gibbs energy change in a deprotonation reaction in a solvent X, as defined by equation D.4

$$pK_a = -\frac{\Delta G(solv, X)}{RT\, ln(10)} = \frac{G(\mathrm{HA}, X) - G(\mathrm{A}^-, X) - G(\mathrm{H}^+, X)}{RT\, ln(10)} \tag{D.4}$$

where R is the gas constant, T the temperature, G the Gibbs free energy, and X any solvent.

To take into account systematic errors coming from the method selected for the computation, Matsui et al. introduced a scaling factor s to adjust the Gibbs energy between reactants and products. Then equation (D.4) can be re-written as

$$pK_a = -\frac{s\,[G(\mathrm{HA}, X) - G(\mathrm{A}^-, X) - G(\mathrm{H}^+, X)]}{RT\, ln(10)} \tag{D.5}$$

which can be further simplified as

$$pK_a = k\left[G(\mathrm{A}^-, X) - G(\mathrm{HA}, X)\right] + C_0 = k\Delta G_0 + C_0 \tag{D.6}$$

being

$$k = \frac{s}{RT\,ln(10)} \tag{D.7}$$

and

$$C_0 = \frac{s\,G(\mathrm{H}^+, X)}{RT\,ln(10)} \tag{D.8}$$

As previously stated, equation (D.6) should establish a linear correlation between ΔG_0 and the corresponding experimental pK_a values.

Also note how C_0/k corresponds to the Gibbs energy of the proton $G(\mathrm{H}^+, X)$, which can be used to compare the values obtained after the linear regression employing this method, to that one experimentally reported as accepted.

D.2 Computations

We first defined the set of molecules from where the experimental values of pK_a were taken (see Figure D.1). Several of these molecules, despite being structurally different than the foldamers studied throughout this manuscript, are able to successfully catalyze the same kind of reactions than the foldamers do. Furthermore, it has to be pointed out that their pK_a values were reported in the experiment with Dimethyl Sulfoxide (DMSO) as solvent. Therefore, all the computations related to the theoretical estimation of the pK_a in this manuscript will be done with DMSO as solvent, and the B3LYP-D3/6-31G(d,p) methodology.

The regression made with the experimental pK_a and the computed Gibbs free energy is represented in Figure D.2. From there, the parameters k and C_0 for equation (D.6) can be extracted, thus, the pK_a of the foldamers studied throughout the manuscript can be estimated by the following expression

$$pK_a = 0.36\ \Delta G_0 - 98.57 \tag{D.9}$$

It is interesting to take into account that the solvation Gibbs energy of the proton, which can be estimated as C_0/k has a value of

$$G(\mathrm{H}^+, X) = \frac{C_0}{k} = -272.38\ kcal/mol \tag{D.10}$$

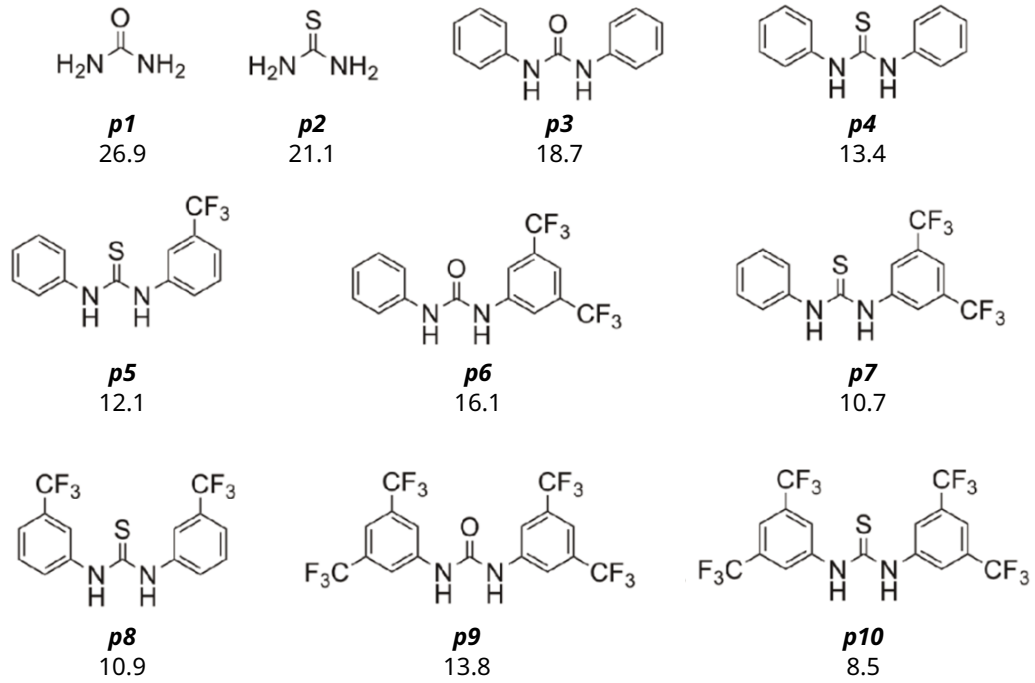


Figure D.1. Set of urea and thiourea based low mass molecules selected to compute the linear regression between their experimental pK_a values [257] and the Gibbs energy difference.

which is in good agreement with the experimentally measured and agreed of $-268.16\ kcal/mol$ in DMSO [249].

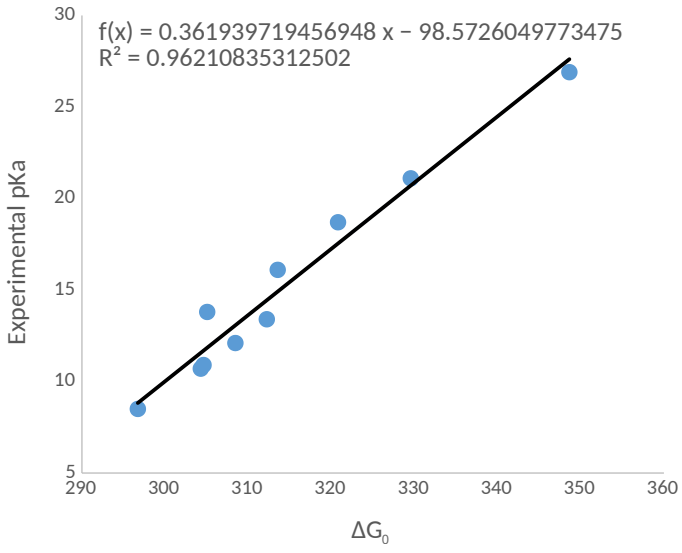


Figure D.2. Regression made to obtain the values k and C_0 in equation (D.6) from the experimental values of pK_a reported in Figure D.1, and their computed Gibbs free energy.

Finally, we applied equation (D.9) to the theoretically obtained values of ΔG_0 for the molecules **p1** to **p10**, finding that in average the deviation from the experimental

Molecule	Experimental pK_a	Estimated pK_a	$\Delta\, pK_a$
p1	26.9	27.6	0.7
p2	21.1	20.7	0.4
p3	18.7	17.5	1.2
p4	13.4	14.4	1.0
p5	12.1	13.1	1.0
p6	16.1	14.9	1.2
p7	10.7	11.6	0.9
p8	10.9	11.7	0.8
p9	13.8	11.8	2.0
p10	8.5	8.8	0.3
Average deviation = mean($\Delta\, pK_a$)		$\rightarrow$	0.9

Table D.1. Experimental and theoretically computes pK_a values following the linear regression method proposed by Matsui et al. [254].

value is around $0.9\, pK_a$ units (see Table D.1). Therefore, we conclude that the use of equation (D.9) to compute a theoretical estimation of the pK_a should report satisfactory values to compare as a function of the foldamers' number of residues.

Bibliography

[1] Victor Gold, ed. *The IUPAC Compendium of Chemical Terminology: The Gold Book*. 4th ed. Research Triangle Park, NC: International Union of Pure and Applied Chemistry (IUPAC), 2019.

[2] Ryoji Noyori. "Asymmetric Catalysis: Science and Opportunities". *Angewandte Chemie International Edition* 41.12 (June 17, 2002), pp. 2008–2022.

[3] Israel Agranat and Hava Caner. "Intellectual Property and Chirality of Drugs". *Drug Discovery Today* 4.7 (July 1, 1999), pp. 313–321.

[4] Israel Agranat and Silvya R. Wainschtein. "The Strategy of Enantiomer Patents of Drugs". *Drug Discovery Today* 15.5 (Mar. 1, 2010), pp. 163–170.

[5] Andrea Calcaterra and Ilaria D'Acquarica. "The Market of Chiral Drugs: Chiral Switches versus de Novo Enantiomerically Pure Compounds". *Journal of Pharmaceutical and Biomedical Analysis*. Review Issue 2017 147 (Jan. 5, 2018), pp. 323–340.

[6] A. Nova and F. Maseras. "Enantioselective Synthesis". *Comprehensive Inorganic Chemistry II (Second Edition)*. Ed. by Jan Reedijk and Kenneth Poeppelmeier. Amsterdam: Elsevier, Jan. 1, 2013, pp. 807–831.

[7] Maja Heitbaum, Frank Glorius, and Iris Escher. "Asymmetric Heterogeneous Catalysis". *Angewandte Chemie International Edition* 45.29 (July 17, 2006), pp. 4732–4762.

[8] Eric N. Jacobsen, ed. *Comprehensive Asymmetric Catalysis*. Berlin: Springer, 1999.

[9] Patrick J. Walsh and Marisa C. Kozlowski. *Fundamentals of Asymmetric Catalysis*. University Science Books, Jan. 2, 2009. 692 pp.

[10] Yvonne Gnas and Frank Glorius. "Chiral Auxiliaries - Principles and Recent Applications". *Synthesis* 2006.12 (June 2006), pp. 1899–1930.

[11] Kurt Faber. *Biotransformations in Organic Chemistry: A Textbook*. In collab. with SpringerLink Service en ligne. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2011.

[12] Albrecht Berkessel and Harald Gröger. *Asymmetric Organocatalysis: From Biomimetic Concepts to Applications in Asymmetric Synthesis*. Weinheim: Wiley-VCH, 2005. 440 pp.

[13] Benjamin List. "Introduction: Organocatalysis". *Chemical Reviews* 107.12 (Dec. 1, 2007), pp. 5413–5415.

[14] Joshua C. Worch, Hannah Prydderch, Sètuhn Jimaja, et al. "Stereochemical Enhancement of Polymer Properties". *Nature Reviews Chemistry* 3.9 (Sept. 2019), pp. 514–535.

[15] Axel H. E. Müller and Krzysztof Matyjaszewski, eds. *Controlled and Living Polymerizations: From Mechanisms to Applications*. 1st ed. Wiley, Aug. 19, 2009.

[16] Xia Chen, Lucia Caporaso, Luigi Cavallo, et al. "Stereoselectivity in Metallocene-Catalyzed Coordination Polymerization of Renewable Methylene Butyrolactones: From Stereo-random to Stereo-perfect Polymers". *Journal of the American Chemical Society* 134.17 (May 2, 2012), pp. 7278–7281.

[17] Deborah K. Schneiderman and Marc A. Hillmyer. "*50th Anniversary Perspective* : There Is a Great Future in Sustainable Polymers". *Macromolecules* 50.10 (May 23, 2017), pp. 3733–3749.

[18] Shady Farah, Daniel G. Anderson, and Robert Langer. "Physical and Mechanical Properties of PLA, and Their Functions in Widespread Applications — A Comprehensive Review". *Advanced Drug Delivery Reviews* 107 (Dec. 2016), pp. 367–392.

[19] L.T. Lim, R. Auras, and M. Rubino. "Processing Technologies for Poly(Lactic Acid)". *Progress in Polymer Science* 33.8 (Aug. 2008), pp. 820–852.

[20] Andrew P. Dove, Russell C. Pratt, Bas G. G. Lohmeijer, et al. "Thiourea-Based Bifunctional Organocatalysis: Supramolecular Recognition for Living Polymerization". *Journal of the American Chemical Society* 127.40 (Oct. 1, 2005), pp. 13798–13799.

[21] Russell C. Pratt, Bas G. G. Lohmeijer, David A. Long, et al. "Exploration, Optimization, and Application of Supramolecular Thiourea-Amine Catalysts for the Synthesis of Lactide (Co)Polymers". *Macromolecules* 39.23 (Nov. 1, 2006), pp. 7863–7871.

[22] Andrew P. Dove. "Organic Catalysis for Ring-Opening Polymerization". *ACS Macro Letters* 1.12 (Dec. 18, 2012), pp. 1409–1412.

[23] David J. Hill, Matthew J. Mio, Ryan B. Prince, et al. "A Field Guide to Foldamers". *Chemical Reviews* 101.12 (Dec. 1, 2001), pp. 3893–4012.

[24] Anthony J. Metrano, Alex J. Chinn, Christopher R. Shugrue, et al. "Asymmetric Catalysis Mediated by Synthetic Peptides, Version 2.0: Expansion of Scope and Mechanisms". *Chemical Reviews* 120.20 (Oct. 28, 2020), pp. 11479–11615.

[25] Bartosz Lewandowski and Helma Wennemers. "Asymmetric Catalysis with Short-Chain Peptides". *Current Opinion in Chemical Biology*. Synthetic Biology - Synthetic Biomolecules 22 (Oct. 1, 2014), pp. 40–46.

[26] Kengo Akagawa and Kazuaki Kudo. "Development of Selective Peptide Catalysts with Secondary Structural Frameworks". *Accounts of Chemical Research* 50.10 (Oct. 17, 2017), pp. 2429–2439.

[27] Elizabeth A. Colby Davie, Steven M. Mennen, Yingju Xu, et al. "Asymmetric Catalysis Mediated by Synthetic Peptides". *Chemical Reviews* 107.12 (Dec. 1, 2007), pp. 5759–5812.

[28] Vincent Diemer, Lucile Fischer, Brice Kauffmann, et al. "Anion Recognition by Aliphatic Helical Oligoureas". *Chemistry – A European Journal* 22.44 (Oct. 24, 2016), pp. 15684–15692.

[29] Diane Bécart, Vincent Diemer, Arnaud Salaün, et al. "Helical Oligourea Foldamers as Powerful Hydrogen Bonding Catalysts for Enantioselective C–C Bond-Forming Reactions". *Journal of the American Chemical Society* 139.36 (Sept. 13, 2017), pp. 12524–12532.

[30] Tomotaka Okino, Yasutaka Hoashi, and Yoshiji Takemoto. "Enantioselective Michael Reaction of Malonates to Nitroolefins Catalyzed by Bifunctional Organocatalysts". *Journal of the American Chemical Society* 125.42 (Oct. 1, 2003), pp. 12672–12673.

[31] Yoshiji Takemoto. "Recognition and Activation by Ureas and Thioureas: Stereoselective Reactions Using Ureas and Thioureas as Hydrogen-Bonding Donors". *Organic & Biomolecular Chemistry* 3.24 (Dec. 2, 2005), pp. 4299–4306.

[32] Abigail G. Doyle and Eric N. Jacobsen. "Small-Molecule H-Bond Donors in Asymmetric Catalysis". *Chemical Reviews* 107.12 (Dec. 1, 2007), pp. 5713–5743.

[33] Christopher R. Jones, G. Dan Pantoş, Angus J. Morrison, et al. "Plagiarizing Proteins: Enhancing Efficiency in Asymmetric Hydrogen-Bonding Catalysis through Positive Cooperativity". *Angewandte Chemie International Edition* 48.40 (2009), pp. 7391–7394.

[34] Ana Crespo-Peña, David Monge, Eloísa Martín-Zamora, et al. "Asymmetric Formal Carbonyl-Ene Reactions of Formaldehyde Tert-Butyl Hydrazone with α-Keto Esters: Dual Activation by Bis-urea Catalysts". *Journal of the American Chemical Society* 134.31 (Aug. 8, 2012), pp. 12912–12915.

[35] Nicolas Probst, Ádám Madarász, Arto Valkonen, et al. "Cooperative Assistance in Bifunctional Organocatalysis: Enantioselective Mannich Reactions with Aliphatic and Aromatic Imines". *Angewandte Chemie International Edition* 51.34 (2012), pp. 8495–8499.

[36] C. Rose Kennedy, Dan Lehnherr, Naomi S. Rajapaksa, et al. "Mechanism-Guided Development of a Highly Active Bis-thiourea Catalyst for Anion-Abstraction Catalysis". *Journal of the American Chemical Society* 138.41 (Oct. 19, 2016), pp. 13525–13528.

[37] Yongho Park, Kaid C. Harper, Nadine Kuhl, et al. "Macrocyclic Bis-Thioureas Catalyze Stereospecific Glycosylation Reactions". *Science* 355.6321 (Jan. 13, 2017), pp. 162–166.

[38] Gabriele Pupo, Francesco Ibba, David M. H. Ascough, et al. "Asymmetric Nucleophilic Fluorination under Hydrogen Bonding Phase-Transfer Catalysis". *Science* 360.6389 (May 11, 2018), pp. 638–642.

[39] Francesco Ibba, Gabriele Pupo, Amber L. Thompson, et al. "Impact of Multiple Hydrogen Bonds with Fluoride on Catalysis: Insight from NMR Spectroscopy". *Journal of the American Chemical Society* 142.46 (Nov. 18, 2020), pp. 19731–19744.

[40] Alexander S. Kucherenko, Alexey A. Kostenko, Andrey N. Komogortsev, et al. "C $_2$ -Symmetric Chiral Squaramide, Recyclable Organocatalyst for Asymmetric Michael Reactions". *The Journal of Organic Chemistry* 84.7 (Apr. 5, 2019), pp. 4304–4311.

[41] Kenya Tamaribuchi, Jiaqi Tian, Kengo Akagawa, et al. "Enantioselective Nitro-Michael Addition Catalyzed by N-Terminal Guanidinylated Helical Peptide". *Advanced Synthesis & Catalysis* 364.1 (Jan. 4, 2022), pp. 82–86.

[42] Kazuki Sato, Tomohiro Umeno, Atsushi Ueda, et al. “Asymmetric 1,4-Addition Reactions Catalyzed by N-Terminal Thiourea-Modified Helical L -Leu Peptide with Cyclic Amino Acids”. *Chemistry – A European Journal* 27.43 (Aug. 2, 2021), pp. 11216–11220.

[43] Carlos Vila, Francisco Cernicharo-Toledo, Gonzalo Blay, et al. “Nitroenynes as Electrophiles in Organocatalysis and Their Application in the Synthesis of Chiral Heterocycles”. *European Journal of Organic Chemistry* 2021.16 (Apr. 26, 2021), pp. 2255–2267.

[44] Mégane Bornerie, Anaïs Brion, Gilles Guichard, et al. “Delivery of siRNA by Tailored Cell-Penetrating Urea-Based Foldamers”. *Chemical Communications* 57.12 (2021), pp. 1458–1461.

[45] M.Umair Tariq and Wesley J. Moran. “Design and Synthesis of Chiral Urea-Derived Iodoarenes and Their Assessment in the Enantioselective Dearomatizing Cyclization of a Naphthyl Amide”. *Tetrahedron* 76.45 (Nov. 2020), p. 131634.

[46] Zhe An, Yuanzhong Tang, Yitao Jiang, et al. “Enhanced Enantioselectivity in Heterogeneous Manganese-Catalyzed Asymmetric Epoxidation with Nanosheets Modified Amino Acid Schiff Bases as Ligands by Modulating the Orientation and the Arrangement Order”. *Journal of Catalysis* 402 (Oct. 2021), pp. 22–36.

[47] K. N. Houk and Paul Ha-Yeon Cheong. “Computational Prediction of Small-Molecule Catalysts”. *Nature* 455.7211 (7211 Sept. 2008), pp. 309–313.

[48] S. Bahmanyar and K. N. Houk. “Transition States of Amine-Catalyzed Aldol Reactions Involving Enamine Intermediates: Theoretical Studies of Mechanism, Reactivity, and Stereoselectivity”. *Journal of the American Chemical Society* 123.45 (Nov. 1, 2001), pp. 11273–11283.

[49] Kathryn N. Rankin, James W. Gauld, and Russell J. Boyd. “Density Functional Study of the Proline-Catalyzed Direct Aldol Reaction”. *The Journal of Physical Chemistry A* 106.20 (May 1, 2002), pp. 5155–5159.

[50] Fernando R. Clemente and K. N. Houk. “Computational Evidence for the Enamine Mechanism of Intramolecular Aldol Reactions Catalyzed by Proline”. *Angewandte Chemie International Edition* 43.43 (2004), pp. 5766–5768.

[51] Susumu Mitsumori, Haile Zhang, Paul Ha-Yeon Cheong, et al. “Direct Asymmetric Anti-Mannich-Type Reactions Catalyzed by a Designed Amino Acid”. *Journal of the American Chemical Society* 128.4 (Feb. 1, 2006), pp. 1040–1041.

[52] Ming Wen, Shihai Li, Jianjun Han, et al. "Mechanism and Enantioselectivity of the Asymmetric [3+2]-Annulation between N-methylindole and Enoldiazoacetamide Catalyzed by Prolinate-Coordinated Dirhodium: A Theoretical Study". *Journal of Molecular Graphics and Modelling* 94 (Jan. 2020), p. 107489.

[53] Jin Lin, Fang Wang, Jing Tian, et al. "Theoretical and Experimental Investigations of the Enantioselective Epoxidation of Olefins Catalyzed by Manganese Complexes". *Chinese Chemical Letters* (Aug. 2021), S1001841721006665.

[54] Pavel A. Dub, Nikolay V. Tkachenko, Vijyesh K. Vyas, et al. "Enantioselectivity in the Noyori–Ikariya Asymmetric Transfer Hydrogenation of Ketones". *Organometallics* 40.9 (May 10, 2021), pp. 1402–1410.

[55] Min-Can Wang, Zhi-kang Liu, Song Li, et al. "An Experimental and Theoretical Study on Free Ligand Conformational Preferences and Enantioselectivity Relationship for the Asymmetric Addition of Diethylzinc to Benzaldehyde". *Tetrahedron: Asymmetry* 21.4 (Mar. 2010), pp. 486–493.

[56] Ira N. Levine. *Quantum Chemistry*. 7th ed. 1 vols. United States of America: Pearson, 2014. 700 pp.

[57] M. Born and R. Oppenheimer. "Zur Quantentheorie Der Molekeln". *Annalen der Physik* 389.20 (1927), pp. 457–484.

[58] Wendy D. Cornell, Piotr Cieplak, Christopher I. Bayly, et al. "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules". *Journal of the American Chemical Society* 117.19 (May 1995), pp. 5179–5197.

[59] A. D. MacKerell, D. Bashford, M. Bellott, et al. "All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins". *The Journal of Physical Chemistry B* 102.18 (Apr. 1, 1998), pp. 3586–3616.

[60] K. Vanommeslaeghe and A.D. MacKerell. "CHARMM Additive and Polarizable Force Fields for Biophysics and Computer-Aided Drug Design". *Biochimica et Biophysica Acta (BBA) - General Subjects* 1850.5 (May 2015), pp. 861–871.

[61] William L. Jorgensen and Julian Tirado-Rives. "The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin". *Journal of the American Chemical Society* 110.6 (Mar. 1, 1988), pp. 1657–1666.

[62] William L. Jorgensen, David S. Maxwell, and Julian Tirado-Rives. "Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids". *Journal of the American Chemical Society* 118.45 (Nov. 13, 1996), pp. 11225–11236.

[63] A. K. Rappe, C. J. Casewit, K. S. Colwell, et al. "UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations". *Journal of the American Chemical Society* 114.25 (Dec. 1, 1992), pp. 10024–10035.

[64] Raju P. Gupta. "Lattice Relaxation at a Metal Surface". *Physical Review B* 23.12 (June 15, 1981), pp. 6265–6270.

[65] P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas". *Physical Review* 136 (3B Nov. 9, 1964), B864–B871.

[66] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects". *Physical Review* 140 (4A Nov. 15, 1965), A1133–A1138.

[67] Dennis R. Salahub and Michael C. Zerner, eds. *The Challenge of d and f Electrons: Theory and Computation*. Vol. 394. ACS Symposium Series. Washington, DC: American Chemical Society, June 8, 1989.

[68] W Yang and Robert G Parr. "Density Functional Theory of Atoms and Molecules". *Oxford University Press* 1 (1989), p. 989.

[69] E Hückel. "Quantum Contributions to the Benzene Problem". *Z Phys* 70 (1931), pp. 204–286.

[70] Michael J. S. Dewar, Eve G. Zoebisch, Eamonn F. Healy, et al. "Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model". *Journal of the American Chemical Society* 107.13 (June 1985), pp. 3902–3909.

[71] M Scholten. PhD thesis. Germany: Universitat Dusseldorf, 2003.

[72] James J. P. Stewart. "Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements". *Journal of Molecular Modeling* 13.12 (Dec. 2007), pp. 1173–1213.

[73] James J. P. Stewart. "Optimization of Parameters for Semiempirical Methods VI: More Modifications to the NDDO Approximations and Re-Optimization of Parameters". *Journal of Molecular Modeling* 19.1 (Jan. 2013), pp. 1–32.

[74] Michael J. S. Dewar and Walter Thiel. "Ground States of Molecules. 38. The MNDO Method. Approximations and Parameters". *Journal of the American Chemical Society* 99.15 (June 1977), pp. 4899–4907.

[75] Michael J. S. Dewar and Walter Thiel. "Ground States of Molecules. 39. MNDO Results for Molecules Containing Hydrogen, Carbon, Nitrogen, and Oxygen". *Journal of the American Chemical Society* 99.15 (June 1977), pp. 4907–4917.

[76] Tell Tuttle and Walter Thiel. "OMx-D: Semiempirical Methods with Orthogonalization and Dispersion Corrections. Implementation and Biochemical Application". *Physical Chemistry Chemical Physics* 10.16 (Apr. 10, 2008), pp. 2159–2166.

[77] D. R. Hartree. "The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion". *Mathematical Proceedings of the Cambridge Philosophical Society* 24.1 (Jan. 1928), pp. 111–132.

[78] J. C. Slater. "Note on Hartree's Method". *Physical Review* 35.2 (Jan. 15, 1930), pp. 210–211.

[79] Lucjan Piela. *Ideas of Quantum Chemistry*. Second Edition. 1 vols. Elsevier, 2014. 1037 pp.

[80] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. "Generalized Gradient Approximation Made Simple". *Physical Review Letters* 77.18 (Oct. 28, 1996), pp. 3865–3868.

[81] Erin R. Johnson, Shahar Keinan, Paula Mori-Sánchez, et al. "Revealing Noncovalent Interactions". *Journal of the American Chemical Society* 132.18 (May 12, 2010), pp. 6498–6506.

[82] Holly J. Davis and Robert J. Phipps. "Harnessing Non-Covalent Interactions to Exert Control over Regioselectivity and Site-Selectivity in Catalytic Reactions". *Chemical Science* 8.2 (2017), pp. 864–877.

[83] Axel D. Becke. "A New Mixing of Hartree–Fock and Local Density-functional Theories". *The Journal of Chemical Physics* 98.2 (Jan. 15, 1993), pp. 1372–1377.

[84] Axel D. Becke. "Density-functional Thermochemistry. I. The Effect of the Exchange-only Gradient Correction". *The Journal of Chemical Physics* 96.3 (Feb. 1992), pp. 2155–2160.

[85] Axel D. Becke. “Density-functional Thermochemistry. II. The Effect of the Perdew–Wang Generalized-gradient Correlation Correction”. *The Journal of Chemical Physics* 97.12 (Dec. 15, 1992), pp. 9173–9177.

[86] Axel D. Becke. “Density-functional Thermochemistry. III. The Role of Exact Exchange”. *The Journal of Chemical Physics* 98.7 (Apr. 1, 1993), pp. 5648–5652.

[87] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, et al. “Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields”. *The Journal of Physical Chemistry* 98.45 (Nov. 1, 1994), pp. 11623–11627.

[88] A. D. Becke. “Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior”. *Physical Review A* 38.6 (Sept. 1, 1988), pp. 3098–3100.

[89] Chengteh Lee, Weitao Yang, and Robert G. Parr. “Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density”. *Physical Review B* 37.2 (Jan. 15, 1988), pp. 785–789.

[90] Burkhard Miehlich, Andreas Savin, Hermann Stoll, et al. “Results Obtained with the Correlation Energy Density Functionals of Becke and Lee, Yang and Parr”. *Chemical Physics Letters* 157.3 (May 1989), pp. 200–206.

[91] S. H. Vosko, L. Wilk, and M. Nusair. “Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis”. *Canadian Journal of Physics* 58.8 (Aug. 1, 1980), pp. 1200–1211.

[92] Yan Zhao and Donald G. Truhlar. “The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-class Functionals and 12 Other Functionals”. *Theoretical Chemistry Accounts* 120.1 (May 1, 2008), pp. 215–241.

[93] B. Liu and A. D. McLean. “Accurate Calculation of the Attractive Interaction of Two Ground State Helium Atoms”. *The Journal of Chemical Physics* 59.8 (Oct. 15, 1973), pp. 4557–4558.

[94] S.F. Boys and F. Bernardi. “The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors”. *Molecular Physics* 19.4 (Oct. 1970), pp. 553–566.

[95] Abby L. Parrill and Kenny B. Lipkowitz. *Reviews in Computational Chemistry, Volume 29*. John Wiley & Sons, Apr. 11, 2016. 486 pp.

[96] Martin Stöhr, Troy Van Voorhis, and Alexandre Tkatchenko. "Theory and Practice of Modeling van Der Waals Interactions in Electronic-Structure Calculations". *Chemical Society Reviews* 48.15 (2019), pp. 4118–4154.

[97] Jianwei Sun, Richard C. Remsing, Yubo Zhang, et al. "Accurate First-Principles Structures and Energies of Diversely Bonded Systems from an Efficient Density Functional". *Nature Chemistry* 8.9 (Sept. 2016), pp. 831–836.

[98] Clemence Corminboeuf. "Minimizing Density Functional Failures for Non-Covalent Interactions Beyond van Der Waals Complexes". *Accounts of Chemical Research* 47.11 (Nov. 18, 2014), pp. 3217–3224.

[99] Stefan Grimme. "Semiempirical GGA-type density functional constructed with a long-range dispersion correction". *Journal of Computational Chemistry* 27.15 (2006), pp. 1787–1799.

[100] Stefan Grimme, Jens Antony, Stephan Ehrlich, et al. "A Consistent and Accurate *Ab Initio* Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu". *The Journal of Chemical Physics* 132.15 (Apr. 21, 2010), p. 154104.

[101] Stefan Grimme, Stephan Ehrlich, and Lars Goerigk. "Effect of the Damping Function in Dispersion Corrected Density Functional Theory". *Journal of Computational Chemistry* 32.7 (May 2011), pp. 1456–1465.

[102] Axel D. Becke and Erin R. Johnson. "Exchange-Hole Dipole Moment and the Dispersion Interaction". *The Journal of Chemical Physics* 122.15 (Apr. 15, 2005), p. 154104.

[103] Axel D. Becke and Erin R. Johnson. "A Density-Functional Model of the Dispersion Interaction". *The Journal of Chemical Physics* 123.15 (Oct. 15, 2005), p. 154101.

[104] Erin R. Johnson and Axel D. Becke. "A Post-Hartree-Fock Model of Intermolecular Interactions: Inclusion of Higher-Order Corrections". *The Journal of Chemical Physics* 124.17 (May 7, 2006), p. 174104.

[105] Erich Runge and E. K. U. Gross. "Density-Functional Theory for Time-Dependent Systems". *Physical Review Letters* 52.12 (Mar. 19, 1984), pp. 997–1000.

[106] Florian Weigend and Reinhart Ahlrichs. "Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy". *Physical Chemistry Chemical Physics* 7.18 (2005), p. 3297.

[107] Florian Weigend. "Accurate Coulomb-fitting Basis Sets for H to Rn". *Physical Chemistry Chemical Physics* 8.9 (2006), p. 1057.

[108] Austin Pounder, William Tam, and Leanne D. Chen. "The Mechanism and Origin of Enantioselectivity in the Rhodium-Catalyzed Asymmetric Ring-Opening Reactions of Oxabicyclic Alkenes with Organoboronic Acids: A DFT Investigation". *Organometallics* 40.11 (June 14, 2021), pp. 1588–1597.

[109] Jakob Wåhlander, Mohamed Amedjkouh, and David Balcells. "A DFT Perspective on Diels-Alder Organocatalysts Based on Substituted Phosphoramides: A DFT Perspective on Diels-Alder Organocatalysts Based on Substituted Phosphoramides". *European Journal of Organic Chemistry* 2019.2-3 (Jan. 23, 2019), pp. 442–450.

[110] Naoki Sakai, Kyohei Kawashima, Masashi Kajitani, et al. "Combined Computational and Experimental Studies on the Asymmetric Michael Addition of α-Aminomaleimides to β-Nitrostyrenes Using an Organocatalyst Derived from *Cinchona* Alkaloid". *Organic Letters* 23.15 (Aug. 6, 2021), pp. 5714–5718.

[111] Jun-Ling Zhu, Yong Zhang, Chong Liu, et al. "Insights into the Dual Activation Mechanism Involving Bifunctional Cinchona Alkaloid Thiourea Organocatalysts: An NMR and DFT Study". *The Journal of Organic Chemistry* 77.21 (Nov. 2, 2012), pp. 9813–9825.

[112] Shengwei Wei, Denis A. Yalalov, Svetlana B. Tsogoeva, et al. "New Highly Enantioselective Thiourea-Based Bifunctional Organocatalysts for Nitro-Michael Addition Reactions". *Catalysis Today* 121.1-2 (Mar. 2007), pp. 151–157.

[113] Rossana Fanelli, Dénes Berta, Tamás Földes, et al. "Organocatalytic Access to a *Cis* -Cyclopentyl-γ-amino Acid: An Intriguing Model of Selectivity and Formation of a Stable 10/12-Helix from the Corresponding γ/α-Peptide". *Journal of the American Chemical Society* 142.3 (Jan. 22, 2020), pp. 1382–1393.

[114] Josué Vazquez-Chavez, Socorro Luna-Morales, Diego A. Cruz-Aguilar, et al. "The Effect of Chiral *N* -Substituents with Methyl or Trifluoromethyl Groups on the Catalytic Performance of Mono- and Bifunctional Thioureas". *Organic & Biomolecular Chemistry* 17.47 (2019), pp. 10045–10051.

[115] Andrea Hamza, Gábor Schubert, Tibor Soós, et al. "Theoretical Studies on the Bifunctionality of Chiral Thiourea-Based Organocatalysts: Competing Routes to C-C Bond Formation". *Journal of the American Chemical Society* 128.40 (Oct. 1, 2006), pp. 13151–13160.

[116] Joseph A. Izzo, Yaroslaw Myshchuk, Jennifer S. Hirschi, et al. "Transition State Analysis of an Enantioselective Michael Addition by a Bifunctional Thiourea Organocatalyst". *Organic & Biomolecular Chemistry* 17.16 (2019), pp. 3934–3939.

[117] Prapussorn Yingcharoen, Wuttichai Natongchai, Albert Poater, et al. "Intertwined Chemistry of Hydroxyl Hydrogen-Bond Donors, Epoxides and Isocyanates in the Organocatalytic Synthesis of Oxazolidinones *versus* Isocyanurates: Rational Catalytic Investigation and Mechanistic Understanding". *Catalysis Science & Technology* 10.16 (2020), pp. 5544–5558.

[118] Ping Wang, Yun Gao, Yang Zhao, et al. "Unveiling Mechanism of a Quinine-Squaramide Catalyzed Enantioselective Aza-Friedel–Crafts Reaction between Cyclic Trifluoromethyl Ketimine and Naphthol: A DFT Study". *The Journal of Organic Chemistry* 82.24 (Dec. 15, 2017), pp. 13109–13114.

[119] Lori A. Burns, Álvaro Vázquez- Mayagoitia, Bobby G. Sumpter, et al. "Density-Functional Approaches to Noncovalent Interactions: A Comparison of Dispersion Corrections (DFT-D), Exchange-Hole Dipole Moment (XDM) Theory, and Specialized Functionals". *The Journal of Chemical Physics* 134.8 (Feb. 28, 2011), p. 084107.

[120] Jonathon Witte, Jeffrey B. Neaton, and Martin Head-Gordon. "Push It to the Limit: Characterizing the Convergence of Common Sequences of Basis Sets for Intermolecular Interactions as Described by Density Functional Theory". *The Journal of Chemical Physics* 144.19 (May 21, 2016), p. 194306.

[121] Lars Onsager. "Electric Moments of Molecules in Liquids". *Journal of the American Chemical Society* 58.8 (Aug. 1936), pp. 1486–1493.

[122] Jacopo Tomasi, Benedetta Mennucci, and Roberto Cammi. "Quantum Mechanical Continuum Solvation Models". *Chemical Reviews* 105.8 (Aug. 1, 2005), pp. 2999–3094.

[123] S. Miertus, E. Scrocco, and J. Tomasi. "Electrostatic Interaction of a Solute with a Continuum. A Direct Utilizaion of AB Initio Molecular Potentials for the Prevision of Solvent Effects". *Chemical Physics* 55.1 (Feb. 1981), pp. 117–129.

[124] S. Miertus and J. Tomasi. “Approximate Evaluations of the Electrostatic Free Energy and Internal Energy Changes in Solution Processes”. *Chemical Physics* 65.2 (Mar. 1982), pp. 239–245.

[125] J. L. Pascual-ahuir, E. Silla, and I. Tuñon. “GEPOL: An Improved Description of Molecular Surfaces. III. A New Algorithm for the Computation of a Solvent-Excluding Surface: GEPOL”. *Journal of Computational Chemistry* 15.10 (Oct. 1994), pp. 1127–1138.

[126] Ming Wah Wong, Michael J. Frisch, and Kenneth B. Wiberg. “Solvent Effects. 1. The Mediation of Electrostatic Effects by Solvents”. *Journal of the American Chemical Society* 113.13 (June 1991), pp. 4776–4782.

[127] Ming Wah Wong, Kenneth B. Wiberg, and Michael Frisch. “Hartree–Fock Second Derivatives and Electric Field Properties in a Solvent Reaction Field: Theory and Application”. *The Journal of Chemical Physics* 95.12 (Dec. 15, 1991), pp. 8991–8998.

[128] Ming Wah Wong, Kenneth B. Wiberg, and Michael J. Frisch. “Solvent Effects. 2. Medium Effect on the Structure, Energy, Charge Density, and Vibrational Frequencies of Sulfamic Acid”. *Journal of the American Chemical Society* 114.2 (Jan. 1992), pp. 523–529.

[129] Ming Wah Wong, Kenneth B. Wiberg, and Michael J. Frisch. “Solvent Effects. 3. Tautomeric Equilibria of Formamide and 2-Pyridone in the Gas Phase and Solution: An Ab Initio SCRF Study”. *Journal of the American Chemical Society* 114.5 (Feb. 1992), pp. 1645–1652.

[130] John G. Kirkwood. “Theory of Solutions of Molecules Containing Widely Separated Charges with Special Application to Zwitterions”. *The Journal of Chemical Physics* 2.7 (July 1934), pp. 351–361.

[131] Aleksandr V. Marenich, Christopher J. Cramer, and Donald 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”. *The Journal of Physical Chemistry B* 113.18 (May 7, 2009), pp. 6378–6396.

[132] Roy L. Johnston. *Atomic and Molecular Clusters*. CRC Press, Apr. 25, 2002. 258 pp.

[133] Thomas Fehlner, Jean-Francois Halet, and Jean-Yves Saillard. *Molecular Clusters: A Bridge to Solid-State Chemistry*. Cambridge University Press, July 5, 2007. 388 pp.

[134] Bernd Hartke. "Global Optimization". *WIREs Computational Molecular Science* 1.6 (2011), pp. 879–887.

[135] Jonathan P. K. Doye and David J. Wales. "Structural Consequences of the Range of the Interatomic Potential A Menagerie of Clusters". *Journal of the Chemical Society, Faraday Transactions* 93.24 (1997), pp. 4233–4243.

[136] Roy L. Johnston. "Evolving Better Nanoparticles: Genetic Algorithms for Optimising Cluster Geometries". *Dalton Transactions* 22 (2003), p. 4193.

[137] Sven Heiles and Roy L. Johnston. "Global Optimization of Clusters Using Electronic Structure Methods". *International Journal of Quantum Chemistry* 113.18 (Sept. 15, 2013), pp. 2091–2109.

[138] Rémi Marchal, Philippe Carbonnière, and Claude Pouchan. "A Global Search Algorithm of Minima Exploration for the Investigation of Low Lying Isomers of Clusters from Density Functional Theory-Based Potential Energy Surfaces: The Example of Sin (N=3,15) as a Test Case". *The Journal of Chemical Physics* 131.11 (Sept. 21, 2009), p. 114105.

[139] H. Bernhard Schlegel. "Optimization of Equilibrium Geometries and Transition Structures". *Journal of Computational Chemistry* 3.2 (Sum. 1982), pp. 214–218.

[140] Xiaosong Li and Michael J. Frisch. "Energy-Represented Direct Inversion in the Iterative Subspace within a Hybrid Geometry Optimization Method". *Journal of Chemical Theory and Computation* 2.3 (May 9, 2006), pp. 835–839.

[141] M. J. Frisch, G. W. Trucks, H. B. Schlegel, et al. *Gaussian 16 Rev. C.01.* Wallingford, CT, 2016.

[142] George Maroulis. "On the Accurate Theoretical Determination of the Static Hyperpolarizability of *Trans* -Butadiene". *The Journal of Chemical Physics* 111.2 (July 8, 1999), pp. 583–591.

[143] Philippe Carbonniere, Sandrine Thicoipe, and Claude Pouchan. "Theoretical Strategy to Build Structural Models of Microhydrated Inorganic Systems for the Knowledge of Their Vibrational Properties: The Case of the Hydrated Nitrate Aerosols". *The Journal of Physical Chemistry A* 117.18 (May 9, 2013), pp. 3826–3834.

[144] Sandrine Thicoipe, Philippe Carbonniere, and Claude Pouchan. "The Use of the GSAM Approach for the Structural Investigation of Low-Lying Isomers of Molecular Clusters from Density-Functional-Theory-Based Potential Energy Surfaces: The Structures of Microhydrated Nucleic Acid Bases". *The Journal of Physical Chemistry A* 117.32 (Aug. 15, 2013), pp. 7236–7245.

[145] Sandrine Thicoipe, Philippe Carbonniere, and Claude Pouchan. "DFT Modelling of the Infrared Spectra for Isolated and Aqueous Forms of Adenine". *Theoretical Chemistry Accounts* 136.4 (Apr. 2017), p. 44.

[146] Tomáš Hrivnák, Šimon Budzák, Heribert Reis, et al. "Electric Properties of Hydrated Uracil: From Micro- to Macrohydration". *Journal of Molecular Liquids* 275 (Feb. 1, 2019), pp. 338–346.

[147] Masao Iwamatsu. "Global Geometry Optimization of Silicon Clusters Using the Space-Fixed Genetic Algorithm". *The Journal of Chemical Physics* 112.24 (June 22, 2000), pp. 10976–10983.

[148] Jun Zhang and Vassiliki-Alexandra Glezakou. "Global Optimization of Chemical Cluster Structures: Methods, Applications, and Challenges". *International Journal of Quantum Chemistry* 121.7 (Apr. 5, 2021).

[149] Zhanghui Chen, Xiangwei Jiang, Jingbo Li, et al. "PDECO: Parallel Differential Evolution for Clusters Optimization". *Journal of Computational Chemistry* 34.12 (2013), pp. 1046–1059.

[150] Anastassia N. Alexandrova and Alexander I. Boldyrev. "Search for the $Li_n^{0/+1/-1}$ (n = 5-7) Lowest-Energy Structures Using the Ab Initio Gradient Embedded Genetic Algorithm (GEGA). Elucidation of the Chemical Bonding in the Lithium Clusters". *Journal of Chemical Theory and Computation* 1.4 (July 1, 2005), pp. 566–580.

[151] Yanchao Wang, Jian Lv, Li Zhu, et al. "CALYPSO: A Method for Crystal Structure Prediction". *Computer Physics Communications* 183.10 (Oct. 2012), pp. 2063–2070.

[152] Jun Zhang and Michael Dolg. "Global Optimization of Clusters of Rigid Molecules Using the Artificial Bee Colony Algorithm". *Physical Chemistry Chemical Physics* 18.4 (2016), pp. 3003–3010.

[153] Jun Zhang and Michael Dolg. "ABCluster: The Artificial Bee Colony Algorithm for Cluster Global Optimization". *Physical Chemistry Chemical Physics* 17.37 (2015), pp. 24173–24181.

[154] Leihou Shao, Kaiwei Wan, Hui Wang, et al. "A Non-Conjugated Polyethylenimine Copolymer-Based Unorthodox Nanoprobe for Bioimaging and Related Mechanism Exploration". *Biomaterials Science* 7.7 (2019), pp. 3016–3024.

[155] Jianbao Wu and Lin-Wang Wang. "2D Framework C_2N as a Potential Cathode for Lithium–Sulfur Batteries: An *Ab Initio* Density Functional Study". *Journal of Materials Chemistry A* 6.7 (2018), pp. 2984–2994.

[156] Xiangliang Ma, Shangguo Liu, and Shiping Huang. "Hydrogen Adsorption and Dissociation on the TM-doped (TM=Ti, Nb) Mg55 Nanoclusters: A DFT Study". *International Journal of Hydrogen Energy* 42.39 (Sept. 2017), pp. 24797–24810.

[157] Yucheng Xu, Ning Wang, Xiangyu Guo, et al. "Effects of the Coordination Number on H2O Dissociation Reaction on the Surface of Zr5nO10n (N=4–9) Nanoparticles: A DFT Approach". *International Journal of Hydrogen Energy* 44.59 (Nov. 2019), pp. 31029–31040.

[158] Alhadji Malloum, Jean J. Fifen, and Jeanet Conradie. "Large-Sized Ammonia Clusters and Solvation Energies of the Proton in Ammonia". *Journal of Computational Chemistry* 41.1 (Jan. 5, 2020), pp. 21–30.

[159] Zhiyan Wu, Kama Huang, and Xiaoyu Kuang. "Dielectric Properties of Pyridine–Ethanol Mixtures: Density Functional Theory and Experiments". *RSC Advances* 6.70 (2016), pp. 66007–66010.

[160] Jun Zhang, Eric T. Baxter, Manh-Thuong Nguyen, et al. "Structure and Stability of the Ionic Liquid Clusters $[EMIM]_n[BF_4]_{n+1}^-$ ($n = 1$–9): Implications for Electrochemical Separations". *The Journal of Physical Chemistry Letters* 11.16 (Aug. 20, 2020), pp. 6844–6851.

[161] Jilai Li, Shaodong Zhou, Jun Zhang, et al. "Mechanistic Variants in Gas-Phase Metal-Oxide Mediated Activation of Methane at Ambient Conditions". *Journal of the American Chemical Society* 138.35 (Sept. 7, 2016), pp. 11368–11377.

[162] Stanislav K. Ignatov, Andrey I. Okhapkin, Oleg B. Gadzhiev, et al. "Adsorption and Diffusion of Hydrogen on the Surface of the Pt_{24} Subnanoparticle. A DFT Study". *The Journal of Physical Chemistry C* 120.33 (Aug. 25, 2016), pp. 18570–18587.

[163] Van Tan Tran and Quoc Tri Tran. "Geometric and Electronic Structures of $VB_4^{0/+}$ Clusters and Reactivity of the Cationic Cluster with Methane from Quantum Chemical Calculations". *The Journal of Physical Chemistry A* 123.42 (Oct. 24, 2019), pp. 9223–9233.

[164] Gentoku Takasao, Toru Wada, Ashutosh Thakur, et al. "Machine Learning-Aided Structure Determination for $TiCl_4$-Capped $MgCl_2$ Nanoplate of Heterogeneous Ziegler–Natta Catalyst". *ACS Catalysis* 9.3 (Mar. 1, 2019), pp. 2599–2609.

[165] Huanchen Zhai, Philippe Sautet, and Anastassia N. Alexandrova. "Global Optimization of Adsorbate Covered Supported Cluster Catalysts: The Case of $Pt_7H_{10}CH_3$ on α-Al_2O_3". *ChemCatChem* 12.3 (Feb. 6, 2020), pp. 762–770.

[166] Sneha A. Akhade, Austin Winkelman, Vanessa Lebarbier Dagle, et al. "Influence of Ag Metal Dispersion on the Thermal Conversion of Ethanol to Butadiene over Ag-ZrO_2/SiO_2 Catalysts". *Journal of Catalysis* 386 (June 2020), pp. 30–38.

[167] Abdallah S. Abdelsattar, Alyaa Dawoud, and Mohamed A. Helal. "Interaction of Nanoparticles with Biological Macromolecules: A Review of Molecular Docking Studies". *Nanotoxicology* 15.1 (Jan. 2, 2021), pp. 66–95.

[168] Darwin Y. Fu and Jens Meiler. "Predictive Power of Different Types of Experimental Restraints in Small Molecule Docking: A Review". *Journal of Chemical Information and Modeling* 58.2 (Feb. 26, 2018), pp. 225–233.

[169] Gareth Jones, Peter Willett, and Robert C. Glen. "Molecular Recognition of Receptor Sites Using a Genetic Algorithm with a Description of Desolvation". *Journal of Molecular Biology* 245.1 (Jan. 1995), pp. 43–53.

[170] Gareth Jones, Peter Willett, Robert C Glen, et al. "Development and Validation of a Genetic Algorithm for Flexible Docking 1 1Edited by F. E. Cohen". *Journal of Molecular Biology* 267.3 (Apr. 1997), pp. 727–748.

[171] Garrett M Morris, David S Goodsell, Robert S Halliday, et al. "Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function". *Journal of computational chemistry* 19.14 (1998), pp. 1639–1662.

[172] Jeffrey S Taylor and Roger M Burnett. "DARWIN: A Program for Docking Flexible Molecules". *Proteins: Structure, Function, and Bioinformatics* 41.2 (2000), pp. 173–191.

[173] Ana Maria Faisca Phillips, Martin H. G. Prechtl, and Armando J. L. Pombeiro. "Non-Covalent Interactions in Enantioselective Organocatalysis: Theoretical and Mechanistic Studies of Reactions Mediated by Dual H-Bond Donors, Bifunctional Squaramides, Thioureas and Related Catalysts". *Catalysts* 11.5 (Apr. 29, 2021), p. 569.

[174] Matej Žabka and Radovan Šebesta. "Experimental and Theoretical Studies in Hydrogen-Bonding Organocatalysis". *Molecules* 20.9 (Aug. 26, 2015), pp. 15500–15524.

[175] Angel Avila, Rafael Chinchilla, Béla Fiser, et al. "Enantioselective Michael Addition of Isobutyraldehyde to Nitroalkenes Organocatalyzed by Chiral Primary Amine-Guanidines". *Tetrahedron: Asymmetry* 25.5 (Mar. 15, 2014), pp. 462–467.

[176] Pei-Gang Ding, Feng Zhou, Xin Wang, et al. "H-Bond Donor-Directed Switching of Diastereoselectivity in the Michael Addition of α-Azido Ketones to Nitroolefins". *Chemical Science* 11.15 (2020), pp. 3852–3861.

[177] Stephan J. Zuend and Eric N. Jacobsen. "Cooperative Catalysis by Tertiary Amino-Thioureas: Mechanism and Basis for Enantioselectivity of Ketone Cyanosilylation". *Journal of the American Chemical Society* 129.51 (Dec. 1, 2007), pp. 15872–15883.

[178] Stephan J. Zuend and Eric N. Jacobsen. "Mechanism of Amido-Thiourea Catalyzed Enantioselective Imine Hydrocyanation: Transition State Stabilization via Multiple Non-Covalent Interactions". *Journal of the American Chemical Society* 131.42 (Oct. 28, 2009), pp. 15358–15374.

[179] Adrien Quintard, Diana Cheshmedzhieva, Maria del Mar Sanchez Duque, et al. "Origin of the Enantioselectivity in Organocatalytic Michael Additions of β-Ketoamides to α,β-Unsaturated Carbonyls: A Combined Experimental, Spectroscopic and Theoretical Study". *Chemistry - A European Journal* 21.2 (Jan. 7, 2015), pp. 778–790.

[180] Dariusz G. Piekarski, Pascal Steinforth, Melania Gómez-Martínez, et al. "Insight into the Folding and Cooperative Multi-Recognition Mechanism in Supramolecular Anion-Binding Catalysis". *Chemistry – A European Journal* 26.72 (Dec. 23, 2020), pp. 17598–17603.

[181] Marco Durini, Florian A. Sahr, Michael Kuhn, et al. "Bifunctional 2,5-Diketopiperazines as Efficient Organocatalysts for the Enantioselective Conjugate Addition of Aldehydes to Nitroolefins". *European Journal of Organic Chemistry* 2011.28 (Oct. 2011), pp. 5599–5607.

[182] Antonio De Nino, Loredana Maiuolo, Pedro Merino, et al. "Efficient Organocatalyst Supported on a Simple Ionic Liquid as a Recoverable System for the Asymmetric Diels-Alder Reaction in the Presence of Water". *ChemCatChem* 7.5 (Mar. 2015), pp. 830–835.

[183] Yosuke Demizu, Mitsunobu Doi, Masaaki Kurihara, et al. "One-Handed Helical Screw Direction of Homopeptide Foldamer Exclusively Induced by Cyclic α-Amino Acid Side-Chain Chiral Centers". *Chemistry - A European Journal* 18.8 (Feb. 20, 2012), pp. 2430–2439.

[184] Takashi Misawa, Mitsuyoshi Imamura, Yuto Ozawa, et al. "Development of Helix-Stabilized Antimicrobial Peptides Composed of Lysine and Hydrophobic α,α-Disubstituted α-Amino Acid Residues". *Bioorganic & Medicinal Chemistry Letters* 27.17 (Sept. 2017), pp. 3950–3953.

[185] Karl N. Blodgett, Xiao Zhu, Patrick S. Walsh, et al. "Conformer-Specific and Diastereomer-Specific Spectroscopy of $A\beta\alpha$ Synthetic Foldamers: Ac-Ala-β_{ACHC}-Ala-NHBn". *The Journal of Physical Chemistry A* 122.14 (Apr. 12, 2018), pp. 3697–3710.

[186] S. M. Blinder. "Basic Concepts of Self-Consistent-Field Theory". *American Journal of Physics* 33.6 (June 1965), pp. 431–443.

[187] Chunyang Peng and H. Bernhard Schlegel. "Combining Synchronous Transit and Quasi-Newton Methods to Find Transition States". *Israel Journal of Chemistry* 33.4 (1993), pp. 449–454.

[188] Chunyang Peng, Philippe Y Ayala, H Bernhard Schlegel, et al. "Using Redundant Internal Coordinates to Optimize Equilibrium Geometries and Transition States". *Journal of Computational Chemistry* 17.1 (1996), pp. 49–56.

[189] Gilbert Newton Lewis. *Valence and the Structure of Atoms and Molecules*. 14. New York: Chemical Catalog Company, Incorporated, 1923.

[190] Johannes Diderik van der Waals. "Doctoral Dissertation". Leiden, 1873.

[191] Klaus Müller-Dethlefs and Pavel Hobza. "Noncovalent Interactions: A Challenge for Experiment and Theory". *Chemical Reviews* 100.1 (Jan. 1, 2000), pp. 143–168.

[192] Steven E. Wheeler, Trevor J. Seguin, Yanfei Guan, et al. "Noncovalent Interactions in Organocatalysis and the Prospect of Computational Catalyst Design". *Accounts of Chemical Research* 49.5 (May 17, 2016), pp. 1061–1069.

[193] Robert R. Knowles and Eric N. Jacobsen. "Attractive Noncovalent Interactions in Asymmetric Catalysis: Links between Enzymes and Small Molecule Catalysts". *Proceedings of the National Academy of Sciences* 107.48 (Nov. 30, 2010), pp. 20678–20685.

[194] F. Dean Toste, Matthew S. Sigman, and Scott J. Miller. "Pursuit of Noncovalent Interactions for Strategic Site-Selective Catalysis". *Accounts of Chemical Research* 50.3 (Mar. 21, 2017), pp. 609–615.

[195] Weijun Tang, Steven Johnston, Jonathan A. Iggo, et al. "Cooperative Catalysis through Noncovalent Interactions". *Angewandte Chemie International Edition* 52.6 (Feb. 4, 2013), pp. 1668–1672.

[196] Stephanie C. C. van der Lubbe and Célia Fonseca Guerra. "The Nature of Hydrogen Bonds: A Delineation of the Role of Different Energy Components on Hydrogen Bond Strengths and Lengths". *Chemistry – An Asian Journal* (July 19, 2019), asia.201900717.

[197] Corentin Lefebvre, Gaëtan Rubez, Hassan Khartabil, et al. "Accurately Extracting the Signature of Intermolecular Interactions Present in the NCI Plot of the Reduced Density Gradient versus Electron Density". *Physical Chemistry Chemical Physics* 19.27 (2017), pp. 17928–17936.

[198] R. F. W. Bader. *Atoms in Molecules. A Quantum Theory*. New York: Oxford University Press, 1990.

[199] Richard F. W. Bader. "The Quantum Mechanical Basis of Conceptual Chemistry". *Monatshefte für Chemie - Chemical Monthly* 136.6 (June 2005), pp. 819–854.

[200] P. Kumar, V. Raghavendra, and V. Subramanian. "Bader's Theory of Atoms in Molecules (AIM) and Its Applications to Chemical Bonding". *Journal of Chemical Sciences* 128.10 (Oct. 1, 2016), pp. 1527–1536.

[201] J. P. Foster and F. Weinhold. "Natural Hybrid Orbitals". *Journal of the American Chemical Society* 102.24 (Nov. 1980), pp. 7211–7218.

[202] Alan E. Reed and Frank Weinhold. "Natural Bond Orbital Analysis of near-Hartree–Fock Water Dimer". *The Journal of Chemical Physics* 78.6 (Mar. 15, 1983), pp. 4066–4073.

[203] Alan E. Reed, Robert B. Weinstock, and Frank Weinhold. "Natural Population Analysis". *The Journal of Chemical Physics* 83.2 (July 15, 1985), pp. 735–746.

[204] Alan E. Reed and Frank Weinhold. "Natural Localized Molecular Orbitals". *The Journal of Chemical Physics* 83.4 (Aug. 15, 1985), pp. 1736–1740.

[205] J. E. Carpenter. "Extension of Lewis Structure Concepts to Open-Shell and Excited-State Molecular Species". PhD thesis. Madison, Wisconsin: University of Wisconsin, 1987.

[206] J.E. Carpenter and F. Weinhold. "Analysis of the Geometry of the Hydroxymethyl Radical by the "Different Hybrids for Different Spins" Natural Bond Orbital Procedure". *Journal of Molecular Structure: THEOCHEM* 169 (Aug. 1988), pp. 41–62.

[207] Ron Naaman and Zeev Vager, eds. *The Structure of Small Molecules and Ions*. Boston, MA: Springer US, 1988.

[208] Jane S. Murray, Pat Lane, Tore Brinck, et al. "Relationships between Computed Molecular Properties and Solute-Solvent Interactions in Supercritical Solutions". *The Journal of Physical Chemistry* 97.19 (May 1993), pp. 5144–5148.

[209] Jane S. Murray, Pat Lane, Tore Brinck, et al. "Relationships of Critical Constants and Boiling Points to Computed Molecular Surface Properties". *The Journal of Physical Chemistry* 97.37 (Sept. 1993), pp. 9369–9373.

[210] Tore Brinck, Jane S. Murray, and Peter Politzer. "Quantitative Determination of the Total Local Polarity (Charge Separation) in Molecules". *Molecular Physics* 76.3 (June 20, 1992), pp. 609–617.

[211] Jane S. Murray, Tore Brinck, Pat Lane, et al. "Statistically-Based Interaction Indices Derived from Molecular Surface Electrostatic Potentials: A General Interaction Properties Function (GIPF)". *Journal of Molecular Structure: THEOCHEM* 307 (Apr. 1994), pp. 55–64.

[212] Bogumil Jeziorski, Robert Moszynski, and Krzysztof Szalewicz. "Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van Der Waals Complexes". *Chemical Reviews* 94.7 (Nov. 1994), pp. 1887–1930.

[213] Edward G. Hohenstein and C. David Sherrill. "Wavefunction Methods for Noncovalent Interactions: Noncovalent Interactions". *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2.2 (Mar. 2012), pp. 304–326.

[214] Tian Lu and Feiwu Chen. "Revealing the Nature of Intermolecular Interaction and Configurational Preference of the Nonpolar Molecular Dimers (H2)2, (N2)2, and (H2)(N2)". *Journal of Molecular Modeling* 19.12 (Dec. 2013), pp. 5387–5395.

[215] Trent M. Parker, Lori A. Burns, Robert M. Parrish, et al. "Levels of Symmetry Adapted Perturbation Theory (SAPT). I. Efficiency and Performance for Interaction Energies". *The Journal of Chemical Physics* 140.9 (Mar. 7, 2014), p. 094106.

[216] Yinchun Jiao, Yi Liu, Wenjing Zhao, et al. "Theoretical Study on the Interactions of Halogen-Bonds and Pnicogen-Bonds in Phosphine Derivatives with Br_2 , BrCl, and BrF". *International Journal of Quantum Chemistry* 117.22 (Nov. 15, 2017), e25443.

[217] Anthony J. Stone. "Natural Bond Orbitals and the Nature of the Hydrogen Bond". *The Journal of Physical Chemistry A* 121.7 (Feb. 23, 2017), pp. 1531–1534.

[218] B. Silvi and A. Savin. "Classification of Chemical Bonds Based on Topological Analysis of Electron Localization Functions". *Nature* 371.6499 (Oct. 1994), pp. 683–686.

[219] Saeedreza Emamian, Tian Lu, Holger Kruse, et al. "Exploring Nature and Predicting Strength of Hydrogen Bonds: A Correlation Analysis Between Atoms-in-Molecules Descriptors, Binding Energies, and Energy Components of Symmetry-Adapted Perturbation Theory". *Journal of Computational Chemistry* 40.32 (2019), pp. 2868–2881.

[220] Joseph R. Lane, Julia Contreras-García, Jean-Philip Piquemal, et al. "Are Bond Critical Points Really Critical for Hydrogen Bonding?" *Journal of Chemical Theory and Computation* 9.8 (Aug. 13, 2013), pp. 3263–3266.

[221] Waro Nakanishi, Satoko Hayashi, and Kenji Narahara. "Atoms-in-Molecules Dual Parameter Analysis of Weak to Strong Interactions: Behaviors of Electronic Energy Densities versus Laplacian of Electron Densities at Bond Critical Points". *The Journal of Physical Chemistry A* 112.51 (Dec. 25, 2008), pp. 13593–13599.

[222] Waro Nakanishi, Satoko Hayashi, and Kenji Narahara. "Polar Coordinate Representation of Hb(Rc) versus $(\hbar 2/8\mathrm{m})\bigtriangledown 2\rho\mathrm{b}(\mathrm{Rc})$ at BCP in AIM Analysis: Classification and Evaluation of Weak to Strong Interactions". *The Journal of Physical Chemistry A* 113.37 (Sept. 17, 2009), pp. 10050–10057.

[223] Karlheinz Drauz, Harald Gröger, and Oliver May. *Enzyme Catalysis in Organic Synthesis, 3 Volume Set*. John Wiley & Sons, Mar. 26, 2012. 2143 pp.

[224] Pratul K. Agarwal. "A Biophysical Perspective on Enzyme Catalysis". *Biochemistry* 58.6 (Feb. 12, 2019), pp. 438–449.

[225] Zebediah C. Girvin and Samuel H. Gellman. "Foldamer Catalysis". *Journal of the American Chemical Society* 142.41 (Oct. 14, 2020), pp. 17211–17223.

[226] David R. Kelly and Stanley M. Roberts. "The Mechanism of Polyleucine Catalysed Asymmetric Epoxidation". *Chemical Communications* 18 (Sept. 14, 2004), pp. 2018–2020.

[227] Stefano Colonna, Dario Perdicchia, and Ernesto Di Mauro. "Enantioselective Reactions Catalyzed by Synthetic Enzymes. A Model for Chemical Evolution". *Tetrahedron: Asymmetry* 20.15 (Aug. 12, 2009), pp. 1709–1714.

[228] Juliette Fremaux, Laura Mauran, Karolina Pulka-Ziach, et al. "α-Peptide–Oligourea Chimeras: Stabilization of Short α-Helices by Non-Peptide Helical Foldamers". *Angewandte Chemie International Edition* 127.34 (2015), pp. 9954–9958.

[229] Bryden A. F. Le Bailly, Liam Byrne, and Jonathan Clayden. "Refoldable Foldamers: Global Conformational Switching by Deletion or Insertion of a Single Hydrogen Bond". *Angewandte Chemie International Edition* 55.6 (2016), pp. 2132–2136.

[230] Zhiguo Zhang and Peter R. Schreiner. "(Thio)Urea Organocatalysis—What Can Be Learnt from Anion Recognition?" *Chemical Society Reviews* 38.4 (Mar. 24, 2009), pp. 1187–1198.

[231] Tian Lu and Feiwu Chen. "Multiwfn: A Multifunctional Wavefunction Analyzer". *Journal of Computational Chemistry* 33.5 (Feb. 15, 2012), pp. 580–592.

[232] K. Fukui. "Theory of Orientation and Stereoselection". *Orientation and Stereoselection*. Ed. by K. Fukui. Fortschritte Der Chemischen Forschung. Berlin, Heidelberg: Springer, 1970, pp. 1–85.

[233] Maria Besora, Ataualpa A. C. Braga, Gregori Ujaque, et al. "The Importance of Conformational Search: A Test Case on the Catalytic Cycle of the Suzuki–Miyaura Cross-Coupling". *Theoretical Chemistry Accounts* 128.4 (Mar. 1, 2011), pp. 639–646.

[234] Corentin Lefebvre, Hassan Khartabil, Jean-Charles Boisson, et al. "The Independent Gradient Model: A New Approach for Probing Strong and Weak Interactions in Molecules from Wave Function Calculations". *ChemPhysChem* 19.6 (Mar. 19, 2018), pp. 724–735.

[235] Miguel Ponce-Vargas, Corentin Lefebvre, Jean-Charles Boisson, et al. "Atomic Decomposition Scheme of Noncovalent Interactions Applied to Host–Guest Assemblies". *Journal of Chemical Information and Modeling* 60.1 (Jan. 27, 2020), pp. 268–278.

[236] Johanna Klein, Hassan Khartabil, Jean-Charles Boisson, et al. "New Way for Probing Bond Strength". *The Journal of Physical Chemistry A* 124.9 (Mar. 5, 2020), pp. 1850–1860.

[237] Iñigo Iribarren, Marianne Rica Garcia, and Cristina Trujillo. "Catalyst Design within Asymmetric Organocatalysis". *WIREs Computational Molecular Science* (Mar. 18, 2022).

[238] Ping Wu, Lei Yu, Cong-Hui Gao, et al. "Design and Synthesis of Axially Chiral Aryl-Pyrroloindoles via the Strategy of Organocatalytic Asymmetric (2 + 3) Cyclization". *Fundamental Research* (Jan. 2022), S2667325822000413.

[239] Jacqueline Bitai, Matthew T. Westwood, and Andrew D. Smith. "α,β-Unsaturated Acyl Ammonium Species as Reactive Intermediates in Organocatalysis: An Update". *Organic & Biomolecular Chemistry* 19.11 (2021), pp. 2366–2384.

[240] Alexey Yu Sukhorukov, Anna A Sukhanova, and Sergei G Zlotin. "Stereoselective Reactions of Nitro Compounds in the Synthesis of Natural Compound Analogs and Active Pharmaceutical Ingredients". *Tetrahedron* 72.41 (2016), pp. 6191–6281.

[241] Mohamed Samir Zaky. "Paires de catalyseurs organiques à base de (thio)urée chirale pour la polymérisation par ouverture de cycle stéréosélective du lactide racémique". PhD thesis. France: University of Bordeaux, Apr. 11, 2022. 274 pp.

[242] Nagendar Pendem, Céline Douat, Paul Claudon, et al. "Helix-Forming Propensity of Aliphatic Urea Oligomers Incorporating Noncanonical Residue Substitution Patterns". *Journal of the American Chemical Society* 135.12 (Mar. 27, 2013), pp. 4884–4892.

[243] Lucile Fischer, Claude Didierjean, Franck Jolibois, et al. "Propensity for Local Folding Induced by the Urea Fragment in Short-Chain Oligomers". *Organic & Biomolecular Chemistry* 6.14 (2008), p. 2596.

[244] P. C. Hariharan and J. A. Pople. "The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies". *Theoretica chimica acta* 28.3 (Sept. 1, 1973), pp. 213–222.

[245] K Throssel and MJ Frisch. "Evaluation and Improvement of Semiempirical Methods I: PM7R8: A Variant of PM7 with Numerically Stable Hydrogen Bonding Corrections". *manuscript in preparation* ().

[246] Matthew D. Liptak and George C. Shields. "Accurate p K_{a} Calculations for Carboxylic Acids Using Complete Basis Set and Gaussian-n Models Combined with CPCM Continuum Solvation Methods". *Journal of the American Chemical Society* 123.30 (Aug. 1, 2001), pp. 7314–7319.

[247] Ann Marie Toth, Matthew D. Liptak, Danielle L. Phillips, et al. "Accurate Relative pK_a Calculations for Carboxylic Acids Using Complete Basis Set and Gaussian-n Models Combined with Continuum Solvation Methods". *The Journal of Chemical Physics* 114.10 (2001), p. 4595.

[248] Matthew D. Liptak, Kevin C. Gross, Paul G. Seybold, et al. "Absolute pK_a Determinations for Substituted Phenols". *Journal of the American Chemical Society* 124.22 (June 1, 2002), pp. 6421–6427.

[249] Casey P. Kelly, Christopher J. Cramer, and Donald G. Truhlar. "Aqueous Solvation Free Energies of Ions and Ion- Water Clusters Based on an Accurate Value for the Absolute Aqueous Solvation Free Energy of the Proton". *The Journal of Physical Chemistry B* 110.32 (2006), pp. 16066–16081.

[250] Matthew D. Liptak and George C. Shields. "Experimentation with Different Thermodynamic Cycles Used for pKa Calculations on Carboxylic Acids Using Complete Basis Set and Gaussian-n Models Combined with CPCM Continuum Solvation Methods". *International Journal of Quantum Chemistry* 85.6 (2001), pp. 727–741.

[251] Barry R. Bickmore, Christopher J. Tadanier, Kevin M. Rosso, et al. "Bond-Valence Methods for pKa Prediction: Critical Reanalysis and a New Approach". *Geochimica et Cosmochimica Acta* 68.9 (May 2004), pp. 2025–2042.

[252] Rodrigo R. da Silva, Teodorico C. Ramalho, Joana M. Santos, et al. "On the Limits of Highest-Occupied Molecular Orbital Driven Reactions: The Frontier Effective-for-Reaction Molecular Orbital Concept". *The Journal of Physical Chemistry A* 110.3 (Jan. 1, 2006), pp. 1031–1040.

[253] Aziz Habibi-Yangjeh, Mohammad Danandeh-Jenagharad, and Mahdi Nooshyar. "Application of Artificial Neural Networks for Predicting the Aqueous Acidity of Various Phenols Using QSAR". *Journal of Molecular Modeling* 12.3 (Feb. 2006), pp. 338–347.

[254] Toru Matsui, Atsushi Oshiyama, and Yasuteru Shigeta. "A Simple Scheme for Estimating the pKa Values of 5-Substituted Uracils". *Chemical Physics Letters* 502.4-6 (Jan. 2011), pp. 248–252.

[255] Kowit Hengphasatporn, Toru Matsui, and Yasuteru Shigeta. "Estimation of Acid Dissociation Constants (pKa) of N-Containing Heterocycles in DMSO and Transferability of Gibbs Free Energy in Different Solvent Conditions". *Chemistry Letters* 49.3 (Mar. 5, 2020), pp. 307–310.

[256] Toru Matsui, Takeshi Baba, Katsumasa Kamiya, et al. "An Accurate Density Functional Theory Based Estimation of pKa Values of Polar Residues Combined with Experimental Data: From Amino Acids to Minimal Proteins". *Physical Chemistry Chemical Physics* 14.12 (2012), p. 4181.

[257] Gergely Jakab, Carlo Tancon, Zhiguo Zhang, et al. "(Thio)Urea Organocatalyst Equilibrium Acidities in DMSO". *Organic Letters* 14.7 (Apr. 6, 2012), pp. 1724–1727.