



HAL
open science

Structural characterization of chiral molecules : tool development for molecular simulations of vibrational circular dichroism spectra

Jessica Bowles

► **To cite this version:**

Jessica Bowles. Structural characterization of chiral molecules : tool development for molecular simulations of vibrational circular dichroism spectra. Theoretical and/or physical chemistry. Université Paris-Saclay, 2022. English. NNT : 2022UPASF063 . tel-04049880

HAL Id: tel-04049880

<https://theses.hal.science/tel-04049880>

Submitted on 29 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Structural characterisation of chiral molecules: tool development for molecular simulations of vibrational circular dichroism spectra

Caractérisation structurale des molécules chirales : développement d'outils de simulation moléculaire de spectres de dichroïsme circulaire vibrationnel

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 571 : Sciences chimiques : molécules, matériaux, instrumentation et biosystèmes (2MIB)
Spécialité de doctorat : Chimie théorique
Graduate School : Chimie, Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche Institut de Chimie Physique, sous la direction de Carine CLAVAGUERA, directrice de recherche CNRS, et de Florent CALVO, directeur de recherche CNRS.

Thèse soutenue à Paris-Saclay, le 28 septembre 2022, par

Jessica BOWLES

Composition du jury

Céline Toubin Professeure, PhLAM, Université de Lille	Rapporteur & Examinatrice
Benoît Champagne Professeur, Université de Namur, Belgique	Rapporteur & Examinateur
Elise Duboué-Dijon Chargée de recherche CNRS, LBT, IBPC Paris	Examinatrice
Christophe Humbert Directeur de recherche CNRS, ICP, Université Paris-Saclay	Président de jury & Examinateur
Carine Clavaguéra Directrice de recherche CNRS, ICP, Université Paris-Saclay	Directrice de thèse
Florent Calvo, Directeur de recherche CNRS, LIPhy, Université Grenoble Alpes	Directeur de thèse

Structural characterisation of chiral molecules: tool development for molecular simulations of vibrational circular dichroism spectra

Caractérisation structurale des molécules chirales : développement d'outils de simulation moléculaire de spectres de dichroïsme circulaire vibrationnel

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 571 : Sciences chimiques : molécules, matériaux, instrumentation et biosystèmes (2MIB)
Spécialité de doctorat : Chimie théorique
Graduate School : Chimie, Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche Institut de Chimie Physique, sous la direction de Carine CLAVAGUERA, directrice de recherche CNRS, et de Florent CALVO, directeur de recherche CNRS.

Thèse soutenue à Paris-Saclay, le 28 septembre 2022, par

Jessica BOWLES

Composition du jury

Céline Toubin Professeure, PhLAM, Université de Lille	Rapporteur & Examinatrice
Benoît Champagne Professeur, Université de Namur, Belgique	Rapporteur & Examineur
Elise Duboué-Dijon Chargée de recherche CNRS, LBT, IBPC Paris	Examinatrice
Christophe Humbert Directeur de recherche CNRS, ICP, Université Paris-Saclay	Président de jury & Examineur
Anne Zehnacker-Rentien Directrice de recherche CNRS, ISMO, Université Paris-Saclay	Invitée
Carine Clavaguéra Directrice de recherche CNRS, ICP, Université Paris-Saclay	Directrice de thèse
Florent Calvo, Directeur de recherche CNRS, LIPhy, Université Grenoble Alpes	Directeur de thèse

Acknowledgements

First I would like to thank my supervisors, **Carine Clavaguéra** and **Florent Calvo**, for offering me this opportunity, supporting me and giving me the confidence to accomplish my PhD.

Elise Duboué-Dijon and **Christophe Humbert** contributed to this thesis through the suivie de thèse and as examiners in my jury. I thank them for that. I also thank **Benoît Champagne** and **Céline Toubin** who were the reporters in my jury. Thank you Benoît for contributing to my passage from Miss to Dr Bowles.

I would like to thank the people involved in the ANR project Dichroprobe, without whom I would not have obtained all these fantastic results. First, thank you to **Sascha Jähnigen** who helped me tremendously throughout my PhD by answering all my questions, with technical support of ChirPy and with your involvement in the determination of the expression of the magnetic dipole moment. Thank you to **Federica Agostini** for her help with the assignments of spectral modes part of the thesis, **Anne Zehnacker** and **Katia Le Barbu-Debus** for the enthusiastic experimental collaborations and **Roldophe Vuilleumier** for his input to the project.

Thank you to those who recommended to me to use git, a version control system. It save my code several times. Thank you **Dominik Domin** for all the IT support.

The next part, I want to thank close members of the lab; **Rika** for dragging me to Hall of Beer regularly, **Damien** for all your help with coding and motivation to exercise, **Carlotta** for being around to talk last spring, **Valeriu**, for your encouragement through the writing of my thesis and putting up with my complaining about it, and of course for all the food you shared with me, **Jean** for the great times like when we changed your number plates and **Anouchka** for the year together as PhD representatives.

I would like to include **Christian** and **Lou**, for their motivation during their internships with me, **Nitik** for starting the muffin trend on that fateful day after a CESFO meal, the **members** of the Institut de Chimie Physique (**ICP**) laboratory and all the **ThéoSim** group for their kind interactions over the last 3 years.

Thank you to my **parents** for always being there to support me, and thank you dad for the last minute proofreading of my thesis. **Rosalyn**, thank you for putting up with your nerdy sister and for the champagne. Love you, dad, mom and Rosalyn. Thank you to friends and family, that I have not previously mentioned for being part of my life. Umber, you will not be able to read this, as your a horse, but thank you for all the amazing rides we've had together.

Last but definitely not least, **Maxime**, thank you for your profound support and love. The page of my thesis is turning, the next page I wish it to be with you. Je t'aime

Synthèse en français

Caractérisation structurale des molécules chirales : développement d'outils de simulation moléculaire de spectres de dichroïsme circulaire vibrationnel

Mots clés : Dichroïsme circulaire vibrationnel, dynamique moléculaire classique, molécules fluxionnelles

Le dichroïsme circulaire vibrationnel (VCD) est la faible différence d'absorption des molécules chirales entre la lumière polarisée droite et gauche dans le domaine de l'infrarouge (IR). Cette technique spectroscopique est devenue de plus en plus populaire depuis sa découverte dans les années 1970 et a des applications prometteuses en pharmacologie en raison de sa capacité à déterminer les configurations absolues des molécules chirales. Ceci est démontré dans la figure suivante où chaque énantiomère donne un signal identique sur le spectre infrarouge et opposé sur le spectre VCD :

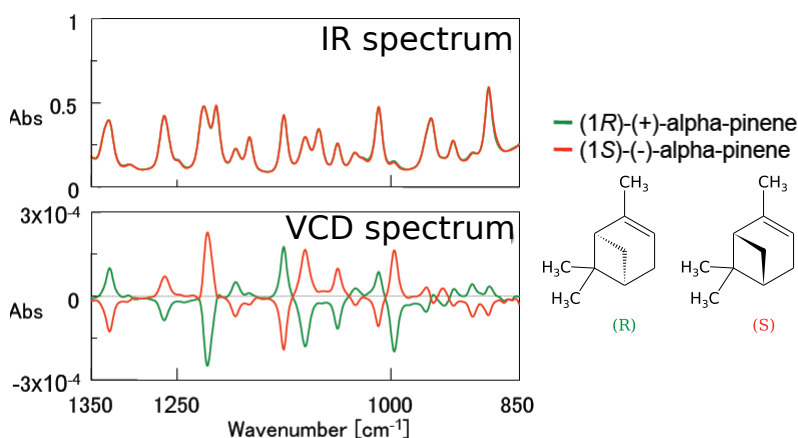


Figure 1 : Spectres IR et VCD de l'alpha-pinène. La figure a été adaptée de la référence [1].

La forme des spectres VCD est très sensible aux changements mineurs de conformation et aux interactions moléculaires, ce qui en fait une sonde de l'isomérisation conformationnelle et de la solvataion. L'attribution des spectres VCD expérimentaux nécessite en général une comparaison avec les calculs théoriques, ce qui implique également que le solvant soit précisément décrit.

Ce manuscrit décrit comment une modélisation basée sur des champs de force polarisables peut être exploitée pour interpréter les signaux VCD. Cette modélisation est appliquée à des molécules flexibles pour lesquelles les approches conventionnelles de

chimie quantique reposant sur des approximations statiques sont limitées. En particulier, il est difficile de prendre en compte les anharmonicités et les effets de température finie, et le solvant est typiquement modélisé de manière simplifiée par un continuum ou par un nombre limité de molécules explicites.

Dans le travail présenté ici, le champ de force AMOEBA mis en oeuvre dans le logiciel Tinker a été utilisé pour modéliser les moments dipolaires électriques et magnétiques de diverses molécules pour des applications en phase gazeuse, ainsi que dans les phases condensées des systèmes solvatés ou cristallins. L'échantillonnage et les informations dynamiques ont été obtenus par simulations de dynamique moléculaire classique. Les spectres IR et VCD anharmoniques obtenus à partir des trajectoires ont été interprétés à l'aide d'une analyse des modes effectifs dans laquelle les coordonnées internes sont décomposées linéairement par projection sur des modes effectifs appropriés, les poids correspondants étant obtenus de telle sorte que le pic de chaque mode soit le plus localisé possible. Cette analyse permet de mettre en évidence les contributions des différents groupes fonctionnels de la molécule aux pics individuels du spectre.

Comme premier exemple, le cas du trans-1-amino-2-indanol solvaté dans le diméthylsulfoxyde (DMSO) a été examiné en combinant des explorations statiques au niveau de la théorie de la fonctionnelle de la densité, des trajectoires courtes de dynamique moléculaire ab initio (AIMD) et un large échantillonnage de la surface d'énergie potentielle à l'aide du champ de force. La comparaison avec les mesures existantes confirme l'importance de la flexibilité pour cette molécule, le rôle des molécules de solvant proches et leur contribution aux signaux VCD. La fluxionalité du système composé d'une molécule de trans-1-amino-2-indanol et d'une molécule de DMSO est illustrée dans la figure de courbes de contour suivante :

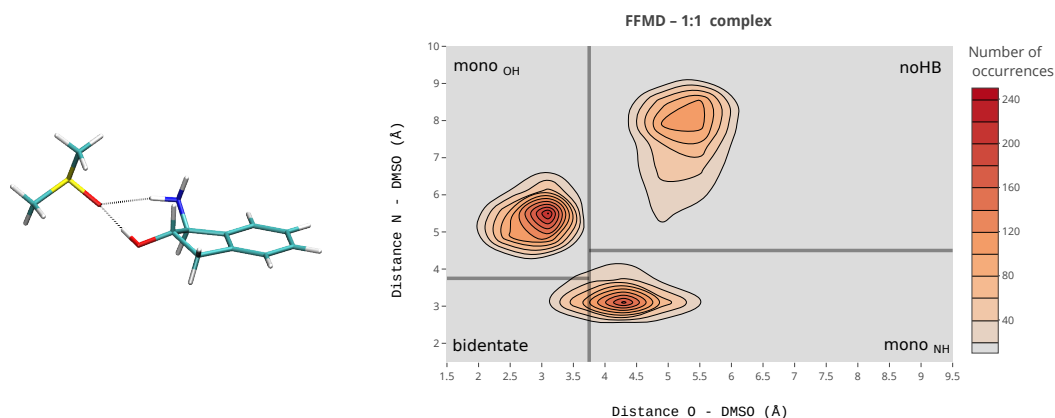


Figure 2 : Courbes de contour de l'occurrence des distances O(trans-AI)-DMSO et N(trans-AI)-DMSO du complexe 1:1 du trans-1-amino-2-indanol à 300 K déterminé à partir de simulations avec le champ de force

La détermination des spectres VCD à partir du champ de force polarisable a ensuite été entreprise et ses performances ont été évaluées par comparaison avec des calculs de structure électronique. Cette comparaison a été effectuée sur la même trajectoire MD,

en utilisant l'acide aminé alanine en phase gazeuse comme système de référence. Les différentes contributions aux moments dipolaires sont discutées.

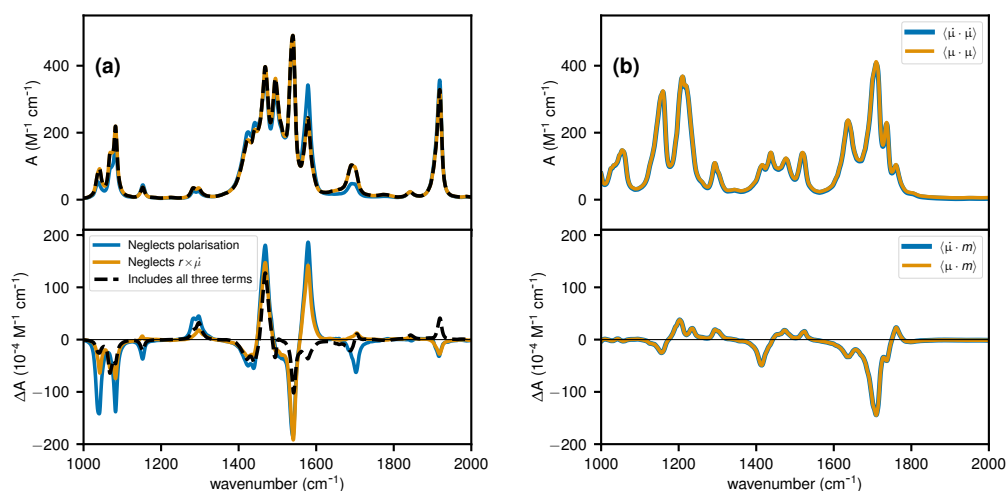


Figure 3 : Spectres IR et VCD de la forme (a) zwitterionique et (b) neutre de l'alanine moyennés sur 40 trajectoires avec un champ de force à 300 K. La figure (a) examine l'addition des différentes contributions au moment dipolaire magnétique et la figure (b) montre les spectres produits avec le moment dipolaire électrique et avec sa dérivée.

Après avoir validé la méthodologie, la convergence statistique des spectres IR et VCD a été quantifiée via des mesures d'erreurs dédiées. Ces erreurs montrent à leur tour que la convergence des spectres VCD est plus lente que celle des spectres IR et quelle n'est atteinte qu'après un échantillonnage bien plus important que celui réalisable par AIMD.

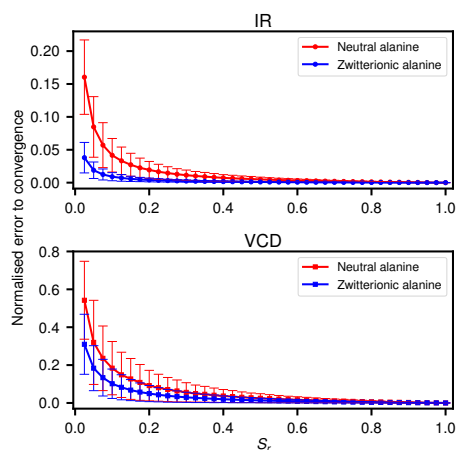


Figure 4 : Erreur normalisée des moyennes des spectres IR et VCD de l'alanine isolé de N_T trajectoires avec un champ de force à 300 K, en fonction de la progression du rapport d'échantillonnage $S_r = N_T/40$.

L'approche par champ de force polarisable a ensuite été appliquée à l'alanine en phases condensées. Là encore, la comparaison avec les simulations AIMD confirme que les moments dipolaires électriques et magnétiques peuvent être évalués de manière fiable.

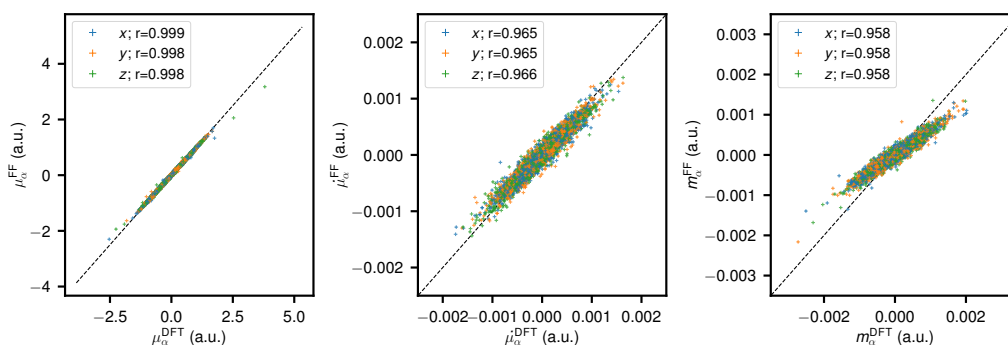


Figure 5 : Corrélation du moment dipolaire électrique $\vec{\mu}$, de sa dérivée $\dot{\vec{\mu}}$ et du moment dipolaire magnétique \vec{m} de l'alanine hydratée.

Ce travail ouvre la voie à la détermination du spectre VCD de systèmes chimiquement complexes inaccessibles via des méthodes qui tiennent compte explicitement de la structure électronique.

Bibliography

- [1] Jasco, *Spectra of alpha-pinene*, 2022. <https://jascoinc.com/applications/measurement-of-vibrational-circular-dichroism-spectra-using-the-fvs-6000/> (accessed April 25th, 2022).

Abbreviations

AIMD	Ab initio molecular dynamics
AMOEBA	Atomic multipole optimized energetics for biomolecular simulation
BOMD	Born-Oppenheimer molecular dynamics
CPMD	Car-Parrinello molecular dynamics
CSVR	Canonical sampling through velocity rescaling
DFT	Density functional theory
DHB	Dihydrogen bond
DMA	Distributed multipole analysis
DMD	Driven molecular dynamics
DMSO	Dimethyl sulfoxide
DO	Distributed origin
ECD	Electronic circular dichroism
FF	Force field
FFMD	Force field-based molecular dynamics
FTIR	Fourier transform infrared
FWHM	Full width at half maximum
IEFPM	Integral equation formalism polarisable model
INMA	Instantaneous normal mode analysis
IR	Infrared
LCP	left-circularly polarised
MCD	Magnetic circular dichroism
MD	Molecular dynamics
MWC	Mass-weighted coordinates
NVE	Microcanonical ensemble
NVPT	Nuclear velocity perturbation theory
NVT	Canonical ensemble
PBC	Periodic boundary conditions
PCM	Polarisable continuum model
PECD	Photoelectron circular dichroism
QM/MM	Quantum mechanics/molecular mechanics
RCP	Right-circularly polarised
RDF	Radial distribution functions
RMSD	Root-mean-square deviation
SERS	Surface-enhanced Raman spectroscopy
TCF	Time-correlation function
trans-AI	Trans-1-amino-2 indanol
VCD	Vibrational circular dichroism
vdW	van der Waals
VPT2	Second-order vibrational perturbation theory

Contents

1	Introduction	11
1.1	Chirality	11
1.2	Vibrational spectroscopy	13
1.2.1	Infrared spectroscopy	13
1.2.2	Raman spectroscopy	14
1.2.3	Circular dichroism	15
1.3	Vibrational circular dichroism	17
1.3.1	Introduction to vibrational circular dichroism	17
1.3.2	Experimental VCD spectroscopy	18
1.3.3	Theoretical VCD and the comparison to experimental VCD spectra	19
1.4	Thesis outline	22
2	Methodology	29
2.1	General theory of vibrational circular dichroism	29
2.2	Harmonic approach	31
2.3	Density-functional theory	33
2.3.1	Basic principles of DFT	33
2.3.2	Application to infrared and vibrational circular dichroism spectroscopies	35
2.4	Anharmonicity effects	36
2.4.1	Perturbation approaches	38
2.4.2	Boltzmann weights and the cluster-in-a-liquid model	39
2.5	Classical molecular dynamics	40
2.5.1	Principle of molecular dynamics	41
2.5.2	Thermostatted molecular dynamics	42
2.5.3	Infrared and VCD absorption spectroscopies from time-dependent properties	43
2.6	Ab initio molecular dynamics	44
2.6.1	Born-Oppenheimer molecular dynamics	45
2.6.2	Car-Parrinello molecular dynamics	45
2.6.3	IR and VCD spectra from nuclear velocity perturbation theory	45
2.6.4	Example of propylene oxide: evidence for chirality transfer	47
2.7	Classical force fields for molecular simulations	49
2.7.1	Nonpolarisable contributions	49
2.7.2	Accounting for polarisation	53
2.7.3	Periodic boundary conditions	56
2.7.4	The AMOEBA force field	56
2.8	Implementation of VCD spectroscopy within AMOEBA	59

2.8.1	Specific difficulties under periodic boundary conditions	60
2.8.2	From the time correlation functions to the spectra	61
2.8.3	Assignment of simulation parameters: the case of trans-1-amino-2 indanol	62
2.9	Assignment of spectral modes	64
2.9.1	Methods to obtain vibrational modes	64
2.9.2	Effective mode analysis	66
2.10	Concluding remarks	67
3	Trans-1-amino-2-indanol: a case study for VCD and fluxionality	79
3.1	Introduction	79
3.1.1	Systems of interest	79
3.1.2	An example of the study of VCD spectra calculated using a cluster-in-a-liquid model: cis-1-amino-2-indanol	80
3.2	Trans-1-amino-2-indanol: cluster-in-a-liquid approach to VCD spectroscopy	82
3.2.1	Nomenclature	82
3.2.2	Cluster-in-a-liquid approach	82
3.3	Beyond cluster-in-a-liquid: exploration of the potential energy surface with molecular dynamics	85
3.3.1	AIMD approach	86
3.3.2	Force field based MD	88
3.4	Locality of vibrational absorption and VCD and the role of the solvent	98
3.5	Concluding remarks	100
4	Validation of the polarisable force field approach for VCD spectroscopy: alanine in the gas phase	105
4.1	Vibrational spectra obtained from force field molecular dynamics (FFMD)	106
4.1.1	Calculating the electric and magnetic dipole moments for the IR and VCD spectra	107
4.1.2	Statistical errors and convergence rates	108
4.2	Comparison with electronic structure calculations	112
4.2.1	Structures and relative energies	112
4.2.2	IR and VCD spectra from static DFT calculations	113
4.3	Assessing moments and spectra from one trajectory	116
4.4	Assignment of the IR spectra into effective modes	118
4.4.1	Analysis of the zwitterionic and neutral forms of alanine	119
4.4.2	Fragment analysis	126
4.5	Concluding remarks	129
5	Applications in condensed phases	133
5.1	Crystalline alanine	134
5.1.1	Vibrational spectra obtained from force field molecular dynamics	134
5.1.2	Assessing moments and spectra from one trajectory	138
5.2	Hydrated alanine	139
5.2.1	Vibrational spectra obtained from force field molecular dynamics	139
5.2.2	Statistical errors and convergence rates	142
5.2.3	Assessing dipole moments and spectra from one trajectory	144
5.3	Effective mode analyses for crystalline and hydrated alanine	146

5.4	Alanine solvated in dinitrogen	150
5.4.1	Vibrational spectra obtained from force field molecular dynamics	151
5.5	Phenylcyclohexanediol in the crystalline phase and solvated in DMSO	152
5.5.1	Vibrational spectra obtained from force field molecular dynamics	153
5.5.2	Statistical errors and convergence rates	155
5.6	Concluding remarks	157
6	Conclusions and perspectives	159

Chapter 1

Introduction

1.1. Chirality

A *chiral* molecule is a type of molecule that cannot be superimposed on its mirror image; like our hands. This resemblance is also shown in the etymology of the word chirality which is derived from the Ancient Greek $\chi\epsilon\iota\rho$ (kheir) "hand". A chiral molecule has two enantiomers denoted R and S, represented in figure 1.1. These names come from *rectus* and *sinister*, the Latin for *right* (as in the sense of correct or virtuous) and *left*.

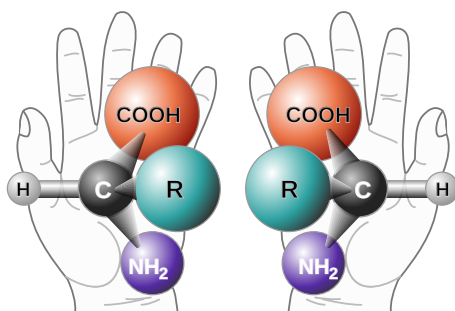


Figure 1.1: One chiral molecule, two enantiomers.

Many occurrences of chirality can be found in Nature. For example, mint and caraway seeds are composed of the same molecule: carvone. The mint contains the R enantiomer and the caraway seed the S enantiomer [1]. The shells of snails provide another case of chirality, with most shells turning anticlockwise when the head of the snail is to the left. A small minority have a shell of opposite chirality, with the shell turning clockwise [2]. Both of these examples are depicted in figure 1.2.

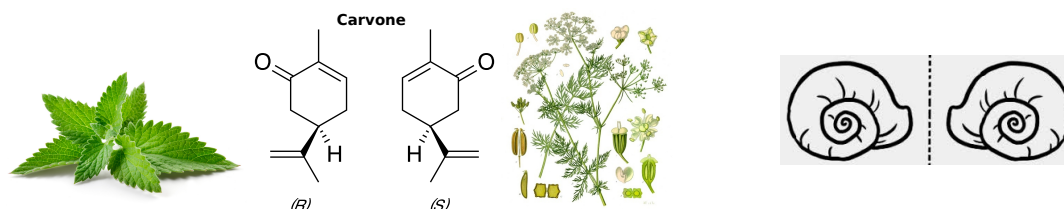


Figure 1.2: (left) The different smells of each enantiomer of carvone; (right) Chirality of the snail's shell.

Amino acids are organic molecules composed of an amino group (-NH₂) and a carboxyl group (-COOH). They are an essential building block of life as they form proteins. Most amino acids are chiral and have a dominant form just like snail shells. The nomenclature for amino acids is different, they are referred to as L (laevus) and D (dexter). To tell which enantiomer is shown, the CORN law can be used. CORN is an acronym for -COOH, the -R group (representation of side chain) and -NH₂. According to the CORN law, looking along the hydrogen-carbon bond represented in figure 1.1, if the -COOH, R and NH₂ are clockwise then the molecule is in the L form. If they are anticlockwise then the molecule is in the D form. The L form corresponds to the R enantiomer and the D form to the S notation. This is independent of the (+) and (-) notation, where the (-) configuration rotates linear polarised light anticlockwise and the (+) configuration rotates this light clockwise. Most amino acids are of the L form, and this enantiomeric excess is also found in meteorites [3]. On the contrary, sugars are found mostly in the D form. One hypothesis for this dissimilarity is that the L amino-acids caused the preference for D sugars by acting as a catalyst [4].

Deoxyribonucleic acid (DNA), which contains the genetic information of living organisms, forms a double helix, as shown in figure 1.3. This structure is chiral, as any kind of helix. There are 3 main configurations of DNA: A-DNA, B-DNA, and Z-DNA. The first two are right-handed and the last one left-handed. In living cells, most of the DNA is in a mixture of A- and B-DNA, with a few small regions capable of forming Z-DNA [5]. Like the helix, some mathematical objects also include chirality such as the Möbius strip.

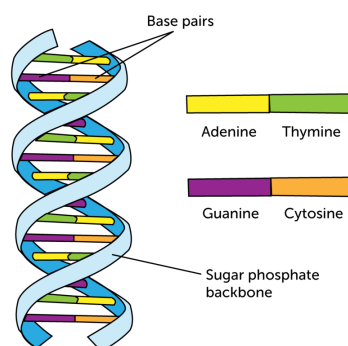


Figure 1.3: Visual representation of deoxyribonucleic acid (DNA). This figure has been reprinted from reference [6].

Finally, chiral molecules are present in medication. In this case, one enantiomer can cure but the other can be harmful. For example, the S enantiomer of ethambutol cures tuberculosis but the other enantiomer causes blindness [7]. It is the same for D-penicillamine, which is used for the treatment of rheumatoid arthritis, while L-penicillamine is toxic as it inhibits the action of pyridoxine, an essential B vitamin [8]. It can thus be crucial to differentiate between the two enantiomers of the same chiral molecule.

1.2. Vibrational spectroscopy

Spectroscopy refers to the response of matter, in absorption or emission, to an excitation by an oscillating electromagnetic field as a function of the wavelength or frequency of this oscillating field. A light beam has its electric and magnetic component both perpendicular to the direction of propagation. A specific direction for these components can be selected with a polariser. A vertical polariser will give linear polarised light. The infrared (IR) light has a wavelength longer than visible light, as shown in figure 1.4. The mid-IR-range corresponds to light with a wavelength between 16 667 nm and 2 500 nm, or between 600 and 4 000 cm^{-1} . Molecules absorb light in this range at wavelengths corresponding to the vibrations between its atoms. Thus, depending on how the atoms are assembled in the structure, different wavelengths will be absorbed.

Spectroscopy that studies molecular vibrations is called vibrational spectroscopy. It is a useful tool to determine the structural composition of complex molecules, including organic molecules. Current practice relies on the (successful) comparison between experimental measurements and theoretically calculated spectra, which is particularly useful to assign the peaks to specific molecular motions. Two methods of vibrational spectroscopy will be discussed here: IR and Raman spectroscopy.

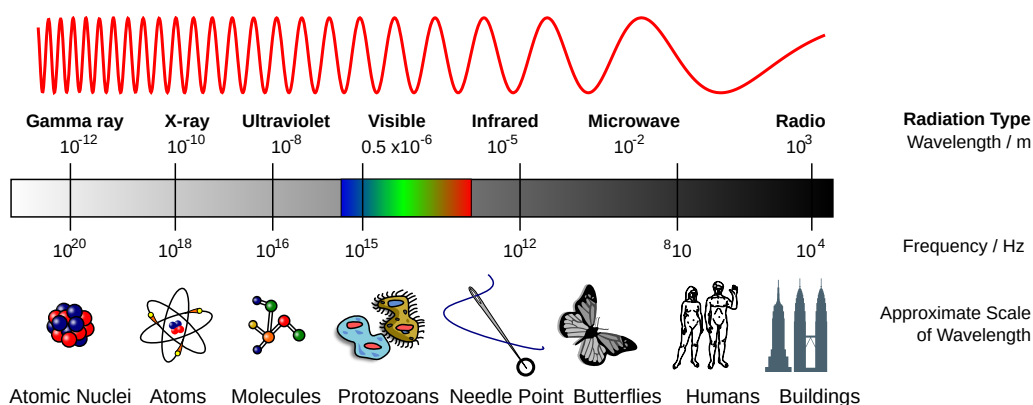


Figure 1.4: Spectrum of electromagnetic waves (Adapted from an image of Inductiveload [CC BY-SA 3.0] [9])

1.2.1. Infrared spectroscopy

The most common vibrational spectroscopic technique relies only on the absorption of IR light. The frequency that is absorbed matches the frequency of the molecular vibration or vibrational mode of a sample. In order for a vibrational mode to be considered as IR active, it must be associated with changes in the electric dipole moment.

IR spectroscopy is exploited in organic and inorganic chemistry. It is used by chemists in the determination of molecular structures and functional groups and in the qualitative and quantitative identification of chemical species. This is possible as, usually, stronger bonds and lighter atoms will vibrate at a higher stretching frequency, thus the different groups will show various peaks at different places in the spectrum. IR spectroscopy is

used in the investigation of matter in the solid, liquid, and gas phases. If a solvent, for example water which has a large electric dipole, is used, it must not absorb IR light in the same region as the molecule of interest as not to hide the signal.

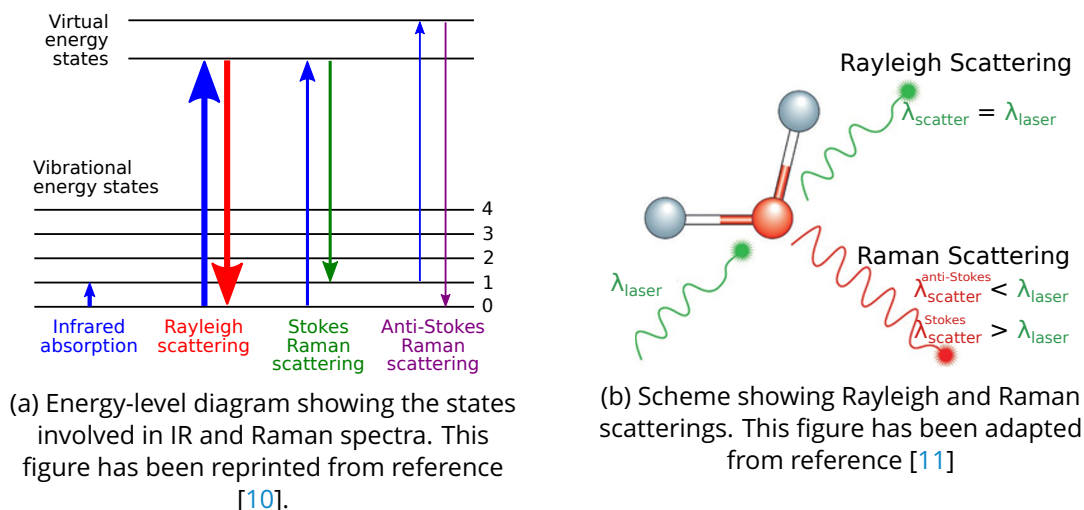


Figure 1.5: Main mechanisms involved in IR and Raman spectroscopies.

1.2.2. Raman spectroscopy

Another vibrational spectroscopic technique uses light scattering, which was first predicted by Smekal in 1923 [12]. In this phenomenon, the incident light of energy $\hbar\omega_i$ is scattered by a sample and subsequently some light of energy $\hbar\omega_s$ is emitted. Most of the scattered light is at the same wavelength as the incident light source and does not provide useful information; this is called Rayleigh scattering. However a small amount of light (typically 0.0000001%) is scattered at different wavelengths, which depend on the chemical structure of the system; this is called Raman scattering. This effect is named after one of its discoverers, who observed it in organic liquids in 1928 together with Krishnan. Independently Landsberg and Mandelstam observed the effect in inorganic crystals [13, 14]. The difference in wavelength in Raman scattering arises from a photon exciting the molecule into a virtual energy state for a short time before the photon is emitted. The emitted photon is of either lower or higher energy than the incident photon, as depicted in figure 1.5. While IR intensities depend on the changes in the electric dipole moment associated with the corresponding molecular motion, Raman scattering probes the higher-order component, namely the electronic polarisability.

Raman scattering is used to identify molecular structures, as IR spectroscopy does. Both spectroscopic methods are used for determining chemical bonds, with IR being sensitive to hetero-nuclear functional group vibrations and polar bonds, especially OH stretching in water, whereas Raman scattering is more sensitive to homo-nuclear molecular bonds. For example, it can distinguish between C-C, C=C and C≡C bonds. In solid-state physics, Raman spectroscopy is used to characterize materials, measure temperatures,

and locate the crystallographic orientation of a sample.

The key advantage of Raman spectroscopy is that it requires little to no sample preparation while the IR method has constraints on sample thickness, uniformity and dilution to avoid saturation. Raman spectra can be collected from a very small volume ($<1\ \mu\text{m}$ in diameter) and also water does not interfere with the signal. However with Raman spectroscopy, fluorescence can obstruct the ability to record the spectra, the intense laser radiation can destroy the sample or swamp the Raman spectrum and Raman scattering is typically very weak. To enhance the weak scattering in the Raman process, nanostructured materials can be used as a support. This is called surface-enhanced Raman spectroscopy (SERS) and is depicted in figure 1.6.

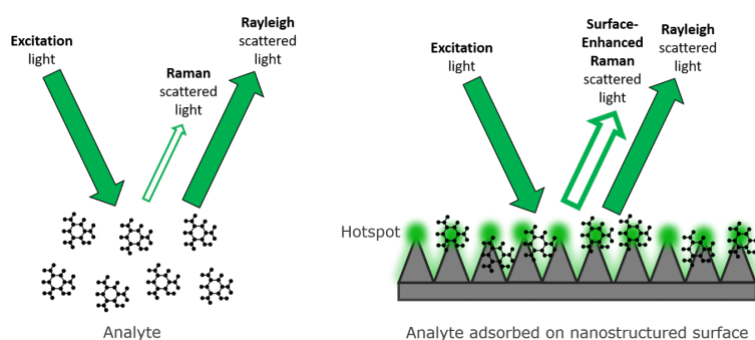


Figure 1.6: Raman scattering enhancement from a surface-enhanced Raman spectroscopy (SERS) substrate. This figure has been reprinted from reference [15].

1.2.3. Circular dichroism

While conventional spectroscopy in the IR or optical ranges provides a wealth of information about molecular structure, it is unable to distinguish between enantiomers. The incident electromagnetic field itself must carry some chirality in order to make the two enantiomers of a same molecule react differently, and this is achieved by using circularly polarised light.

With a quarter-wave plate, linear polarisation can be turned into circular polarisation. Such a plate dephases a given linear component of light relative to its orthogonal linear component. Circularly polarised light turns around the axis of propagation in both directions giving, left- and right- handed circular polarisations. Both of these polarisations form a helix which has a chiral form. Depending of the direction of polarisation, chiral molecules absorb this light differently. This is called circular dichroism (CD) (see figure 1.7).

Circular dichroism is mainly used in two ranges: ultraviolet/visible (UV-vis) and the IR range. The UV-vis range gives rise to electronic circular dichroism (ECD) and the IR range gives vibrational circular dichroism (VCD).

The rotation of polarised light by a molecule can also be induced by the application of a magnetic field. This perturbs the energy levels of the system, as was first observed by

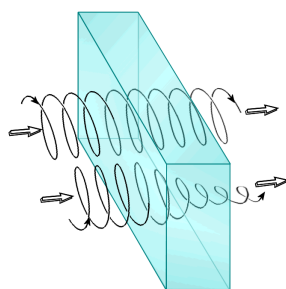


Figure 1.7: Chiral matter absorbing left circularly polarised light but transmitting right circularly polarised light. This defines circular dichroism.

Faraday [16]. This is called the Faraday effect, and is the basis of magnetic circular dichroism (MCD) [17]. It can be used to probe both the electronic and geometric structures of metal-containing systems, such as metal centres in metalloproteins [18].

Circular dichroism can also be observed in photoelectron angular distributions following photoionization by left- and right-circularly polarised light [19]. This is called photoelectron circular dichroism (PECD). The photoelectron angular distributions show the distribution of directions where the electrons are emitted. A study of camphor, by Garcia and coworkers showed changes in the photoelectron angular distributions depending on the enantiomer and on the direction of the polarisation of light, as depicted in figure 1.8.

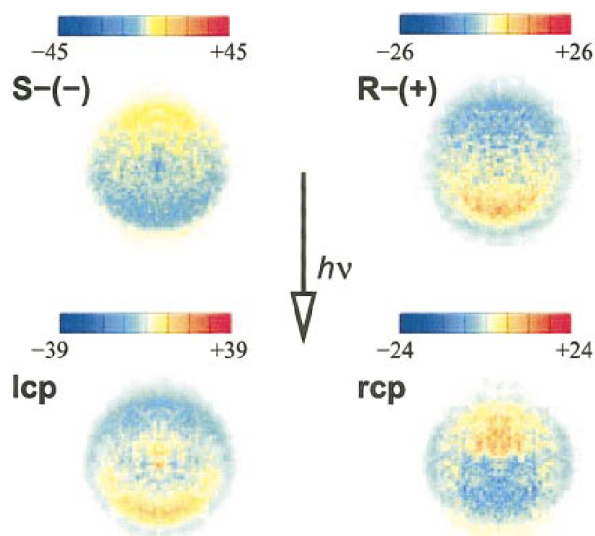


Figure 1.8: (Color) 2D projection images showing dichroism in the photoelectron angular distributions obtained from camphor enantiomers at $h\nu = 9.2$ eV using left- and right-circularly polarised light (labeled lcp and rcp, respectively). Top row: the difference (rcp-lcp) for pure S(-) and R(+) enantiomers; Bottom row: the difference between molecular enantiomers [S(-)-R(+)] with light of a fixed helicity (either lcp or rcp). This figure has been reprinted from reference [19].

Electronic circular dichroism

Circular dichroism was first discovered by the French scientist Cotton in 1896, who measured CD in the near-ultraviolet (UV) spectral region [20, 21]. UV-visible light is used to scrutinize electronic transitions of a system [22].

In the context of biopolymers, ECD is useful to give insight into the local planar amides which have conformational chirality due to their coupling in the polymer. This coupling gives the characteristic ECD bandshapes for helices and sheets. ECD requires lower concentrations than VCD and has little spectral interference with important solvents. However, important transitions of biosystems are present at wavelengths below 250 nm, which extend into a region (under 200 nm) where ECD is less sensitive and where there is more interference from both water and quartz. In addition, spectral simulation of ECD is challenging due to the fact that it involves transitions to excited electronic states, which are demanding for quantum chemistry methods [23].

1.3. Vibrational circular dichroism

1.3.1. Introduction to vibrational circular dichroism

A chiral molecule is considered with its two enantiomers. Each enantiomer absorbs the opposite amount of left or right circularly polarised light. The vibrational circular dichroism intensities arise from the difference between the absorbance of left-circularly polarised (LCP) and right-circularly polarised (RCP) light: $\Delta A = A_L - A_R$, in the IR range. The intensities of the VCD spectrum for each enantiomer are opposite in intensity [24]. This is shown in figure 1.9 with the IR and VCD spectra of alpha-pinene taken as an example. The spectrum of the R enantiomer is represented in green and the one of the S enantiomer in orange. The IR spectra overlap but their VCD spectra are opposite.

VCD has become an invaluable tool in the determination of absolute configurations by assigning recorded spectra to structural information obtained from theoretical calculations [25]. It has been applied in the fields of nanoscience [26], catalysis [27], solid state organisation [28, 29, 30], and has also been employed to detect protein fibrils such as amyloids that are responsible for neurodegenerative disorders [31].

From a more fundamental point of view, VCD provides a very sensitive probe of conformational flexibility and molecular interactions [32, 33, 34, 35, 36, 37, 38, 39, 40]. At room temperature, flexible molecules can adopt many stable conformations corresponding to different local minima, each of them contributing to the VCD spectrum.

Non-covalent interactions have a strong effect on the VCD spectra, even in cases where IR absorption remains unchanged [37, 38, 39, 41]. Consequently, effects like supra-molecular chirality and chirality transfer can be addressed as well [33, 42].

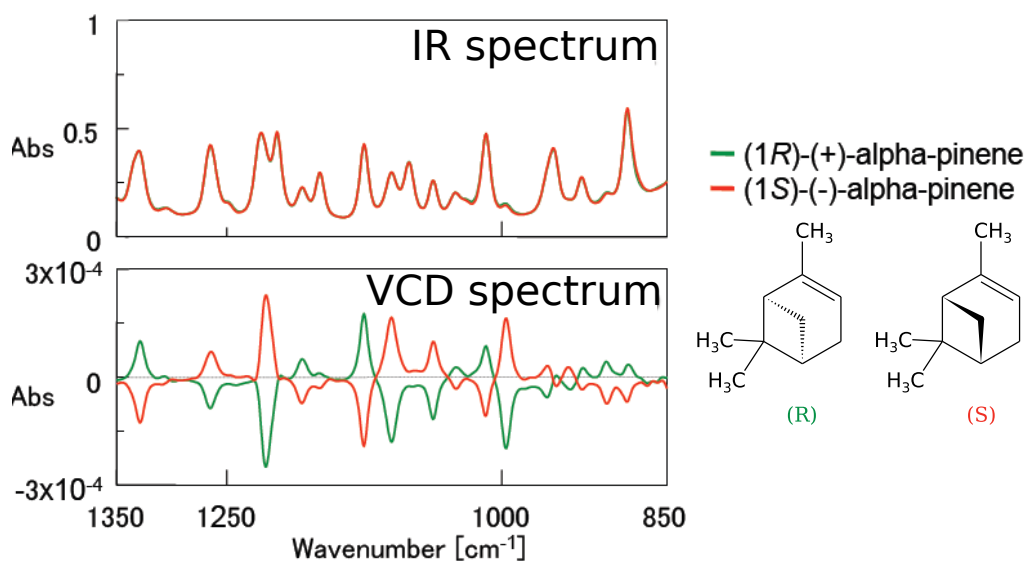


Figure 1.9: Spectra of alpha-pinene (neat, 50 μm pathlength BaF_2 liquid cell). This figure has been adapted from reference [24], with Abs. corresponding to the absorption.

1.3.2. Experimental VCD spectroscopy

While early experimental predictions of VCD were first conceptualised around 1970 [43, 44, 45, 46, 47], experimental work started on instruments capable of measuring VCD at the Stephens laboratory [48] and the Holzwarth laboratory [49]. VCD was first observed in the liquid state by Holzwarth et al. in 1974 [50], and by Nafie et al. in 1975 [51] on 2,2,2-trifluoro-L-phenylethanol.

To make VCD a practical technique for determining the enantiomer of a chiral molecule, two essential tools needed to be developed. First, a spectrometer with a frequency limit lower than 1600 cm^{-1} and a higher sensitivity (i.e. the signal-to-noise ratio) since VCD signals are small ($\sim 10^{-4}$) compared to IR signals ($\sim 10^0$). This first goal was achieved around 1980 with Fourier transform infrared (FTIR) spectrometers.

The usefulness of the technique could be demonstrated for differentiating polypeptides. For example, poly-L-lysine can be converted from a disordered structure to α -helix and to β -sheet conformations by increasing the pH and the temperature [23]. This type of study in an aqueous solution is simplified by exchanging H_2O to D_2O , which involves an H/D exchange of the amides but enables measurements on more dilute solutions. In figure 1.10, the corresponding IR spectra are shown in the lower frames and the VCD spectra in the upper frames. These correspond to the spectra of the three configurations of poly-L-lysine and show that both helix conformations have nearly identical IR spectra but different VCD spectra. The peaks correspond to the amine I region of the spectra.

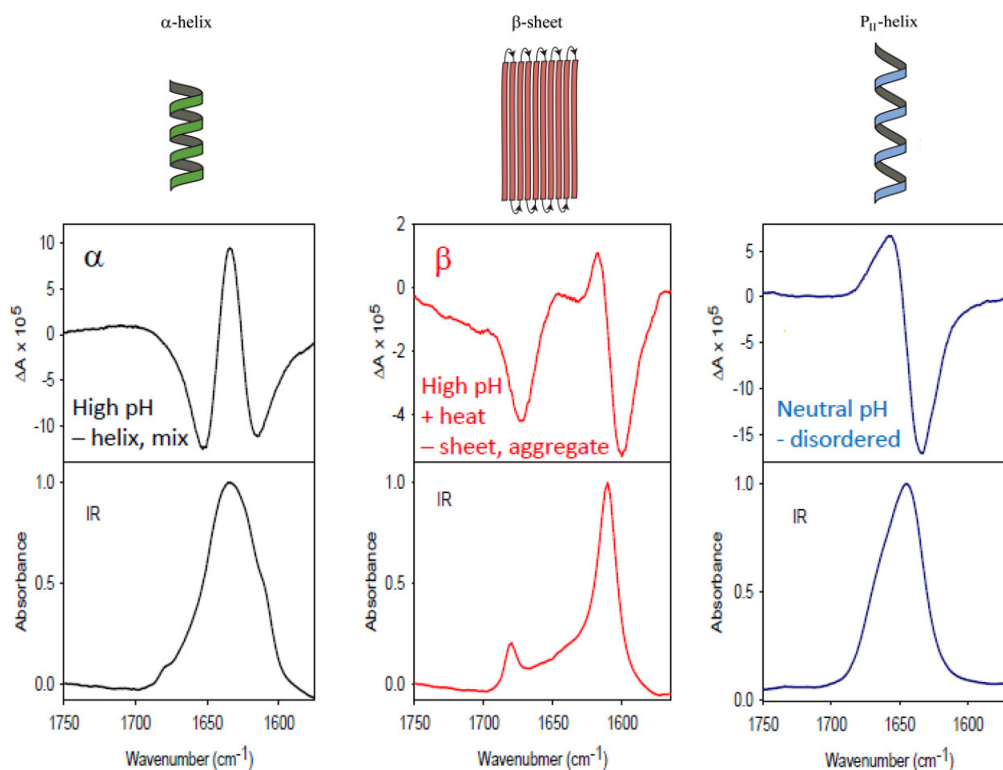


Figure 1.10: IR (lower frames) and VCD (upper frames) spectra in the amide I mode of poly-L-lysine region for 3 conformations of the polypeptide. This figure has been adapted from reference [23].

1.3.3. Theoretical VCD and the comparison to experimental VCD spectra

The second essential tool needed for the development of VCD was efficient theoretical models to interpret and assign the experimental results [52]. The VCD signal can change sign with a slight modification in conformation or non-covalent interactions, so VCD models need to be sufficiently accurate to be able to compare results reliably [53].

From the 1990s, the Stephens theory of VCD using density functional theory (DFT) [54] was implemented in the Gaussian suite of programs by Cheeseman and Frisch [21, Chapter 1]. More details about the DFT method will be given in section 2.3.

Example of the use of VCD spectroscopy to determine absolute molecular configurations

As an illustration of the state-of-the-art in the 2000s, the IR and VCD spectra determined using DFT for molecules of the pyrazole type, shown in Fig. 1.11 [55], are reproduced in Fig. 1.12. The pyrazole type molecules were separated into 2 diastereomers (**2a** and **2b**), for this example, the absolute configuration of diastereomer **2a** was examined.

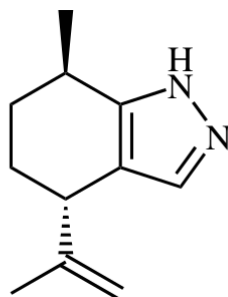


Figure 1.11: (4*S*,7*R*)-(-)-4-isopropylidene-7-methyl-4,5,6,7-tetrahydro-2(1*H*)-indazoles. This figure is adapted from reference [55].

The DFT calculations were performed, for various configurations, with the B3LYP functional and 6-31G(d) basis-set in the Gaussian 03 software. For the experimental spectrum, a solvent of CDCl₃ was used.

The configurations **2a-A** and **2a-B** are the lowest in energy and their spectra are the most similar to the experimental spectrum. It was shown that a mixture of 60% of **2a-A** and 40% of **2a-B** gives the VCD spectrum closest to the experimental spectrum.

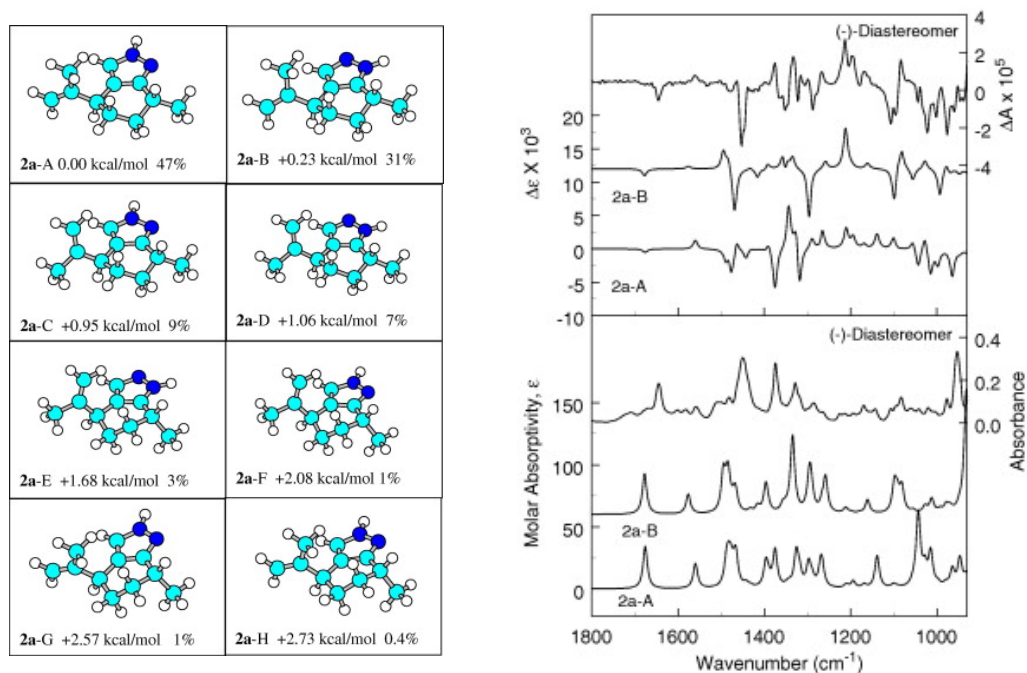


Figure 1.12: (a) Optimized conformers, relative energies and Boltzmann populations for pyrazole diastereomer 2a. (b) Comparison of VCD (upper frame) and observed IR (lower frame) spectra for (-)-diastereomer (right axes) with calculated spectra for diastereomer 2a conformers A and B. These figures have been reprinted from Tur et al. [55].

The study accomplished in the article used for this example did not include solvation for the theoretical part. However in VCD experiments, the molecules of interest are usually solvated and one fundamental issue is to assess the role of the solvent on the molecular conformation and on the spectrum itself. This role can depend on the nature of the solvent and whether it has a strong interaction with the solute. Successful assignment of VCD signals based on computed spectra therefore requires that the solvent is correctly described.

Solvation and theoretical VCD

VCD is known to have a high sensitivity to bulk solvent effects. This was demonstrated by Dezhahang et al. [56], who used a system of the TOLBINAP (2,2'-Bis(di-p-tolylphosphino)-1,1'-binaphthalene) palladium complex Pd(TOLBINAP)Cl₂ with the structure depicted in figure 1.13.

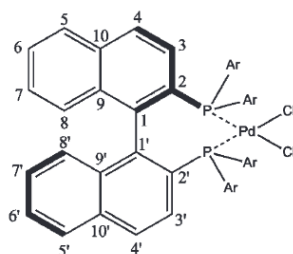


Figure 1.13: Structures of the TOLBINAP (2,2'-Bis(di-p-tolylphosphino)-1,1'-binaphthalene) (Ar=4-CH₃-C₆H₄) palladium complex, as taken from reference [56].

The goal of this study was to determine the conformational changes in the TOLBINAP ligands before and after their coordination with palladium [56]. The incorporation of an implicit polarisable continuum solvation model provided much better agreement between theory and experiment. This is shown in figure 1.14 where the spectra with the solvent described implicitly with a polarisable continuum (PCM) model show better agreement with the experimental spectrum. The regions showing noticeably different VCD patterns in the gas phase and in solution are highlighted with dashed-dotted lines.

In both the studies of Tur et al. [55] and Dezhahang et al [56], the calculations were performed assuming static configurations and without any explicit solvent. A way to include an explicit solvent is through the cluster-in-a-liquid model where the solute-solvent interactions are described explicitly through n solvent molecules that form a cluster with the solute. The surrounding bulk environment is accounted for with an implicit solvent [57]. To go beyond the static view, dynamic simulations such as molecular dynamics (MD) can be employed. MD is a powerful tool to sample the phase space of a molecular system at finite temperature in a realistic way [58] and more details about this technique will be given in chapter 2.

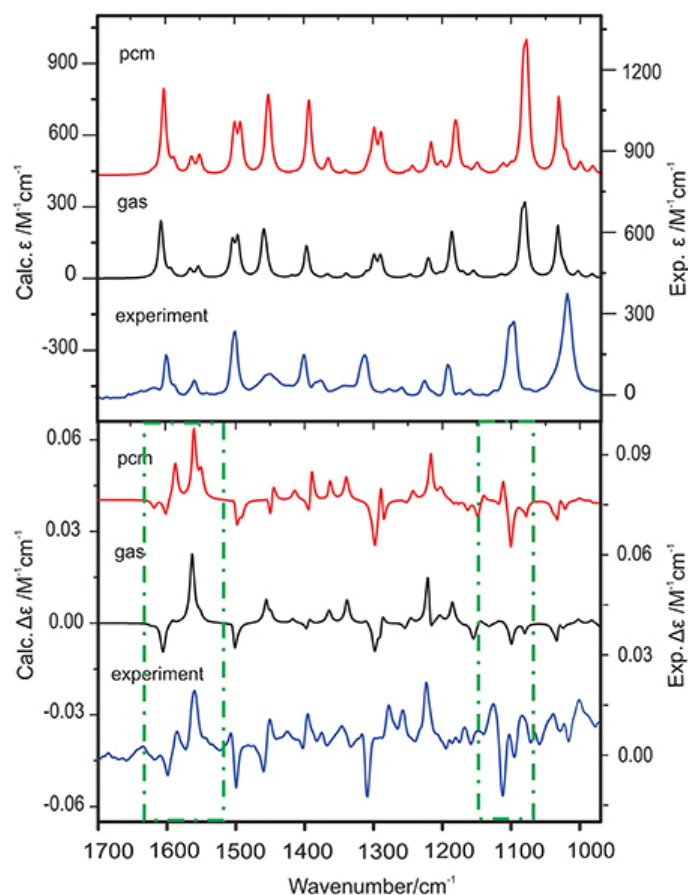


Figure 1.14: Comparison of the simulated IR and VCD spectra in the gas phase and with PCM solvent model for CDCl_3 with the corresponding experimental spectra of $\text{Pd}(\text{TOLBINAP})\text{Cl}_2$ in CDCl_3 , as reproduced from reference [56].

1.4. Thesis outline

After having introduced VCD spectroscopy in the previous sections, chapter 2 will illustrate various methods used to compute VCD spectra, pointing out the limitations of static approximations. This will bring us to the utility of the development of a method to obtain VCD spectroscopy through classical molecular dynamics based on force fields to include anharmonicities and finite temperature effects. It is this method that has been implemented during this PhD using the AMOEBA polarisable force field in the Tinker software package.

In chapter 3, a first illustrative case concerns *trans*-1-amino-2-indanol solvated in dimethyl sulfide which is examined using a combination of methods. This case demonstrates the importance of the sampling of the potential energy surface and of the ability to reach VCD spectra through classical molecular dynamics based on advanced force fields.

In chapter 4, this implementation using the AMOEBA polarisable force field is validated first with alanine in the gas phase, revealing the importance of extensive sampling of the system, which goes beyond the capacities of ab initio molecular dynamics. The various contributions of the electric and magnetic dipole moments are discussed, illustrating the importance of including electrostatics effects in the force field. Comparison with density-functional theory results along one MD trajectory confirms that electric and magnetic dipole moments and thus VCD spectra can be reliably evaluated using the polarisable force field. Information about the spectral assignment is not directly included in the molecular dynamics, thus a method to determine effective vibrational modes from the analysis of the MD trajectories is employed. In this method, the internal coordinates are linearly decomposed by projection onto suitable effective modes, the corresponding weights being obtained in such a way that the peak of each mode is as localized as much as possible. This effective mode analysis sheds light onto the various contributions of the different functional groups of the molecule to individual peaks in the vibrational spectrum.

In chapter 5, the force field combined to MD approach is extended and applied to alanine in crystalline and hydrated phases. The method is also validated with a comparison to DFT results and the effective mode analysis gives the vibrational modes corresponding to the peaks in the spectrum. Alanine solvated in dinitrogen is also examined, as this solvent binds more weakly to the alanine solute, the absence of hydrogen bonding between them limits the perturbation on the vibrational spectra. The final system studied is phenylcyclohexanediol in the crystalline phase and solvated in dimethyl sulfoxide. This relatively large molecule presents inter-molecular hydrogen bonding between molecules in the crystal and between the solvent and the solute in the solvated phase. The different environment of the hydroxyl groups are shown to have an impact on the spectra.

In chapter 6, the results are summed up in a conclusion and future work is detailed in the perspectives.

Bibliography

- [1] T. J. Leitereg, D. G. Guadagni, J. Harris, T. R. Mon, and R. Teranishi, "Chemical and sensory data supporting the difference between the odors of the enantiomeric carvones," *Journal of Agricultural and Food Chemistry*, vol. 19, no. 4, pp. 785–787, 1971.
- [2] F. Maderspacher, "Snail chirality: the unwinding," *Current Biology*, vol. 26, no. 5, pp. R215–R217, 2016.
- [3] J. R. Cronin and S. Pizzarello, "Enantiomeric excesses in meteoritic amino acids," *Science*, vol. 275, no. 5302, pp. 951–955, 1997.
- [4] R. Breslow and Z.-L. Cheng, "L-amino acids catalyze the formation of an excess of d-glyceraldehyde, and thus of other d sugars, under credible prebiotic conditions," *Proceedings of the National Academy of Sciences*, vol. 107, no. 13, pp. 5723–5725, 2010.
- [5] R. Dickerson, H. Drew, B. Conner, M. Kopka, and P. Pjura, "Helix geometry and hydration in A-DNA, B-DNA, and Z-DNA," in *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 47, pp. 13–24, Cold Spring Harbor Laboratory Press, 1983.
- [6] PowerSchool Learning, *Nucleic acids*, 2022. https://www.saddlespace.org/whittakerm/science/cms_page/view/7795295 (accessed July 7th, 2022).
- [7] N. Chhabra, M. L. Aseri, and D. Padmanabhan, "A review of drug isomerism and its significance," *International Journal of Applied and Basic Medical Research*, vol. 3, no. 1, p. 16, 2013.
- [8] I. A. Jaffe, K. Altman, P. Merryman, *et al.*, "The antipyridoxine effect of penicillamine in man," *The Journal of Clinical Investigation*, vol. 43, no. 10, pp. 1869–1873, 1964.
- [9] Inductiveload, *Electromagnetic Spectrum*, 2013. https://commons.wikimedia.org/wiki/File:EM_Spectrum_Properties_reflected.svg (accessed April 21st, 2022).
- [10] Moxfyre, *Raman energy levels*, 2009. https://commons.wikimedia.org/wiki/File:Raman_energy_levels.svg (accessed June 13th, 2022).
- [11] H. Scientific, *Raman principle*, 2022. <https://www.horiba.com/gbr/scientific/technologies/raman-imaging-and-spectroscopy/raman-spectroscopy/> (accessed April 21st, 2022).
- [12] A. Smekal, "Zur quantentheorie der dispersion," *Naturwissenschaften*, vol. 11, no. 43, pp. 873–875, 1923.
- [13] G. Landsberg and L. Mandelstam, "Über die lichtzerstreuung in kristallen," *Zeitschrift für Physik*, vol. 50, no. 11, pp. 769–780, 1928.
- [14] G. Landsberg and L. Mandelstam, "Sur des faits nouveaux relatifs à la diffusion de la lumière dans les cristaux," *CR Acad, Sci*, vol. 187, p. 109, 1928.
- [15] Edinburgh Instruments, *Optimisation of SERS for Glucose Sensing*, 2022. <https://www.edinst.com/optimisation-of-sers-for-glucose-sensing/> (accessed July 7th, 2022).

- [16] M. Faraday, "I. On the magnetic relations and characters of the metals," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 27, no. 177, pp. 1–3, 1845.
- [17] R. Crichton and R. Louro, *Practical approaches to biological inorganic chemistry*. Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands: Elsevier, 2019.
- [18] J. McMaster and V. S. Oganessian, "Magnetic circular dichroism spectroscopy as a probe of the structures of the metal sites in metalloproteins," *Current opinion in Structural Biology*, vol. 20, no. 5, pp. 615–622, 2010.
- [19] G. A. Garcia, L. Nahon, M. Lebech, J.-C. Houver, D. Doweck, and I. Powis, "Circular dichroism in the photoelectron angular distribution from randomly oriented enantiomers of camphor," *The Journal of Chemical Physics*, vol. 119, no. 17, pp. 8781–8784, 2003.
- [20] A. Cotton, "Absorption and dispersion of light," *Annales Chimie Physique*, vol. 8, no. 347, 1896.
- [21] P. J. Stephens, F. J. Devlin, and J. R. Cheeseman, *VCD spectroscopy for organic chemists*. 6000 Broken Sound Parkway NW, Suite 300 Boca Raton, FL 33487-2742: CRC Press, 2012.
- [22] N. Berova, L. Di Bari, and G. Pescitelli, "Application of electronic circular dichroism in configurational and conformational analysis of organic compounds," *Chemical Society Reviews*, vol. 36, no. 6, pp. 914–931, 2007.
- [23] T. A. Keiderling, "Structure of condensed phase peptides: insights from vibrational circular dichroism and raman optical activity techniques," *Chemical reviews*, vol. 120, no. 7, pp. 3381–3419, 2020.
- [24] Jasco, *Spectra of alpha-pinene*, 2022. <https://jascoinc.com/applications/measurement-of-vibrational-circular-dichroism-spectra-using-the-fvs-6000/> (accessed April 25th, 2022).
- [25] L. A. Nafie, *Vibrational optical activity: principles and applications*. The Atrium, Southern Gate, Chichester, West Sussex: John Wiley & Sons, 2011.
- [26] S. R. Bhattacharya and T. Bürgi, "Amplified vibrational circular dichroism as a manifestation of the interaction between a water soluble gold nanocluster and cobalt salt," *Nanoscale*, vol. 11, no. 48, pp. 23226–23233, 2019.
- [27] T. P. Golub and C. Merten, "Stereochemistry of the reaction intermediates of prolinol ether catalyzed reactions characterized by vibrational circular dichroism spectroscopy," *Chemistry – A European Journal*, vol. 26, no. 11, pp. 2349–2353, 2020.
- [28] A. Pérez-Mellor, K. L. Barbu-Debus, and A. Zehnacker, "Solid-state synthesis of cyclo LD-diphenylalanine: A chiral phase built from achiral subunits," *Chirality*, vol. 32, no. 5, pp. 693–703, 2020.

- [29] A. Pérez-Mellor and A. Zehnacker, "Vibrational circular dichroism of a 2,5-diketopiperazine (DKP) peptide: Evidence for dimer formation in cyclo LL or LD diphenylalanine in the solid state," *Chirality*, vol. 29, no. 2, pp. 89–96, 2017.
- [30] V. Declerck, A. Pérez-Mellor, R. Guillot, D. J. Aitken, M. Mons, and A. Zehnacker, "Vibrational circular dichroism as a probe of solid-state organisation of derivatives of cyclic β -amino acids: *Cis*- and *trans*-2-aminocyclobutane-1-carboxylic acid," *Chirality*, vol. 31, no. 8, pp. 547–560, 2019.
- [31] S. Ma, X. Cao, M. Mak, A. Sadik, C. Walkner, T. B. Freedman, I. K. Lednev, R. K. Dukor, and L. A. Nafie, "Vibrational circular dichroism shows unusual sensitivity to protein fibril formation and development in solution," *Journal of the American Chemical Society*, vol. 129, no. 41, pp. 12364–12365, 2007.
- [32] C. Merten, "Vibrational optical activity as probe for intermolecular interactions," *Physical Chemistry Chemical Physics*, vol. 19, no. 29, pp. 18803–18812, 2017.
- [33] J. Sadlej, J. C. Dobrowolski, and J. E. Rode, "VCD spectroscopy as a novel probe for chirality transfer in molecular interactions," *Chemical Society Reviews*, vol. 39, no. 5, pp. 1478–1488, 2010.
- [34] C. L. Covington, V. P. Nicu, and P. L. Polavarapu, "Determination of the absolute configurations using exciton chirality method for vibrational circular dichroism: Right answers for the wrong reasons?," *The Journal of Physical Chemistry A*, vol. 119, no. 42, pp. 10589–10601, 2015.
- [35] V. P. Nicu and E. J. Baerends, "Effects of complex formation on vibrational circular dichroism spectra," *The Journal of Physical Chemistry A*, vol. 112, no. 30, pp. 6978–6991, 2008.
- [36] T. B. Freedman, X. Cao, R. K. Dukor, and L. A. Nafie, "Absolute configuration determination of chiral molecules in the solution state using vibrational circular dichroism," *Chirality*, vol. 15, no. 9, pp. 743–758, 2003.
- [37] A. Bouchet, T. Brotin, M. Linares, H. Ågren, D. Cavagnat, and T. Buffeteau, "Enantioselective complexation of chiral propylene oxide by an enantiopure water-soluble cryptophane," *The Journal of Organic Chemistry*, vol. 76, no. 10, pp. 4178–4181, 2011.
- [38] T. Brotin, D. Cavagnat, and T. Buffeteau, "Conformational changes in cryptophane having C_1 -symmetry studied by vibrational circular dichroism," *The Journal of Physical Chemistry A*, vol. 112, no. 36, pp. 8464–8470, 2008.
- [39] T. Buffeteau, D. Pitrat, N. Daugey, N. Calin, M. Jean, N. Vanthuyne, L. Ducasse, F. Wien, and T. Brotin, "Chiroptical properties of cryptophane-111," *Physical Chemistry Chemical Physics*, vol. 19, no. 28, pp. 18303–18310, 2017.
- [40] Y. He, W. Bo, R. K. Dukor, and L. A. Nafie, "Determination of absolute configuration of chiral molecules using vibrational optical activity: A review," *Applied Spectroscopy*, vol. 65, no. 7, pp. 699–723, 2011.

- [41] T. Bürgi and A. Baiker, "Conformational behavior of cinchonidine in different solvents: a combined NMR and ab initio investigation," *Journal of the American Chemical Society*, vol. 120, no. 49, pp. 12920–12926, 1998.
- [42] C. Merten and Y. Xu, "Chirality transfer in a methyl lactate-ammonia complex observed by matrix-isolation vibrational circular dichroism spectroscopy," *Angewandte Chemie International Edition*, vol. 52, no. 7, pp. 2073–2076, 2013.
- [43] C. W. Deutsche and A. Moscowitz, "Optical Activity of Vibrational Origin. I. A Model Helical Polymer," *The Journal of Chemical Physics*, vol. 49, no. 7, pp. 3257–3272, 1968.
- [44] C. W. Deutsche and A. Moscowitz, "Optical Activity of Vibrational Origin. II. Consequences of Polymer Conformation," *The Journal of Chemical Physics*, vol. 53, no. 7, pp. 2630–2644, 1970.
- [45] G. Holzwarth and I. Chabay, "Optical Activity of Vibrational Transitions: A Coupled Oscillator Model," *The Journal of Chemical Physics*, vol. 57, no. 4, pp. 1632–1635, 1972.
- [46] J. A. Schellman, "Vibrational optical activity," *The Journal of Chemical Physics*, vol. 58, no. 7, pp. 2882–2886, 1973.
- [47] T. R. Faulkner, A. Moscowitz, G. Holzwarth, E. C. Hsu, and H. S. Mosher, "Infrared circular dichroism of carbon-hydrogen and carbon-deuterium stretching modes. calculations," *Journal of the American Chemical Society*, vol. 96, no. 1, pp. 252–253, 1974.
- [48] G. Osborne, J. Cheng, and P. Stephens, "A near-infrared circular dichroism and magnetic circular dichroism instrument," *Review of Scientific Instruments*, vol. 44, no. 1, pp. 10–15, 1973.
- [49] I. Chabay and G. Holzwarth, "Infrared circular dichroism and linear dichroism spectrophotometer," *Applied Optics*, vol. 14, no. 2, pp. 454–459, 1975.
- [50] G. Holzwarth, E. C. Hsu, H. S. Mosher, T. R. Faulkner, and A. Moscowitz, "Infrared circular dichroism of carbon-hydrogen and carbon-deuterium stretching modes. observations," *Journal of the American Chemical Society*, vol. 96, no. 1, pp. 251–252, 1974.
- [51] L. Nafie, J. Cheng, and P. Stephens, "Vibrational circular dichroism of 2, 2, 2-trifluoro-1-phenylethanol," *Journal of the American Chemical Society*, vol. 97, no. 13, pp. 3842–3843, 1975.
- [52] L. A. Nafie, "Vibrational optical activity: From discovery and development to future challenges," *Chirality*, vol. 32, no. 5, pp. 667–692, 2020.
- [53] K. Le Barbu-Debus, J. Bowles, S. Jähnigen, C. Clavaguéra, F. Calvo, R. Vuilleumier, and A. Zehnacker, "Assessing cluster models of solvation for the description of vibrational circular dichroism spectra: synergy between static and dynamic approaches," *Physical Chemistry Chemical Physics*, vol. 22, no. 45, pp. 26047–26068, 2020.
- [54] P. J. Stephens, "Theory of vibrational circular dichroism," *The Journal of Physical Chemistry*, vol. 89, no. 5, pp. 748–752, 1985.

- [55] E. Tur, G. Vives, G. Rapenne, J. Crassous, N. Vanthuyne, C. Roussel, R. Lombardi, T. Freedman, and L. Nafie, "HPLC separation and VCD spectroscopy of chiral pyrazoles derived from (5R)-dihydrocarvone," *Tetrahedron: Asymmetry*, vol. 18, no. 16, pp. 1911–1917, 2007.
- [56] Z. Dezhahang, C. Merten, M. R. Poopari, and Y. Xu, "Vibrational circular dichroism spectroscopy of two chiral binaphthyl diphosphine ligands and their palladium complexes in solution," *Dalton Transactions*, vol. 41, no. 35, pp. 10817–10824, 2012.
- [57] A. S. Perera, J. Thomas, M. R. Poopari, and Y. Xu, "The clusters-in-a-liquid approach for solvation: new insights from the conformer specific gas phase spectroscopy and vibrational optical activity spectroscopy," *Frontiers in Chemistry*, vol. 4, p. 9, 2016.
- [58] S. Abbate, G. Longhi, K. Kwon, and A. Moscovitz, "The use of cross-correlation functions in the analysis of circular dichroism spectra," *The Journal of Chemical Physics*, vol. 108, no. 1, pp. 50–62, 1998.

Chapter 2

Methodology

2.1. General theory of vibrational circular dichroism

Circular dichroism is the difference in absorption of left- and right- circularly polarised light. It originates from molecular interactions with the electric and magnetic fields of the light, which causes an oscillation of the electrons. These oscillations are in phase with the light's fields for one circular polarisation and out of phase for the other [1].

The interaction Hamiltonian, which expresses the molecule's energy in the light's electromagnetic field, is

$$H = -\vec{\mu} \cdot \vec{E} - \vec{m} \cdot \vec{B} \quad (2.1)$$

where $\vec{\mu}$ is the molecule's electric dipole moment and \vec{m} is the molecule's magnetic dipole moment. \vec{E} is the externally applied electric field and \vec{B} is the externally applied magnetic field. The electric field (E-field) and the magnetic field (B-field) of linear polarised light, propagating toward the positive x-axis vary as:

$$\vec{E}(t) = E_0 \hat{e} e^{i(\kappa x - \omega t)} \quad \vec{B}(t) = B_0 \hat{e} e^{i(\kappa x - \omega t)} \quad (2.2)$$

where E_0 and B_0 are the wave amplitudes, \hat{e} is a polarisation vector with unit magnitude, κ is the wavenumber, related to the wavelength λ through $\kappa = 2\pi/\lambda$, and ω is the temporal frequency [2]. These fields are represented in red for the E-field and in blue for the B-field in figure 2.1A.

For circularly polarised (CP) light, the E-field rotates around the propagation axis. For right-circularly polarised (RCP) light, the E-field rotates anti-clockwise as the viewer looks down the propagation axis toward the origin, as shown in figure 2.1B. For left-circularly polarised (LCP) light the rotation is clockwise. This convention is reversed if the viewer looks up the propagation axis toward the detector as in 2.1C and D. These rotations can be described by adapting the polarisation vector \hat{e} . Therefore the electric field of right- and left- circularly polarised light propagating toward the x-axis can be expressed as:

$$\vec{E}(t)_{\text{RCP}} = \frac{E_0}{\sqrt{2}} (\hat{y} - i\hat{z}) e^{i(\kappa x - \omega t)}, \quad \vec{E}(t)_{\text{LCP}} = \frac{E_0}{\sqrt{2}} (\hat{y} + i\hat{z}) e^{i(\kappa x - \omega t)} \quad (2.3)$$

where \hat{y} and \hat{z} are unit vectors on the y-axis and z-axis, respectively.

For CP light, the B-field rotates 90° out of phase, behind for RCP and ahead for LCP as depicted in figure 2.1C and D. The B-field amplitude is equal to $\frac{E_0}{c}$, thus the magnetic field for left- and right- circularly polarised light can be written:

$$\vec{B}(t)_{\text{RCP}} = \frac{i}{c} \vec{E}(t)_{\text{RCP}}, \quad \vec{B}(t)_{\text{LCP}} = -\frac{i}{c} \vec{E}(t)_{\text{LCP}} \quad (2.4)$$

where c is the velocity of light.

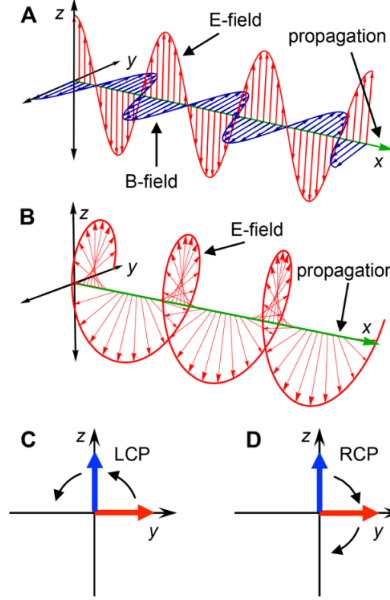


Figure 2.1: Light polarisation, shown with the E-field in red and B-field in blue. (A) Vertically polarised light with polarisation vector $\hat{e} = \hat{z}$; (B) E-field of right-circularly polarised light (RCP) with the B-field not shown for clarity; (C) fields at the origin for LCP, facing backward along the propagation direction; (D) fields at the origin for RCP. This figure has been reprinted from reference [1].

The molecular extinction coefficient (ε) conveys how strongly a molecule absorbs light at a given wavelength. This extinction coefficient is constructed from Fermi's golden rule which gives the transition probability between an initial state $|i\rangle$ and a final state $|f\rangle$ [3]:

$$\varepsilon = \frac{\pi N_A \omega \delta(\omega - \omega_{fi})}{\varepsilon_0 \hbar c \log 10} \frac{1}{E_0^2} |\langle f | H | i \rangle|^2 \quad (2.5)$$

where N_A is Avogadro's number, ε_0 is the electric permittivity of free space, $\hbar = \frac{h}{2\pi}$ with h being Planck's constant and $\delta(\omega - \omega_{fi})$ a Dirac δ function.

The Hamiltonian can be adopted using the previous equations for the electric and magnetic dipole moments. Although these equations come from classical physics, they also apply in quantum mechanics via the replacement of the position and velocity vectors with their quantum operators. These operators can be written as:

$$\vec{\mu} = \begin{bmatrix} \vec{\mu}_{ii} & \vec{\mu}_{fi} \\ \vec{\mu}_{if} & \vec{\mu}_{ff} \end{bmatrix} \quad \vec{m} = \begin{bmatrix} \vec{m}_{ii} & \vec{m}_{fi} \\ \vec{m}_{if} & \vec{m}_{ff} \end{bmatrix} \quad (2.6)$$

where each matrix element represents a bracket. In the electric dipole matrix, $\vec{\mu}_{ii} = \langle i | \vec{\mu} | i \rangle$ is the static electric dipole moment of the molecule in its initial state, and $\vec{\mu}_{ff} =$

$\langle f | \vec{\mu} | f \rangle$, is the static electric dipole moment of the molecule in its final state. The off-diagonal elements, $\vec{\mu}_{fi}$ and $\vec{\mu}_{if}$, are the electric transition dipole moments, which enable coupling between the states.

After including the equations for electric and magnetic fields and the electric and magnetic dipole moment operators in equation (2.5) and following the rotational averaging procedure described by Andrews [4], the following expression for the extinction coefficient for circular dichroism is found:

$$\epsilon_{CD} = \frac{4\pi N_A \omega \delta(\omega - \omega_{fi})}{3\epsilon_0 \hbar c^2 \log 10} \Im(\vec{\mu}_{fi} \cdot \vec{m}_{if}). \quad (2.7)$$

This expression corresponds to the Rosenfeld equation [5]. Different versions of the prefactor exist depending on the units chosen [6, 7, 8]. The magnetic dipole moment is taken into account for the VCD representation so we can consider the rotation around the axes of propagation, characteristic of circularly polarised light. The corresponding equation of the infrared (IR) signal [9, 10] is:

$$\epsilon_{IR} = \frac{4\pi N_A \omega \delta(\omega - \omega_{fi})}{3\epsilon_0 \hbar c^2 \log 10} (\vec{\mu}_{fi} \cdot \vec{\mu}_{if}) \quad (2.8)$$

where $(\vec{\mu}_{fi} \cdot \vec{\mu}_{if})$ corresponds to the dipole strength [11].

2.2. Harmonic approach

A transition from an initial to a final state occurs due to the absorption or emission of light. A transition has an associated frequency depending on the energy of the light absorbed or emitted ΔE , given by Planck's relation:

$$\Delta E = h\omega_{fi} \quad (2.9)$$

With light in the IR range, this transition corresponds to a vibrational transition, from one vibrational energy level to another, as shown in figure 2.2.

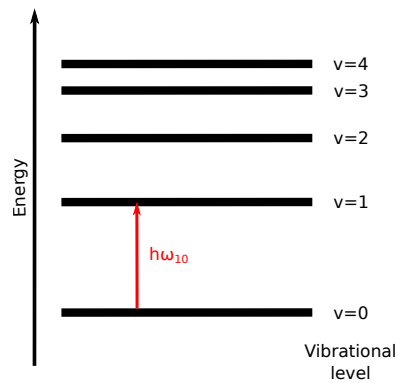


Figure 2.2: Energy level diagram of vibrational transitions.

The energy of a molecular system is the sum of its potential (V) and kinetic (T) energies. For molecular conformations that reside near a local minimum of the potential energy surface, a Taylor expansion can be used to approximate the potential energy as:

$$V = V_0 + \sum_{j=1}^{3N} \left(\frac{\partial V}{\partial p_j} \right)_0 p_j + \frac{1}{2} \sum_{j,k}^{3N} \left(\frac{\partial^2 V}{\partial p_j \partial p_k} \right)_0 p_j p_k + \dots \quad (2.10)$$

where p_j are the mass-weighted coordinates (MWC) of atom j . MWCs are given relative to the equilibrium position $(x_{j,0}, y_{j,0}, z_{j,0})$ and are multiplied by the square root of the mass of the corresponding atom: $p_1 = \sqrt{m_j}(x_j - x_{j,0})$, $p_2 = \sqrt{m_j}(y_j - y_{j,0})$, $p_3 = \sqrt{m_j}(z_j - z_{j,0})$, ... As we consider an equilibrium geometry, the geometry corresponds to a minimum of the potential energy surface, where the gradient term vanishes. The consideration of only equilibrium geometries is indicated via $(\)_0$. The harmonic approximation ignores the higher terms of the expansion. Thus the potential is reduced to:

$$V_{\text{harmonic}} = V_0 + \frac{1}{2} \sum_{j,k}^{3N} \left(\frac{\partial^2 V}{\partial p_j \partial p_k} \right)_0 p_j p_k = V_0 + \frac{1}{2} \sum_{j,k}^{3N} f_{ij}^{\text{MWC}} p_j p_k \quad (2.11)$$

where $f_{ij}^{\text{MWC}} = f_{ji}^{\text{MWC}}$ is the ij component of the mass-weighted Hessian denoted H^{MWC} .

To obtain IR and VCD spectra, the vibrational frequencies are computed by determining the second partial derivatives of the potential energy V with respect to displacement of the atoms in MWC, at the equilibrium geometry. This corresponds to the ij component of the mass-weighted Hessian, f_{ij}^{MWC} , mentioned in equation (2.11):

$$f_{ij}^{\text{MWC}} = \left(\frac{\partial^2 V}{\partial p_i \partial p_j} \right)_0 \quad (2.12)$$

The mass-weighted Hessian matrix, also called the dynamical matrix, is diagonalised to obtain the $3N$ eigenvalues corresponding to the harmonic frequencies of the system and the $3N$ eigenvectors corresponding to the atomic displacements associated with each frequency. These eigenvectors are named normal modes and correspond to linear combinations of atomic displacements for which the dynamical matrix is diagonal. The normal modes corresponding to global rotations and translations of the molecule are associated with vanishing eigenvalues. Removing these modes leaves $N_{\text{vib}} = 3N - 6$ independent normal modes, or $3N - 5$ if the molecule is linear.

The IR intensities arise from the dipole strength D_i and VCD intensities come from the rotational strength R_i [12]. As shown in equations (2.8) and (2.7), the vibrational intensities arise from the determination of the electric ($\vec{\mu}_{\text{if}}$) and magnetic transition dipole moments (\vec{m}_{if}) [13].

A second harmonic approximation is also taken where fundamental properties describing the spectroscopic intensities, in our case the dipole moment surfaces, are treated as harmonic [14].

2.3. Density-functional theory

2.3.1. Basic principles of DFT

To describe an atomic or molecular system and to determine its properties, various computational modelling methods can be used. The first one explored here is density-functional theory (DFT).

Density-functional theory is based on the first and second Hohenberg and Kohn theorems [15]. The first theorem declares that all physical properties of a molecular system can be obtained from its electronic density $\rho(\vec{r})$. Thus the energy is a density functional that depends only on the three spatial coordinates :

$$E[\rho] = T[\rho] + V_{ee}[\rho] + V_{Ne}[\rho] \quad (2.13)$$

where $T[\rho]$ is the kinetic energy of the electrons, $V_{ee}[\rho]$ is the electron-electron repulsion energy and $V_{Ne}[\rho]$ the electron-nuclei attraction energy. This dependence of the energy on the electronic density differs from the Hartree-Fock (HF) method, where the energy is a function of the wavefunction that depends on $3N$ space coordinates and N spin coordinates (for N electrons).

The second Hohenberg and Kohn theorem states that the variational principle can be applied. Thus the ground-state energy, E_0 , is always less than or equal to the expected value of the energy E :

$$E[\rho] \geq E[\rho_0] = E_0. \quad (2.14)$$

However the exact analytical expression of $T[\rho]$ and $V_{ee}[\rho]$ is not known. To attempt to access the energy, the real system of interacting electrons is substituted by a fictitious system of independent electrons having the same density, following the Kohn-Sham approach [16]. One-electron wavefunctions χ_n are introduced and obtained by solving the Kohn-Sham equation, where v_{eff} is an effective potential:

$$\left[\frac{-1}{2} \nabla^2 + v_{\text{eff}}(r) \right] \chi_n = \varepsilon_n \chi_n. \quad (2.15)$$

Hence, the corresponding non-interacting Hamiltonian does not contain the explicit electron-electron repulsion term but rather an effective potential. This equation is solved iteratively until self-consistency is achieved.

The effective potential contains an exchange-correlation potential, V_{XC} . The exchange-correlation can be decomposed into the exchange term V_X , which represents the interaction of electrons with the same spin, and into the correlation term V_C , which represents the interaction of electrons with opposite spins. The exchange-correlation energy functional corresponds to the integration of the exchange-correlation potential over the electronic density:

$$E_{XC}[\rho] = \int \rho(\vec{r}) V_{XC} d\vec{r}. \quad (2.16)$$

Unfortunately, the exact exchange-correlation functional is not known either. Thus, approximate exchange-correlation functionals are employed, leading to approximate densities and energies. Over the last decades, many exchange-correlation functionals have been proposed and tested [17, 18, 19, 20, 21, 22, 23, 24]. They are usually distinguished as the rungs of a ladder, going from basic/simple expressions to improved ones [25], as shown in figure 2.3.

The family of least accurate exchange-correlation functionals are functionals that use the local density approximation (LDA), they depend only on the electronic density. Then there are the functionals that employ the generalized gradient approximation (GGA) that depend on the density and its gradient. If the functional also depends on the Laplacian of the density, it is a meta-GGA. The functionals with the highest accuracy exploit exact exchange where the functional can be expressed in terms of occupied orbitals.

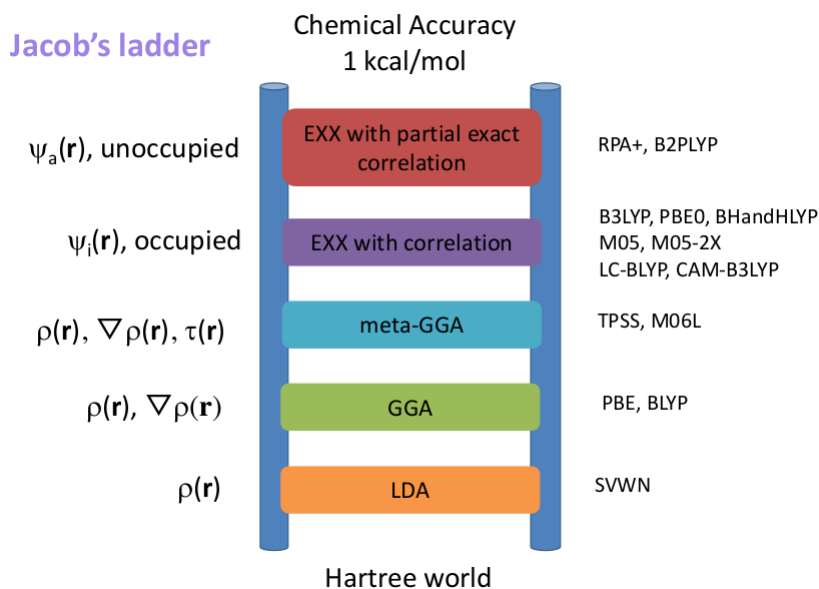


Figure 2.3: Jacob's ladder of density functional approximations. The figure has been adapted from reference [25].

During this thesis, if nothing is mentioned, the functionals B3LYP and the basis-set 6-311++G(d,p) were used. The comparison of the quality of mid-IR spectra of 4-methyl-2-oxetanone predicted using various density functionals demonstrates the quality of the B3LYP functional [26], though this reference is dated. More recent work on vibrational spectroscopy also used this functional [27, 28].

For the solvated systems, the D3BJ correction was added to the B3LYP functional to include dispersion effects. The term D3BJ corresponds to the D3 version of Grimme's dispersion with Becke-Johnson (BJ) damping [29]. Grimme's dispersion is an atom-pair wise potential for the treatment of the dispersion energy which helps to reach the so-called chemical accuracy, currently assumed to be 1 kcal/mol for the energies [30, 31]. Including

the BJ damping shows better results for nonbonded distances, which is important when considering the interactions between the solute and the solvent.

2.3.2. Application to infrared and vibrational circular dichroism spectroscopies

An early application of density-functional theory to obtain IR and VCD intensities has been reported by Cheeseman et al. [32]. The vibrational dipole and rotational strengths of 2,3-trans-d₂-oxirane were determined within the harmonic approximation and compared to experimental values. The structure of the molecule is shown in figure 2.4. The experimental study was performed in different solutions depending on the mode under scrutiny, namely, a C₂Cl₄ solution was used for the C–H and C–D stretching modes and a CS₂ solution otherwise.



Figure 2.4: Structure of 2,3-trans-d₂-oxirane.

The calculated rotational strengths using HF and DFT employing the B3LYP functional are given against the experimental values in figure 2.5. The mean absolute deviation from experiment is reduced from 10.8 ($\times 10^{-44}$ esu² cm²) for the HF calculations to 8.4 for the DFT calculations.

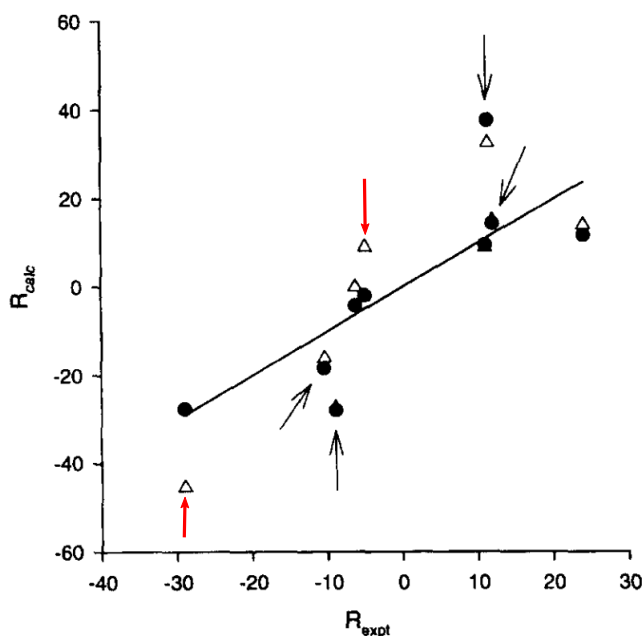


Figure 2.5: Comparison of calculated, Hartree Fock (Δ) and DFT (\bullet) rotational strengths (R_{calc}) of 2,3-trans-d₂-oxirane to experimental values (R_{expt}). Experimental rotational strengths are from solution spectra. Arrows indicate C–H and C–D stretching modes. The line is of unit slope. Adapted from Cheeseman et al. [32].

The two modes shown with a red arrow are the main contributors to the reduction of the mean absolute deviation from experiment, as their correlation with experiment is excellent for DFT. The arrows point to the HF values and the corresponding DFT point are on the same vertical line. These modes are assigned experimentally to hydrogenic bending modes [33]. In contrast, the DFT agreement with experimental results is worse for 3 of the 4 hydrogenic stretching modes, shown with black arrows. These modes hinder the reduction of the mean absolute deviation from experiment. Ignoring these modes, the mean absolute deviation drops from 9.7 for HF to 4.0 for the DFT calculations.

By comparing frequencies and dipole strengths computed at different levels of theory and comparing theory and experiment for other achiral isotopomers of oxirane, the authors also deduced that an important source of error of the calculated results compared to experimental results is the neglect of anharmonicity. Thus, in the next section, anharmonicity will be discussed.

2.4. Anharmonicity effects

Within the harmonic approximation, the potential energy between 2 atoms can be described as the following potential:

$$V_{\text{harmonic}}^{N=2} = \frac{1}{2}kr^2, \quad \text{with } k = \frac{\partial^2 V}{\partial r^2} \quad (2.17)$$

and $r^2 = p_1 p_2$, with p_i the mass-weighted coordinates of atom i [34].

A more accurate model involves the use of the Morse potential, which is described with the following equation [35]:

$$V_{\text{Morse}}^{N=2} = D_e \left(1 - e^{-a(r-r_e)} \right)^2 \quad (2.18)$$

Here r is the distance between the atoms, r_e is the equilibrium bond distance, D_e is dissociation energy, and a controls the width of the potential, the smaller a is, the larger the well.

The harmonic and Morse potentials are compared to each other, together with their vibrational levels, in figure 2.6. The energy spacing between vibrational levels and thus the frequency, decreases with increasing energy of the vibrations. The harmonic and Morse potentials are close to each other only for the ground vibrational level ($v = 0$). Also, the further the value of the internuclear separation r is from the value of equilibration r_e , the more the two potentials differ. Hence the harmonic approximation is only valid for small amplitudes of vibration.

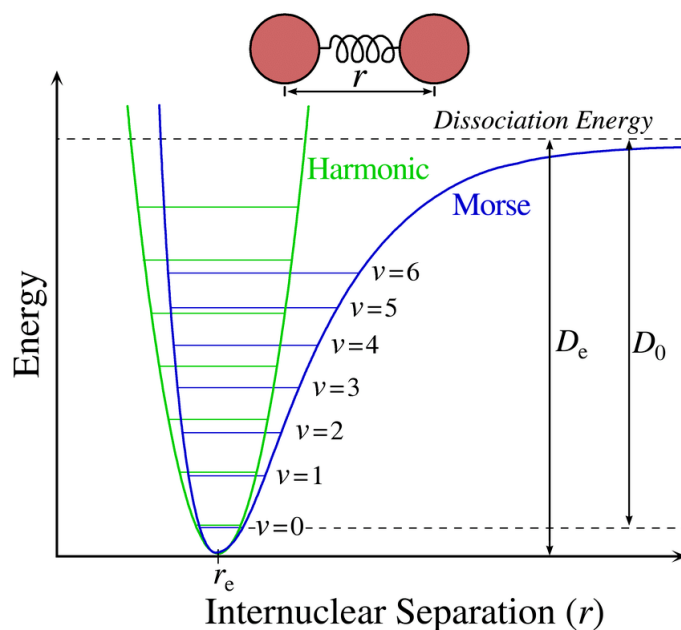


Figure 2.6: Comparison of the Morse and the harmonic oscillator potentials. This figure has been reprinted from reference [36].

For the value of the vibration levels higher than 1, there is a shift between the corresponding energies for the Morse and harmonic values. This leads to an overestimation of the harmonic frequencies which is also found in larger systems where the harmonic frequencies are higher compared with experimental results. A first solution to go beyond the harmonic approximation is to apply a scaling factor that shifts the values of the frequencies to correct for anharmonicities. Scaling factors also empirically correct for the various deficiencies of the underlying quantum chemical methods, including basis set incompleteness. Their ideal values have been tabulated for many methods. [37, 38, 39, 40, 41]. Depending on the functional and basis set used, these scaling factors range between 0.8 and 0.99, for common organic molecules. For the functional B3LYP and the basis-set 6-311++G(d,p), a scaling factor of 0.98 was used in accordance with similar theoretical methods [28].

The potential surfaces of fluxional molecules, namely molecules that have multiple conformations, are composed of multiple wells that can hardly be described with a single harmonic potential. Also, at sufficiently high temperature, the system has enough energy to access higher vibrational modes which are misrepresented with the harmonic potential, as shown in figure 2.6.

In a second step, approaches that improve over the harmonic approximation are discussed. Two complementary methods will be considered, namely second-order perturbation theory [42] and the so-called cluster-in-a-liquid approach [43].

2.4.1. Perturbation approaches

Within second-order vibrational perturbation theory (VPT2), the potential energy can be expanded up to the next non-vanishing term (order 4) as:

$$V_{\text{anharm}} = \frac{1}{2} \sum_{j,k}^{3N} \left(\frac{\partial^2 V}{\partial p_j \partial p_k} \right)_0 p_j p_k + \frac{1}{6} \sum_{j,k,l}^{3N} \left(\frac{\partial^3 V}{\partial p_j \partial p_k \partial p_l} \right)_0 p_j p_k p_l + \frac{1}{24} \sum_{j,k,l,m}^{3N} \left(\frac{\partial^4 V}{\partial p_j \partial p_k \partial p_l \partial p_m} \right)_0 p_j p_k p_l p_m + \dots \quad (2.19)$$

This equation can be reduced to :

$$V_{\text{anharm}} = \frac{1}{2} \sum_{j,k}^{3N} f_{ij}^{\text{MWC}} p_j p_k + \frac{1}{6} \sum_{j,k,l}^{3N} k_{jkl} p_j p_k p_l + \frac{1}{24} \sum_{j,k,l,m}^{3N} k_{jklm} p_j p_k p_l p_m + \dots \quad (2.20)$$

where k_{jkl} and k_{jklm} are the anharmonic force constants.

The idea of VPT2 is to solve the vibrational Hamiltonian within a 2nd order perturbation approach. The vibrational energy [44] of mode n out of a total of M modes, about an equilibrium geometry, is found as a so-called Dunham expansion expressed as:

$$E_n = E_n^{(0)} + \hbar \sum_i^M \omega_i (n_i + \frac{1}{2}) + \hbar \sum_{i \leq j}^M \chi_{ij} (n_i + \frac{1}{2})(n_j + \frac{1}{2}) \quad (2.21)$$

where the ω_i 's are the harmonic wavenumbers and the χ_{ij} 's are the anharmonic contributions that can be given as a function of the anharmonic force constants [45].

The fundamental bands ν_i , overtones $[2\nu_i]$, and combination bands $[\nu_i \nu_j]$ are then expressed by:

$$\nu_i = \omega_i + 2\chi_{ii} + \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^N \chi_{ij} \quad (2.22)$$

$$[2\nu_i] = 2\omega_i + 6\chi_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^N \chi_{ij} = 2\nu_i + 2\chi_{ii} \quad (2.23)$$

$$[\nu_i \nu_j] = \omega_i + \omega_j + 2\chi_{ii} + 2\chi_{jj} + \sum_{\substack{k=1 \\ j \neq j, i}}^N (\chi_{ik} + \chi_{jk}) \quad (2.24)$$

Foldes et al. used VPT2 to show that hydrogen bonding has a profound effect on anharmonicity [46]. For example, the formation of the hydrogen bond induces a red-shift in the OH and NH vibrations.

2.4.2. Boltzmann weights and the cluster-in-a-liquid model

The VPT2 method generally improves the positions of vibrational bands, however, it still remains unable to account for fluxionality. In this section, the anharmonicity due to the fluxional character is accounted for through a large number of conformations of the system. Each conformation represents a well of the potential energy surface. For each one, a spectrum and a Boltzmann weight, depending on the energy and temperature of the configuration, can be determined. Thus an average spectrum can be obtained by multiplying each individual harmonic spectrum by its Boltzmann weight and summing all of the contributions. In the examples given in this work, the properties of each conformation are determined using DFT.

Moreover, hydrogen bonding is poorly described with second-order perturbation theory. To consider hydrogen bonds and solvation, Perera et al. [43] introduced the use of a cluster-in-a-liquid model where the hydrogen bonds are described explicitly through solute-(solvent)_n clusters. Thus the explicit system includes a solute molecule surrounded by *n* solvent molecules. The surrounding bulk environment is accounted for with an implicit solvent. Including the whole solvent as a explicit bulk would be too expensive for DFT calculations.

Continuum models of solvation represent the solvent as a continuous medium, with the same dielectric constant as the pure solvent, surrounding the solute, instead of individual explicit solvent molecules [47]. In the polarisable continuum model (PCM), the effective solute-solvent interaction potential is determined from the dielectric medium polarised by the charge distribution of the molecular solute [48]. An integral equation formalism variant of the implicit polarisable continuum model (IEFPCM) also exists [49] and is the current version of PCM applied in common quantum chemistry packages.

Merten et al. [50] demonstrated the importance of hydrogen bonding for VCD in the case of methylbenzyl-amine-borane in CDCl₃. In this molecule there is a special kind of hydrogen bond, named dihydrogen bond (DHB): N-H^{δ+}⋯^{δ-}H-B. This bond was identified to be responsible for the unusual stability of the system [51] and is shown by the dotted lines in the left of figure 2.7.

For the VCD spectra, a better agreement is achieved between the experimental spectrum and the spectrum of the dimer than the one of the monomer. The features at ~1585 cm⁻¹ and 1400–1350 cm⁻¹ are only reproduced in the dimer spectrum, when the DHB is present. Both the monomer and dimer are likely to coexist in solution, though the stronger agreement of the dimer's spectrum to experimental results supports the proposal that a majority of the dimer configuration exists in solution. This confirms that VCD is sensitive to the effect of hydrogen bonding and thus this effect needs to be accounted for when calculating VCD spectra.

Weirich et al. [52] employed the cluster-in-a-liquid model to study how hydrogen bonding affects the VCD spectral signatures of chiral alcohols. They found that hydrogen bonding to the OH-group alters mainly the vibrational modes that have contributions of the H-O-C bending and C-O related deformation motions. They also showed that the farther away the OH-group from the stereocenter, the smaller the effect on the VCD

spectrum.

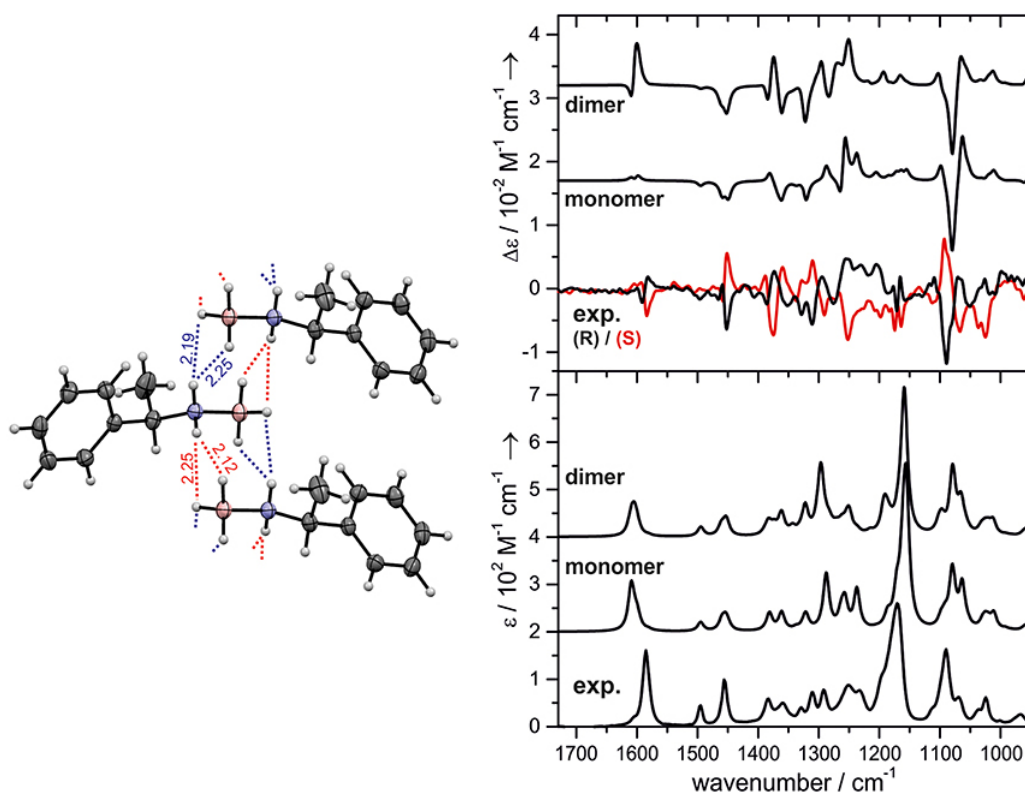


Figure 2.7: Crystal structure of the DHB dimer (left) and experimental and calculated VCD and IR spectra of methylbenzyl-amine-borane (right). This figure has been adapted from Merten et al. [50].

However VCD spectra obtained by Boltzmann-averaging of the different minima contributions do not show optimal agreement with experiment for floppy molecules [52, 53]. Moreover Ghidinelli et al. [54] claimed that solvents with donor capability like methanol seem particularly difficult to treat satisfactorily with the cluster-in-a-liquid model. According to these authors, the inclusion of anharmonicity and temperature effects through the Boltzmann weights provides a flimsy result. It relies on determining enough structures to describe the system fully and the temperature is considered after the determination of the individual spectrum. Finally the cluster-in-a-liquid model does not include long-range interactions with an explicit solvent.

2.5. Classical molecular dynamics

In the previous section, various methods were discussed that offer an improvement on the harmonic approximation, though they can be inaccurate for very floppy molecules which have a very large number of local minima in their potential energy surface. In the cluster-in-the-liquid approach, the temperature is considered through the Boltzmann

weights, subsequently to the determination of the frequencies and spectral intensities.

Atomistic simulations such as molecular dynamics (MD) or Monte Carlo are powerful tools to sample the phase space of a molecular system at finite temperature, realistically.

2.5.1. Principle of molecular dynamics

Molecular dynamics simulation methods naturally account for anharmonic and temperature effects and thus the fluxionality of the system. In MD, successive configurations of the system are generated to give a classical trajectory that specifies how the positions and velocities of the particles vary with time. This is completed by integrating Newton's second law of motion, shown here for only one spatial coordinate x :

$$\frac{d^2 x_i}{dt^2} = \frac{F_{x_i}}{m_i} \quad (2.25)$$

where m_i is the mass of particle i , $\frac{d^2 x_i}{dt^2}$ is its acceleration along the x -axis and F_{x_i} is the force on the particle along the x direction. To integrate Newton's equation of motion, the positions and the dynamical properties (velocities, acceleration, etc.) are approximated as Taylor expansions:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \delta t \vec{v}(t) + \frac{1}{2} \delta t^2 \vec{a}(t) + \frac{1}{6} \delta t^3 \vec{b}(t) + \frac{1}{24} \delta t^4 \vec{c}(t) + \dots \quad (2.26)$$

$$\vec{r}(t - \delta t) = \vec{r}(t) - \delta t \vec{v}(t) + \frac{1}{2} \delta t^2 \vec{a}(t) - \frac{1}{6} \delta t^3 \vec{b}(t) + \frac{1}{24} \delta t^4 \vec{c}(t) + \dots \quad (2.27)$$

$$\vec{v}(t + \delta t) = \vec{v}(t) + \delta t \vec{a}(t) + \frac{1}{2} \delta t^2 \vec{b}(t) + \frac{1}{6} \delta t^3 \vec{c}(t) + \dots \quad (2.28)$$

$$\vec{a}(t + \delta t) = \vec{a}(t) + \delta t \vec{b}(t) + \frac{1}{2} \delta t^2 \vec{c}(t) + \dots \quad (2.29)$$

$$\vec{b}(t + \delta t) = \vec{b}(t) + \delta t \vec{c}(t) + \dots \quad (2.30)$$

where $\vec{r} = (x, y, z)$ is the position, \vec{v} the velocity, \vec{a} the acceleration and \vec{b} and \vec{c} the third and fourth derivatives of the position with respect to time [55]. To propagate the particles and determine their positions and velocities at the subsequent time step $t + \delta t$ from the knowledge at time t , the Verlet algorithm is adopted. In this algorithm, the sum of equations (2.26) and (2.27) is used to determine the positions at the next step :

$$\vec{r}(t + \delta t) = 2\vec{r}(t) - \vec{r}(t - \delta t) + \delta t^2 \vec{a}(t) \quad (2.31)$$

Then to determine the velocities a simple approach is to divide the difference of the positions at time $t + \delta t$ and time $t - \delta t$ by $2\delta t$:

$$\vec{v}(t) = \frac{\vec{r}(t + \delta t) - \vec{r}(t - \delta t)}{2\delta t} \quad (2.32)$$

However with this algorithm, which is straightforward and for which the storage requirements are modest, the velocities are only available when the positions of the next step have been computed. In this work, it is the Beeman algorithm [56] that was chosen, as

it is naturally implemented in the Tinker software package on which our work relies. The key equations for the Beeman method are given in equations (2.33) and (2.34) for the positions and velocities, respectively:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \delta t \vec{v}(t) + \frac{2}{3} \delta t^2 \vec{a}(t) - \frac{1}{6} \delta t^2 \vec{a}(t - \delta t) \quad (2.33)$$

$$\vec{v}(t + \delta t) = \vec{v}(t) + \frac{1}{3} \delta t \vec{a}(t) + \frac{5}{6} \delta t \vec{a}(t) - \frac{1}{6} \delta t \vec{a}(t - \delta t) \quad (2.34)$$

This expression of the velocities is also more precise than the Verlet algorithm, which is important when determining the temperature of the system, as explained in the following.

2.5.2. Thermostatted molecular dynamics

To control the temperature, an NVT or canonical ensemble is employed. In this ensemble the number of particles N and the volume V are constant and the system is maintained at a temperature T . This is useful to equilibrate our systems at the desired temperature, generally 300 K. The temperature of the system is related to the time average of the kinetic energy [55] by:

$$\langle E_{\text{kin}} \rangle = \frac{3}{2} N k_B T \quad (2.35)$$

The average kinetic energy is fixed at a certain temperature, the instantaneous kinetic energy of an N -body system fluctuates with the velocity:

$$E_{\text{kin}}(t) = \frac{1}{2} \sum_{i=1}^N m_i v_i^2(t) \quad (2.36)$$

Thus the instantaneous temperature [55, 57] can be given by:

$$T(t) = \frac{\sum_{i=1}^N m_i v_i^2(t)}{3Nk_B} \quad (2.37)$$

Hence by modifying the velocities is possible to control the temperature. This influence on the MD trajectory to keep the temperature constant is called a thermostat. Three main thermostats exist: Berendsen, Andersen and Nosé-Hoover.

The Berendsen thermostat [58] couples the system to an external heat bath fixed at the desired temperature. At each step, the velocities are scaled such that the rate of change in temperature is proportional to the difference in instantaneous temperature between the bath and the system. However the kinetic energy converges towards a constant with this thermostat, thus the fluctuations of the instantaneous kinetic energy become zero. This absence of fluctuation does not respect the canonical ensemble.

The canonical sampling through velocity rescaling (CSVR) thermostat [59] achieves sampling of the canonical ensemble by propagating the positions and velocities using an areapreserving NVE integrator, then updates the kinetic energy for one time step using some stochastic dynamics in the canonical ensemble, and finally rescales all velocities using the updated kinetic energy.

The next two thermostats respect the canonical ensemble as the instantaneous kinetic energy fluctuates around its average value. The Andersen thermostat [60] chooses a particle at random and changes its velocity to a value randomly selected from a Maxwell-Boltzmann distribution. This is equivalent to random particles of the system undergoing stochastic collisions with a fictitious particle.

The Nosé-Hoover thermostat [61, 62] corrects the velocities by exchanging energy with an extra degree of freedom. This method produces correct fluctuations in the velocities and reproduces well the desired canonical ensemble. It is this thermostat that will be adopted in this work.

NVT ensembles influence the MD trajectory via their thermostat. This influence on the trajectory could interfere with the time series of dipole moments computed. Thus to explore the potential energy surface in the goal of retrieving the electric and magnetic dipole moments, the NVE, or microcanonical ensemble is used once the system is equilibrated at the wanted temperature.

Another way to explore the potential energy surface is through the Monte Carlo method where one or more coordinates of the structure are randomly modified at each step. Generally the Metropolis algorithm [63] is used where if the new structure has a lower energy than the previous one, it is automatically accepted. If the energy is higher, the new structure is only accepted with a Boltzmann probability depending on the temperature of the system. This method does not permit to access dynamical properties of the system and thus will not be exploited here. The next section will develop the acquiring of IR and VCD spectra with time-dependent properties.

2.5.3. Infrared and VCD absorption spectroscopies from time-dependent properties

While IR spectroscopy measures the response of the electric dipole moment to an external excitation, the VCD signal comes from simultaneous changes in electric and magnetic dipole moments of the molecule due to nuclear motion [64, Chapter 4].

To represent the dynamical changes of these two moments, a time-correlation function (TCF) $C(\tau)$ can be used. This function provides a statistical description of the time evolution of one or two random variables at different points in time (t_i and t_j , with $\tau = t_j - t_i$). If the random variables represent the same quantity: $C(\tau) = \langle A(t_i) \cdot A(t_j) \rangle$, then this time-correlation function is called an autocorrelation function. If the variables represent two different quantities ($C(\tau) = \langle A(t_i) \cdot B(t_j) \rangle$) then the time-correlation function is called a cross-correlation function. A and B are dynamical variables [65, Chapter 5.2].

Abbate and co-workers [66] introduced the concept of cross-correlation functions of time-dependent electric and magnetic dipole moments for the computation and analysis of circular dichroism spectroscopy, following the work of Gordon [67] on autocorrelation functions for IR spectroscopy.

According to the Wiener-Khinchine theorem, [68, Chapter 22] the correlation function and the spectral density $S(\omega)$ are the Fourier transform of each other (see Equation 2.38). This theorem is taken within the ergodic hypothesis, which states that for a stationary

random process, a large number of observations made on a single system at n arbitrary instants of time have the same statistical properties as observing n arbitrarily chosen systems at the same time from an ensemble of similar systems:

$$S(\omega) = \int_{-\infty}^{+\infty} C(\tau) e^{-i\omega\tau} d\tau \quad (2.38)$$

Thus for describing the absorption intensity line shape of the IR and VCD signal [8, 66, 69], we can write:

$$I_{\text{IR}}(\omega) = \frac{2\pi\beta\omega^2}{3Vcn(\omega)} \int_{-\infty}^{+\infty} \langle \vec{\mu}(0) \cdot \vec{\mu}(\tau) \rangle e^{-i\omega\tau} d\tau \quad (2.39)$$

$$I_{\text{VCD}}(\omega) = \frac{4\pi\beta\omega}{3Vcn(\omega)} \int_{-\infty}^{+\infty} \langle \vec{\mu}(0) \cdot \vec{m}(\tau) \rangle e^{-i\omega\tau} d\tau \quad (2.40)$$

However the IR spectral line shape can also be obtained from the autocorrelation function of the time derivative of $\vec{\mu}$, and the VCD spectral line shape from the cross correlation function of the time derivative of $\vec{\mu}$ with \vec{m} , which offers a number of practical advantages, including origin independence [8, 12].

In the present work, IR absorption spectroscopy is represented with the time autocorrelation function of the time derivative of electric dipole moment: $C_{\text{IR}}(\tau) = \langle \dot{\vec{\mu}}(0) \cdot \dot{\vec{\mu}}(\tau) \rangle$. To represent VCD, the cross-correlation function $C_{\text{VCD}}(\tau)$ with $A = \dot{\vec{\mu}}(0)$ and $B = \vec{m}(\tau)$ is used, where \vec{m} is the magnetic dipole moment.

Various methods were used to apply Gordon and Abbate's concepts to the determination of IR and VCD spectroscopy from a classical trajectory, starting with MD simulations that employ ab initio harmonic force fields to determine the value of the vibrational frequencies of α -pinene, D-gluconic acid, formaldehyde dimer and acetylprolineamide [9]. The quantum mechanics/molecular mechanics (QM/MM) method was also used to study alanine dipeptide in water [6, 70], (R)-methyloxirane and (L)-alanine in aqueous solutions [71] and (1S)-(-)- β -pinene in CCl_4 [72]. Subsequently ab initio molecular dynamics was adopted to study gas [73, 74, 75] and liquid phases [8]. The last method employed is force field-based molecular dynamics to study camphor [69].

Ab initio molecular dynamics (AIMD) and force field-based molecular dynamics (FFMD) will be the subject of the next two sections.

2.6. Ab initio molecular dynamics

In an ab initio molecular dynamics (AIMD) calculation, to include the electronic structure in MD simulations, electrons are treated quantum mechanically and the nuclei which are heavier are modelled as classical particles. Thus the nuclei are propagated via Newton's second law, equation (2.25) and the energies and the forces are computed at the quantum level.

2.6.1. Born-Oppenheimer molecular dynamics

In the Born-Oppenheimer molecular dynamics (BOMD) method, the static electronic structure problem is solved at each molecular dynamics step given the set of fixed nuclear positions at that instant of time. Thus the electronic structure part is treated within the time-independent Schrödinger equation. This corresponds to an adiabatic treatment where the electronic wavefunction is considered to adapt quasi instantaneously to the nuclei movement.

2.6.2. Car-Parrinello molecular dynamics

Though in BOMD the electrons are treated statically, Car and Parrinello [76] proposed the use of an electronic wavefunction as a variable of the dynamical trajectory, which enables a dynamically evolving electronic subsystem.

While BOMD separates electron and nuclei via time, Car-Parrinello molecular dynamics (CPMD) divides nuclei and electrons through energy to maintain adiabaticity and maps both the electronic and nuclear component classically. Car and Parrinello introduced a new class of Lagrangians that takes into account the energy of the electronic subsystem $\langle \Psi_0 | H_e | \Psi_0 \rangle$ and the possible constraints that can be imposed on the set of orbitals such as orthonormality:

$$L_{CP} = \underbrace{\sum_I \frac{1}{2} M_I \dot{R}_I^2}_{\text{kinetic energy}} + \underbrace{\sum_i \frac{1}{2} \mu_i \langle \dot{\psi}_i | \dot{\psi}_i \rangle}_{\text{potential energy}} - \underbrace{\langle \Psi_0 | H_e | \Psi_0 \rangle}_{\text{potential energy}} + \underbrace{\text{constraints}}_{\text{orthonormality}} \quad (2.41)$$

with M_I and R_I the nuclear masses and positions, respectively. The equations of motion of Car-Parrinello MD are:

$$\begin{cases} M_I \ddot{R}_I = -\frac{\partial}{\partial R_I} \langle \Psi_0 | H_e | \Psi_0 \rangle + \frac{\partial}{\partial R_I} \text{constraints} \\ \sigma \ddot{\psi}_i = -\frac{\delta}{\delta \psi_i^*} \langle \Psi_0 | H_e | \Psi_0 \rangle + \frac{\delta}{\delta \psi_i^*} \text{constraints} \end{cases} \quad (2.42)$$

where σ is a fictitious mass associated with the electronic wavefunction. It ensures an integration of the equations of motion within a reasonable time step and adiabatic separation of the electrons and nuclei. The unit of this mass parameter is energy times a squared time for dimensional balance [77, 78].

2.6.3. IR and VCD spectra from nuclear velocity perturbation theory

The adiabatic treatment is inadequate for predicting VCD [79, 80, 81]. To circumvent this issue, nuclear velocity perturbation theory (NVPT) can be used. In this approach, the total electron-nuclear wavefunction $\Psi(\vec{r}, \vec{R}, t)$ is factorized in the nuclear part $\chi(\vec{R}, t)$ and the electronic part $\Phi_{\vec{R}}(\vec{r}, t)$, which depend on the electronic coordinates \vec{r} and on the nuclear coordinates \vec{R} :

$$\Psi(\vec{r}, \vec{R}, t) = \Phi_{\vec{R}}(\vec{r}, t) \chi(\vec{R}, t) \quad (2.43)$$

In NVPT, the electronic part is taken as the Born-Oppenheimer (BO) ground state $\Phi_{\vec{R}}^{(0)}(\vec{r})$, corrected by a time-dependent, purely imaginary perturbative contribution $i\Phi_{\vec{R}}^{(1)}(\vec{r}, t)$:

$$\Phi_{\vec{R}}(\vec{r}, t) = \Phi_{\vec{R}}^{(0)}(\vec{r}) + i\Phi_{\vec{R}}^{(1)}(\vec{r}, t) \quad (2.44)$$

The electronic Hamiltonian can be partitioned into the standard BO Hamiltonian \hat{H}_{BO} with eigenfunctions $\Phi_{\vec{R}}^{(0)}(\vec{r})$ and a small NVPT perturbation:

$$\hat{H}_{\text{el}}(\vec{r}, \vec{R}) = \hat{H}_{\text{BO}} + \sum_{\nu=1}^{N_n} \lambda_{\nu}(\vec{R}, t) \cdot (-i\hbar\nabla_{\nu}) \quad (2.45)$$

with the nuclear gradient operator acting on the parametric dependence of the electronic wavefunction. Here, the sum runs over the N_n nuclei. In the semiclassical limit for the nuclei, the perturbation parameter $\lambda_{\nu}(\vec{R}, t)$ corresponds to the nuclear velocity \hat{V}_{ν} ,

$$\lambda_{\nu}(\vec{R}, t) = \frac{1}{M_{\nu}} \frac{-i\hbar\nabla_{\nu}\chi(\vec{R}, t)}{\chi(\vec{R}, t)} = \hat{V}_{\nu} \quad (2.46)$$

with the ionic mass M_{ν} . The total perturbative correction to the electronic wavefunction $i\Phi_{\vec{R}}^{(1)}(\vec{r}, t)$ is obtained as a weighted sum over the corrections for each component:

$$\Phi_{\vec{R}}^{(1)}(\vec{r}, t) = \sum_{\nu=1}^{N_n} \sum_{\alpha=1}^3 \lambda_{\alpha}^{\nu}(\vec{R}, t) \Phi_{\vec{R}, \nu\alpha}^{(1)}(\vec{r}) \quad (2.47)$$

Using the NVPT treatment, it is thus possible to access the VCD properties through the following expressions:

$$\hat{\mu} = \hat{\mu}^e + \hat{\mu}^n = - \sum_{i=1}^{N_e} q \hat{r}_i + \sum_{\nu=1}^{N_n} Z_{\nu} q \hat{R}_{\nu} \quad (2.48)$$

and

$$\hat{m} = \hat{m}^e + \hat{m}^n = - \sum_{i=1}^{N_e} \frac{q}{2c} \hat{r}_i \times \hat{v}_i + \sum_{\nu=1}^{N_n} \frac{Z_{\nu} q}{2c} \hat{R}_{\nu} \times \hat{V}_{\nu} \quad (2.49)$$

Here, q is the electronic charge, $Z_{\nu}q$ is the nuclear charge, and c is the velocity of light in vacuum. The position and velocity operators for the electronic subsystem are indicated as \hat{r}_i and \hat{v}_i , respectively, and similar symbols are used for the nuclear operators \hat{R}_{ν} and \hat{V}_{ν} .

An implementation of nuclear velocity perturbation theory (NVPT) was carried out in the plane-wave codes CPMD [82] and in CP2K [75] and tested in the gas [73] and liquid [8] phases.

2.6.4. Example of propylene oxide: evidence for chirality transfer

The NVPT approach was applied to (R)-propylene-oxide (RPO) in water, for which chirality transfer was found from a chiral solute to an achiral solvent. With chirality transfer, an achiral species gains some of the chiral properties of the chiral system. This transfer has been theoretically demonstrated for RPO in water but also methyl lactate in water [83, 84]. The system of study is composed of 1 RPO molecule, represented in figure 2.8 and 53 water molecules.

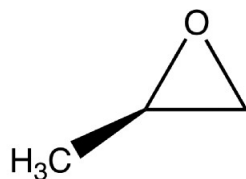


Figure 2.8: Schematic representation of (R)-propylene-oxide (RPO). This figure has been adapted from Scherrer et al. [8].

According to Scherrer et al. [8] the most relevant contributions to the VCD spectra arise from the molecules in the first solvation shell, for RPO in water. Scaled molecular moments were introduced which include only the molecular moments of a finite region around a chosen center. For a chiral molecule K , the surrounding molecules J are included as:

$$\vec{\mu}_K^{\text{scale}}(t) = \vec{\mu}_K(t) + \sum_{J \neq K} P_{KJ}(t) \vec{\mu}_J(t) \quad (2.50)$$

$$\vec{m}_K^{\text{scale}}(t) = \vec{m}_K(t) + \sum_{J \neq K} P_{KJ}(t) \vec{m}_J(t) \quad (2.51)$$

with the damping function being $P_{KJ}(t) = \left(1 + e^{(|\vec{R}_{KJ}(t)| - R_0)/D}\right)^{-1}$, for intermolecular contributions. R_0 corresponds to the distance included from the center of mass of the solute, it will be referred to as the cutoff. D is the sharpness parameter and it is chosen equal to 0.25 Å as in reference [85]. It is this cutoff that will be used in chapter 5 for hydrated alanine.

The IR signal of water and VCD signals of RPO in water are shown in figure 2.9. The blue lines correspond to the experimental data [83], the red lines to the VCD signal considering only the moments of the RPO molecule. The dashed line is the VCD signal of the RPO molecule and the molecular contributions of the water molecules near to the solute.

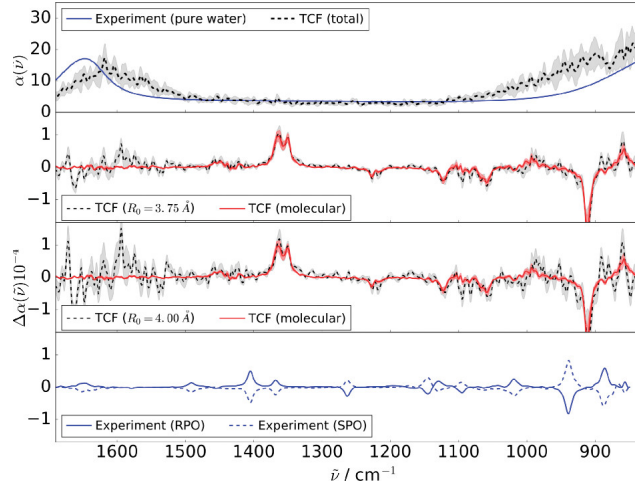


Figure 2.9: Dynamical IR and VCD spectra of a (R)-propylene-oxide molecule solvated in water at 340 K (1 RPO molecule in 53 H₂O molecules). The intensities are in units of km cm mol⁻¹. We show the molecular correlation and damping cutoffs of $R_0 = 3.75 \text{ \AA}$ and $R_0 = 4.0 \text{ \AA}$, which includes the hydrogen bonded water molecules. The shaded areas indicate the standard deviation of the statistically independent microcanonical simulations. The experimental data are taken from reference [83] and are in arbitrary units. The computed spectra are divided by the refractive index of bulk water [86] to compare with the experiment. This figure has been reprinted from Scherrer et al. [8].

At 1650 cm^{-1} the signal corresponding to the bending mode of water is expected. The presence of this VCD signal of achiral water shows the existence of chirality transfer as the VCD spectra of an achiral molecule is expected to be vanishing.

If we consider a cutoff of $R_0 = 3.75 \text{ \AA}$, only a slight contribution of the water is seen, though if the cutoff is increased to 4.0 \AA , the signal becomes more intense and noisy. The disorder can be measured through the mean standard deviation of the VCD spectra:

$$\sigma(R_0) = \sqrt{\frac{\sum^N (I_i^{\text{VCD}}(R_0) - I_{\text{av}}^{\text{VCD}}(R_0))^2}{N}} \quad (2.52)$$

The spectra from $N=8$ microcanonical simulations (I_i^{VCD}) was compared to the average spectrum from all the simulations ($I_{\text{av}}^{\text{VCD}}$). The individual spectra and the average spectrum are considered for different values of R_0 which are given in figure 2.10. Three different systems were studied, neat RPO with 13 RPO and 1 propanal molecules and two solvated systems: one with 7 RPO molecules in 34 H₂O molecules (low solvation) and one with 1 RPO molecule in 52 H₂O molecules (high solvation).

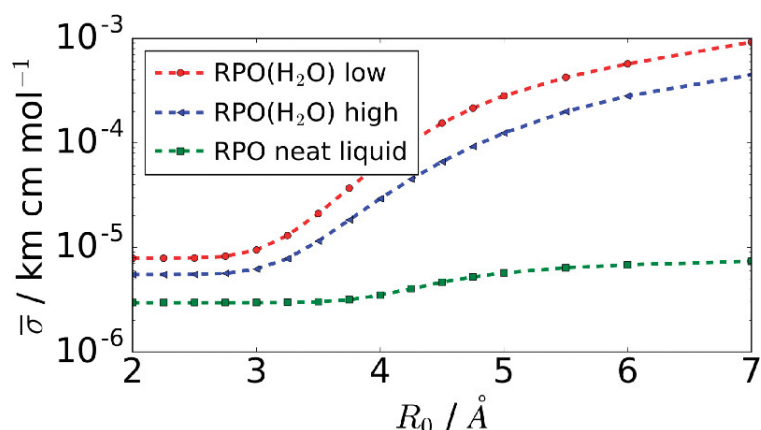


Figure 2.10: Cutoff dependence of the mean standard deviation of the VCD spectra. The magnitudes at the molecular level, i.e., at small distances, originate from the different amounts of sampling. This figure has been reprinted from Scherrer et al. [8].

As the value of the cutoff increases, the VCD spectra get more and more disordered, thus considering a cutoff of $R_0 = 3.75 \text{ \AA}$ is a good compromise as it also includes the first solvation shell.

In BOMD or CPMD studies of vibrational spectroscopy, the solvent is typically modeled by a continuum or very few solvent molecules [81]. This method is also expensive which leads to a short exploration time, a few dozen picoseconds for a system with around 100 atoms.

2.7. Classical force fields for molecular simulations

In the previous section, the electronic part was treated with quantum chemistry methods. Here the whole system is treated with classical dynamics. Thus the potential energy and dipole moment surfaces are calculated using a collection of equations and associated constants describing the dependence of the energy of a system on the coordinates of its particles. This is called a force field (FF) thus the name force field-based molecular dynamics (FFMD) when the collection of equations is used within MD. The use of a force field leads to the ability to explore for longer periods of time as the calculation for each step is less expensive than with AIMD. Larger systems can also be considered to include the solvent explicitly.

2.7.1. Nonpolarisable contributions

Force fields for organic systems can be divided into four key components: bond stretching, angle bending, torsional terms and non-bonded interactions.

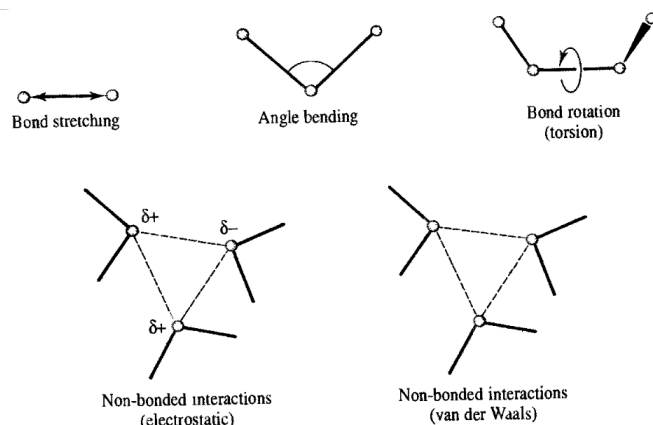


Figure 2.11: Schematic representation of the four key contributions to a molecular mechanics force field: bond stretching, angle bending and torsional terms and non-bonded interactions. This figure has been reprinted from reference [55].

The potential energy V of a system of N atoms can be written as a sum of these four contributions:

$$V(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = V_{\text{bond}} + V_{\text{angle}} + V_{\text{torsion}} + V_{\text{non-bonded}} \quad (2.53)$$

$$\begin{aligned}
 V(\vec{r}^N) = & \sum_{\text{bond}} \frac{k_r}{2} (r - r_0)^2 + \sum_{\text{angle}} \frac{k_\theta}{2} (\theta - \theta_0)^2 \\
 & + \sum_{\text{torsions}} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) \\
 & + \sum_{i=1}^N \sum_{j=i+1}^N \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right)
 \end{aligned} \quad (2.54)$$

where $(r - r_0)$ and $(\theta - \theta_0)$ are the deviation of the bond length and of the angle from their reference values, k_r and k_θ are corresponding force constants. Thus each bond and angle is described by a force constant and a reference value. The torsions are described by a constant V_n , the torsion angle ω and the phase factor γ . The non-bonding part includes a van der Waals term and the electrostatic interactions [55].

Bond length term

This quadratic description of the potential due to the variations of the bond length is appropriate for large systems (proteins, polymers...) and is implemented in the CHARMM [87], AMBER [88], OPLS [89] and GROMOS [90] force fields. For smaller molecules, and to more accurately model the Morse curve, terms with higher orders can be added to V_{bond} , according to a Taylor expansion:

$$V_{\text{bond}}(r) = \frac{k_r}{2} (r - r_0)^2 \left[1 - k'(r - r_0) - k''(r - r_0)^2 - k'''(r - r_0)^3 \dots \right] \quad (2.55)$$

This expression of V_{bond} is used in the MM1 [91], MM2 [92], MM3 [93] and MM4 [94] models developed by Allinger and dedicated to organic molecules. As the cubic term

passes through a maximum, this potential must only be used close to the reference value, where it describes the Morse potential well. To prevent the overly large lengthening of bonds, a quartic term can also be added, which is the case in MM3. The AMOEBA force field used in this thesis, and described below, is based on the MM3 formalism [95].

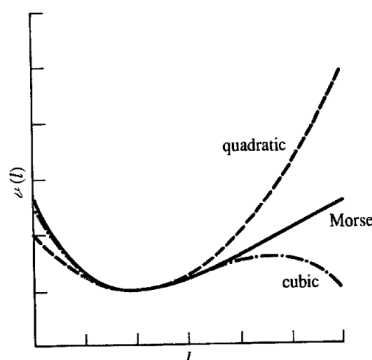


Figure 2.12: A cubic bond-stretching potential passes through a maximum but gives a better approximation to the Morse curve close to the equilibrium structure than the quadratic form. This figure has been reprinted from reference [55].

Angular term

A Taylor expansion can also be applied to the angle bending, through the following expression:

$$V_{\text{angle}}(\theta) = \frac{k_{\theta}}{2}(\theta - \theta_0)^2 \left[1 - k'(\theta - \theta_0) - k''(\theta - \theta_0)^2 - k'''(\theta - \theta_0)^3 \dots \right] \quad (2.56)$$

As the Taylor expansion of the bond length term, this expression of the angle terms is used in the MM1, MM2, MM3 and MM4 series of force fields and in the AMBER, CHARMM and AMOEBA force fields.

Torsional term

Most force fields for organic molecules include the torsional term explicitly. This is mostly achieved through a cosine series expansion:

$$V_{\text{torsions}}(\omega, \gamma) = \sum_{\text{torsions}} \left[\frac{V_1}{2} (1 + \cos(\omega - \gamma)) + \frac{V_2}{2} (1 - \cos(2\omega - \gamma)) + \frac{V_3}{2} (1 + \cos(3\omega - \gamma)) \right] \quad (2.57)$$

Non-bonding terms

The non-bonding terms describe inter- and intra-molecular interactions which are described with a van der Waals (vdW) term and an electrostatic term. The van der Waals term includes the repulsion ($\propto r^{-12}$) and attraction ($\propto r^{-6}$) terms are often described by the 12-6 Lennard-Jones potential:

$$V_{\text{vdW}} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (2.58)$$

where σ is the distance at which $E_{vdW} = 0$, ϵ is the depth of the well represented in figure 2.13 and r is the distance between two atoms. The repulsion term accounts for Pauli's repulsion between electrons which prevents the overlapping of the electron clouds. This term dominates when r is lower than the equilibrium value r_0 . At large values of r , it is the attractive term that dominates. This term accounts for the weak dispersion attraction between two atoms at long range.

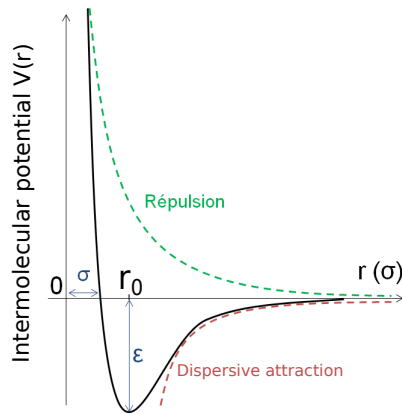


Figure 2.13: Lennard-Jones potential as a function of the particle distance. This figure has been reprinted from reference [96].

The idea of the electrostatic term is to represent the electronic response of the system to its environment, it is represented by Coulomb's law:

$$V_{\text{Coulomb}} = \frac{q_A q_B}{4\pi\epsilon_0 r_{AB}} \quad (2.59)$$

This law describes the interaction between two point charges on two atoms (A and B). The corresponding force $F_{\text{Coulomb}}(r) = -dV_{\text{Coulomb}}/dr$ can be represented by the following, as in figure 2.14.

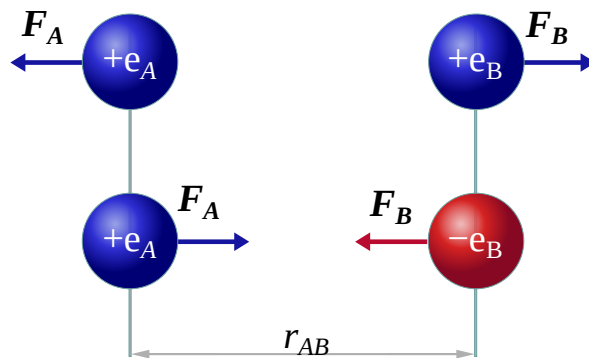


Figure 2.14: Two equal point charges repel each other, and two opposite charges attract each other, with an electrostatic force F . This figure has been reprinted from reference [97].

This description of the vdW term and of the electrostatic interaction terms is used in conventional force fields. However these force fields are not accurate enough for spectroscopy. A reason for this is the neglect of molecular polarisation. This has led to recent FF development to improve the treatment of electrostatics [98, 99, 100]. These improvements include the use of multipoles centered on each atom and permit the introduction of polarisation and charge transfer effects [101, 102].

2.7.2. Accounting for polarisation

Polarisation of electronic clouds is the deformation experienced by the charge on atoms in response to the induced electric field from the charges of the other atoms.

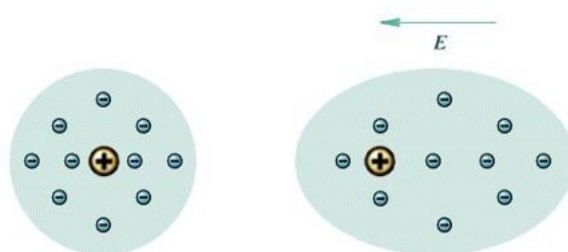


Figure 2.15: Schematic representation of the deformation due to the induced electric field E from the charges of other atoms.

To describe this effect, three models are commonly employed: the fluctuating charges model [103], the Drude model [104] and the induced dipoles model [98, 105]. The manner in which polarisation is accounted for in these models is schematised in figure 2.16.

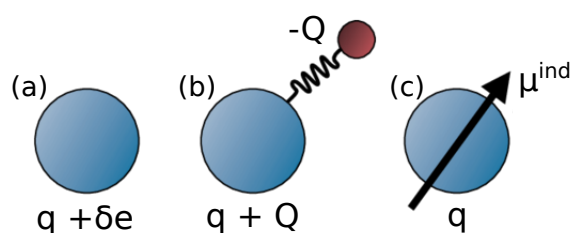


Figure 2.16: Qualitative description of approaches to modelling polarisation. (a) Fluctuating charges model where atomic charges are modified to reproduce the effect of polarisation; (b) Drude oscillator model where a point charge is attached to each atom by means of a spring, and moved, in order to reproduce polarisation effects; (c) induced dipole model where an explicit dipole term is added to each point charge. This figure has been reprinted from Illingworth et al. [106].

Fluctuating charges

In the fluctuating charges model, the fixed charges are replaced with charges that can vary by incorporating an additional electrostatic energy term. This term represents the

energy required to modify the charge of an individual atom. The idea is to redistribute the atomic charges to equalize the electronegativity at each site, according to the electronegativity equalization principle [107]. The total electrostatic energy is minimized, for each time step, with respect to the atomic charges. For a system of N atoms with charges q_1, q_2, \dots, q_N , the electrostatic energy is written as:

$$V_{\text{Fluct}} = \sum_{i=1}^N \left(E_{i0} + \chi_i^0 q_i \right) + \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i}^N q_i q_j J_{ij} \quad (2.60)$$

where E_{i0} is the energy of atom i at zero charge, χ_i^0 is the electronegativity and J_{ij} is a term that mimics the Coulomb interaction at long distance but, typically, smoothly reaches finite values at vanishing distances.

To obtain the charges of the system, equation (2.60) is minimized with respect to the sum of the charges q_1, q_2, \dots, q_N which are fixed for the whole system. This yields a $(N + 1) \times (N + 1)$ linear system to solve:

$$\frac{\partial V_{\text{Fluct}}}{\partial q_1} = \frac{\partial V_{\text{Fluct}}}{\partial q_2} = \dots = \frac{\partial V_{\text{Fluct}}}{\partial q_N} \quad \text{and} \quad \sum_i q_i = Q_{\text{tot}} = \text{const.} \quad (2.61)$$

This model permits the calculation of polarisation without the addition of any new interactions. However one problem with the fluctuating charges method is that it does not reproduce out-of-plane polarisation for planar or linear chemical moieties. This is understandable because electronegativity equalization proceeds along the bonds between atoms.

Drude model

In the Drude model (or shell model), the fluctuation of the charge is represented by a charge (Q) on a harmonic oscillator centred on the atom. The displacement of this charge, by a distance \vec{d} , in response to an electric field creates a dipole moment $\vec{\mu}^{\text{Drude}}$ on atom i :

$$\vec{\mu}_i^{\text{Drude}} = -Q_i \vec{d}_i \quad (2.62)$$

The total energy of a system described with the Drude model can be written:

$$V_{\text{Drude}}(\vec{r}_i, \vec{d}_i) = \sum_{i=1}^N \left\{ \frac{1}{2} k_i d_i^2 + q_i \left[\vec{r}_i \cdot \vec{E}_i^0 - (r_i + d_i) \cdot \vec{E}_i^{0'} \right] \right\} + \sum_{i=1}^N \left\{ \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i}^N \left(\frac{1}{r_{ij}} - \frac{1}{|r_{ij} - d_j|} - \frac{1}{|r_{ij} + d_i|} \right) + \frac{1}{|r_{ij} - d_i + d_j|} \right\} \quad (2.63)$$

where k_i is the spring constant, E_i^0 is the static field at the location of the core charge, r_i , and $E_i^{0'}$ is the static field at the location of the shell charge, $r_i + d_i$. Note that $E_i^0 \neq E_i^{0'}$ in general. This model uses particle-particle interactions, thus keeping the computational cost low.

Induced dipoles

Most polarisable models developed nowadays employ the induced dipole model approach [108, 109, 110, 111, 112, 113]. The induced dipole model adds induced dipoles, centered on each atom i , to the permanent charge:

$$\vec{M}_i = \vec{M}_i^0 + \vec{M}_i^{\text{ind}} \quad (2.64)$$

M_i includes fixed point charges, as well as possibly permanent dipoles and quadrupoles:

$$\vec{M}_i = [q_i, \mu_{ix}, \mu_{iy}, \mu_{iz}, Q_{ixx}, Q_{ixy}, Q_{ixz}, \dots, Q_{izz}] \quad (2.65)$$

with q_i the point charge on atom i , μ the dipole and Q the quadrupoles.

The induced dipole is proportional to the electric field, \vec{E}_i on site i . The electric field arises both from the permanent charges of the system and from the other induced dipoles:

$$\vec{\mu}_i^{\text{ind}} = \alpha_i \vec{E}_i = \alpha_i \left(\sum_j T_{ij}^1 \vec{M}_j^0 + \sum_{j'} T_{ij'}^{11} \vec{\mu}_{j'}^{\text{ind}} \right) \quad (2.66)$$

with α_i the isotropic atomic polarisability, T_{ij}^1 a 3×13 matrix representing the second to the fourth rows of the multipole-multipole interaction matrix T_{ij} and $T_{ij'}^{11}$ is a 3×3 submatrix consisting of elements in T_{ij} corresponding to the dipole moments. The sum over j represents all atomic sites outside the polarisation group of atom i and the sum over j' corresponds to all atomic sites other than i itself. T_{ij} is the interaction matrix between sites i and j :

$$T_{ij} = \begin{bmatrix} 1 & \frac{\partial}{\partial x_j} & \frac{\partial}{\partial y_j} & \frac{\partial}{\partial z_j} & \dots \\ \frac{\partial}{\partial x_i} & \frac{\partial}{\partial x_i x_j} & \frac{\partial}{\partial x_i y_j} & \frac{\partial}{\partial x_i z_j} & \dots \\ \frac{\partial}{\partial y_i} & \frac{\partial}{\partial y_i x_j} & \frac{\partial}{\partial y_i y_j} & \frac{\partial}{\partial y_i z_j} & \dots \\ \frac{\partial}{\partial z_i} & \frac{\partial}{\partial z_i x_j} & \frac{\partial}{\partial z_i y_j} & \frac{\partial}{\partial z_i z_j} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \left(\frac{1}{r_{ij}} \right) \quad (2.67)$$

Since the induced dipoles also contribute to the electric field, they are solved iteratively through a self-consistent field.

The polarisation energy due to the induced dipoles can be expressed as:

$$V_{\text{ind}} = -\frac{1}{2} \sum_{i=1}^N \left(\vec{\mu}_i^{\text{ind}} \right)^t \cdot \vec{E}_i^0 \quad (2.68)$$

where \vec{E}_i^0 is the electric field due to the permanent charges and $()^t$ the transpose.

Most developments in the polarisable force fields for biomolecules relate to explicitly induced dipoles and somewhat less to fluctuating charges and Drude oscillators. This is despite induced dipoles requiring iterative computation, which is a much slower process compared with the other additive models. The main advantage of this method is it only requires the addition of a few parameters to the force field, namely the atomic polarisabilities. It is the induced dipoles model that is used by the AMOEBA force field.

2.7.3. Periodic boundary conditions

To model liquid and solid phases, periodic boundary conditions (PBC) [55] are needed to mimic infinitely extended systems. For a simulation of a system using PBC, the molecules are represented in a unit cell, any molecule that crosses one boundary of the cell reappears on the opposite side, as represented in figure 2.17.

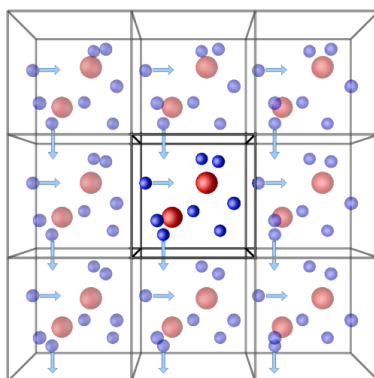


Figure 2.17: Visual representation of the idea of periodic boundary conditions. This figure has been reprinted from reference [114].

In PBC and with a globally neutral system, long distance interactions can be described with the Ewald summation. In this method, the long-range energy contributions (usually selected by a cutoff function) are calculated in Fourier space, where they converge faster than in real space [115]. In the program used for the MD simulations, Ewald summation is used for the polarisable atomic multipoles.

The non-bonded van der Waals interaction is typically of short range relative to, for example, electrostatic forces. This interaction is truncated at a cutoff distance. The value of the corresponding van der Waals potential is brought to zero for distances larger than the cutoff [116].

2.7.4. The AMOEBA force field

The AMOEBA (atomic multipole optimized energetics for biomolecular simulation) polarisable force field [95] has been used in this thesis to describe the molecular systems of interest through molecular dynamics simulations. AMOEBA was originally developed for biomolecules such as proteins and DNA, and some organic molecules. The main advantage of the force field is its ability to reproduce relative conformational energies [117]. Conventional force fields, that do not include polarisation effects, only reproduce ~50 % of the conformations found at the MP2/aug-cc-pVTZ level of theory whereas AMOEBA reproduces ~80 % of the conformations [118]. This leads to a good precision for thermodynamic properties such as free energies of solvation and complexation. The AMOEBA force field has also been extended to IR spectroscopy very satisfactorily [101, 119].

The good reproduction of the relative energy is due to the use of permanent electrostatic multipolar moments on each atom by this force field. These moments include point

charges as well as dipole and quadrupoles moments and are derived from high-level gas phase quantum mechanical calculations [120]. This derivation is performed via Stone's distributed multipole analysis (DMA) where the main idea is to distribute the charge density from the quantum wavefunction into local multipoles moments [121].

In practice, AMOEBA combines the traditional ingredients of nonpolarisable models, including bonded and non-bonded terms, with a multipolar description of electrostatics on each atom up to the quadrupole, and induced dipoles reacting self-consistently to the presence of local electric fields on each polarisable site.

Previously a general overview of the contribution to the energy of the system was discussed, then, the specificities of AMOEBA are now presented [95].

$$V = V_{\text{bond}} + V_{\text{angle}} + V_{b\theta} + V_{\text{oop}} + V_{\text{torsion}} + V_{\text{vdW}} + V_{\text{ElectroStat}} \quad (2.69)$$

where the first five terms correspond to bond stretching, angle bending, bond-angle cross term, out-of-plane bending, and torsional rotation. These are the short-range valence interactions. The last two terms correspond to the non-bonded van der Waals and electrostatic contributions, including Coulomb and polarisation interactions.

Covalent interactions

The bond stretching V_{bond} , angle bending V_{angle} and the coupling between the stretching and bending $V_{b\theta}$ are represented by equations (2.70) to (2.72). These correspond to those of the MM3 force field [93] in which anharmonicity is accounted for by the terms of order 3 and 4.

$$V_{\text{bond}} = K_b(r - r_0)^2 \left[1 - 2.55(r - r_0) + \left(\frac{7}{12}\right) 2.55^2(r - r_0)^2 \right] \quad (2.70)$$

$$V_{\text{angle}} = K_\theta(\theta - \theta_0)^2 \left[1 - 0.014(\theta - \theta_0) + 5.6 \times 10^{-5}(\theta - \theta_0)^2 - 7 \times 10^{-7}(\theta - \theta_0)^3 + 2.2 \times 10^{-8}(\theta - \theta_0)^4 \right] \quad (2.71)$$

$$V_{b\theta} = K_{b\theta} \left[(r - r_0) + (r' - r'_0) \right] (\theta - \theta_0) \quad (2.72)$$

where K_b , K_θ and $K_{b\theta}$ are force constants.

The out-of-plane bending term V_{oop} is controlled by a Wilson-Decius-Cross function [122] for the sp^2 -hybridized trigonal centers only:

$$V_{\text{oop}} = K_\chi \chi^2 \quad (2.73)$$

where K_χ is the force constant and χ the dihedral angle between the 4 atoms that are considered.

The torsion terms here originate from a Fourier expansion of a 1-fold through 6-fold trigonometric form:

$$V_{\text{torsion}} = \sum_{n=1}^6 K_{n\omega} [1 + \cos(n\omega \pm \gamma)] \quad (2.74)$$

The other coupling terms (torsion-bond, torsion-angle and angle-angle) incorporated in the MM3 force field have not been conserved in AMOEBA.

Non-bonded interactions

The first non-bonded interaction is the van der Waals interaction term V_{vdW} which takes the buffered 14-7 [123] functional form instead of the standard 12-6 Lennard-Jones form:

$$V_{\text{vdW}}(ij) = \epsilon_{ij} \left(\frac{1.07}{\rho_{ij} + 0.07} \right)^7 \left(\frac{1.12}{\rho_{ij}^7 + 0.12} - 2 \right) \quad (2.75)$$

with

$$\rho_{ij} = r_{ij}/r_{ij}^0 \quad (2.76)$$

$$r_{ij}^0 = \frac{(r_{ii}^0)^3 + (r_{jj}^0)^3}{(r_{ii}^0)^2 + (r_{jj}^0)^2} \quad (2.77)$$

$$\epsilon_{ij} = \frac{4\epsilon_{ii}\epsilon_{jj}}{(\sqrt{\epsilon_{ii}} + \sqrt{\epsilon_{jj}})^2} \quad (2.78)$$

where ϵ_{ij} is the potential well depth, r_{ij} is the distance between atoms i and j and r_{ij}^0 is the minimum energy distance. The equation applies to heterogeneous atom pairs.

The parameters involved in these expressions are fitted to both gas phase and bulk phase experimental properties. Halgren [123] showed that the conventional 12-6 Lennard-Jones potentials are too repulsive at short contact distances and the buffered 14-7 potentials reproduced well ab initio results of noble gases.

The electrostatic term has two contributions. The first arises from the permanent multipolar moments on each atom, and replaces the Coulomb term described previously. The second contribution emerges from the induced dipole moments.

In the induced dipoles model, the set of induced dipoles is solved iteratively and is accelerated in the AMOEBA force field via successive overrelaxation (SOR) [124] using a default value of $\omega = 0.7$ in:

$$\vec{\mu}_i^{\text{ind}}(n+1) = (1 - \omega)\vec{\mu}_i^{\text{ind}}(n) + \omega\alpha_i \left(\sum_j T_{ij}^1 \vec{M}_j^0 + \sum_{j'} T_{ij'}^{11} \vec{\mu}_{j'}^{\text{ind}}(n) \right) \quad (2.79)$$

In this expression of the induced dipoles, the atomic polarisabilities α_i are present, for which the values have been previously determined through Applequist, Carl, and Fung's model [125].

However Thole [126] pointed out that at small distances, $r_{12} \leq (4\alpha_1\alpha_2)^{1/6}$, the polarisation energy becomes infinite. To circumvent this "polarisation catastrophe", he proposed a damping function to this energy:

$$\rho = \frac{3a}{4\pi} \exp(-au^3) \quad (2.80)$$

with $u = r_{ij}/(\alpha_i\alpha_j)^{1/6}$ and a a dimensionless parameter that controls the damping strength. In the AMOEBA force field this distance-dependent damping function is combined with the exclusion of some of the 1–2 bonded polarisation interactions [127].

Summing all the energy contributions gives the energy of the system as described by the AMOEBA force field.

The AMOEBA force field is distributed with the Tinker software package [128], developed by Ponder and coworkers. It is this software that was modified to incorporate the ingredients needed to calculate VCD spectra through force field molecular dynamics.

The Tinker software [128] provides a good framework to perform molecular dynamics as well as other energy-based and structural manipulation calculations such as local energy minimisations. It is a code composed of relatively small programs that inter-operate to perform complex computations. Both these elements ease the addition of new code.

2.8. Implementation of VCD spectroscopy within AMOEBA

In section 2.5.3, the absorption intensity line shape of an IR spectrum is shown to be obtained via the Fourier transform of the autocorrelation of the electric dipole moment. Likewise, the line shape of a VCD spectrum can be obtained via the Fourier transform of the cross-correlation of the electric and magnetic dipole moments. Thus to reach these spectra, the electric and magnetic dipole moments must be acquired throughout time.

The dynamics of our molecular system is described by an atomistic force field, in which the potential energy and gradient assume a particular form that explicitly accounts for multipolar electrostatics, including polarisation forces and induced dipoles. For any configuration \mathbf{R} of the system, the total electric dipole moment $\vec{\mu}$ contains a contribution at lowest order from the partial charges $\{q_i\}$ distributed at the respective positions \vec{r}_i , and a contribution of the induced dipoles $\vec{\mu}_i^{\text{ind}}$ on each polarisable site i :

$$\vec{\mu}(t) = \sum_i^N \left(q_i \vec{r}_i(t) + \vec{\mu}_i^{\text{ind}}(t) \right) \quad (2.81)$$

where N denotes the number of atoms. Once the electric dipole moment is determined for a molecule, its time derivative is calculated.

However the calculation of the magnetic dipole moment $\vec{m}(t)$ is not included in the force field and software adopted here. Thus we define our magnetic moment by basing the expression on the magnetic moment used for conventional force fields [69]:

$$\vec{m}_{\text{ConvFF}}(t) = \frac{1}{2c} \sum_i^N q_i \vec{r}_i(t) \times \vec{v}_i(t) \quad (2.82)$$

where $\vec{v}_i(t)$ denotes the velocity vector of atom i at time t , q_i its charge and c the velocity of light.

The transformation from conventional force field to polarisable FF involves the addition of the induced dipole moment that needs to participate in the magnetic moment.

Thus a second term is added, corresponding to the moving induced dipole as given by Hnizdo [129]:

$$\vec{m}_{\text{MoveInd}}(t) = \frac{1}{c} \dot{\vec{\mu}}_i^{\text{ind}}(t) \times \vec{v}_i(t) \quad (2.83)$$

The effects of the time varying induced dipoles moments due to the motion of neighbouring charges also need to be considered. Therefore a last term is added to the magnetic moment:

$$\vec{m}_{\text{Fluc}}(t) = \frac{1}{2c} \vec{r}_i(t) \times \dot{\vec{\mu}}_i^{\text{ind}}(t) \quad (2.84)$$

This contribution is also present in the description of the total magnetic dipole moment in equation (41) of reference [8].

Thus we define the magnetic dipole moment as the sum of the moving point charges, the moving induced dipoles and the time varying induced dipoles moving around fixed positions:

$$\vec{m}(t) = \sum_i^N \left[\frac{1}{2c} q_i \vec{r}_i(t) \times \vec{v}_i(t) + \frac{1}{c} \dot{\vec{\mu}}_i^{\text{ind}}(t) \times \vec{v}_i(t) + \frac{1}{2c} \vec{r}_i(t) \times \dot{\vec{\mu}}_i^{\text{ind}}(t) \right] \quad (2.85)$$

It is this expression of the magnetic moment that was added to the Tinker software package, for this work.

Returning to the expression of the electric dipole moment, the velocity form of this moment $\dot{\vec{\mu}}(t) = \frac{d\vec{\mu}(t)}{dt}$ is exploited to avoid the origin dependence that arises through the positions [12, Chapter 4]. This origin dependence also arises within a periodic cell, which will be developed in the next section.

2.8.1. Specific difficulties under periodic boundary conditions

In the case of PBC, the minimum image convention is used. This ensures that the distance between the origin O and any atom i at a position \vec{r}_i is the lowest, taking the periodic image if necessary. This convention is expressed by translating \vec{r}_i by $\Delta_{O_i}^{\text{PBC}}$, a conditional lattice translation, which is a multiple of the size of the PBC cell. Thus, depending on where the origin is placed, the original value of \vec{r}_i or a periodic image value will be used. This changes the values of \vec{r}_i and therefore \vec{m} according to the origin, as shown by the black lines in figure 2.18.

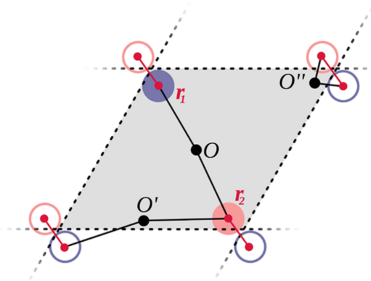


Figure 2.18: Sketch illustrating the common origin problem within periodic boundaries. This figure has been reprinted from Jähnigen et al. [7].

To address this issue, Jähnigen et al [7] have recently proposed an alternative formalism that establishes an origin invariant version of the TCF function. This is achieved by considering the nearest periodic image directly between atoms instead of through the origin, as represented by the red lines in figure 2.18 and by the following equation:

$$C_{\text{indep}}(t) = \left\langle \dot{\vec{\mu}}(0) \cdot \vec{m}(t) \right\rangle + \frac{1}{2c} \sum_i \left\langle \dot{\vec{\mu}}_i(0) \sum_j \left[\left(\vec{r}_j(t) + \frac{1}{2} \Delta_{ij}^{\text{PBC}}(0) \right) \times \dot{\vec{\mu}}_j(t) \right] \right\rangle \quad (2.86)$$

where Δ_{ij}^{PBC} is the conditional lattice translation used for the nearest-image convention between i and j . The first term is the sum of the correlation of moments within and between locally finite cells. These cells, generally corresponding to one molecule, each have a local origin. Partitioning the space as such corresponds to the distributed origin (DO) gauge [12]. The second term is the transformation from DO gauge to the common origin gauge, where all the particles have the same origin.

This equation was implemented in the ChirPy program [130], which was used to proceed from the time correlation functions to the corresponding spectra, as discussed in the next section.

ChirPy is a python package for analysing supramolecular and electronic structure, chirality and dynamics [130]. This package can, among other things, compute time-correlation functions (TCF) of moments obtained from MD trajectories. This package is used to compute the IR and VCD spectra from FFMD trajectories as the TCF function and the origin invariant version of the TCF function (C_{indep}) for the PBC case is already implemented.

2.8.2. From the time correlation functions to the spectra

The moments accumulated over the course of the trajectories were processed using the ChirPy program [130], from which the time correlation functions and resulting IR and VCD spectra were determined.

So far the equations given for the determination of the IR and VCD spectra through MD have been given in arbitrary units, as proportional to the Fourier transform of the appropriate TCF [equations (2.39) and (2.40)]. These correlation functions are classical, thus they are real and even, hence they obey $C_{cl}(-t) = C_{cl}(t)$. The Fourier transform conserves parity, giving the following relation for the classical spectral line shape: $I_{cl}(-\omega) = I_{cl}(\omega)$. However, the spectral line shape should satisfy the detailed balance condition:

$$I(\omega) = e^{\beta\hbar\omega} I(-\omega) \quad (2.87)$$

with $\beta = \frac{1}{k_B T}$, T being the temperature. To restore this condition, a frequency and temperature dependent prefactor is added: $I(\omega) = Q_{\text{QC}}(\omega) e^{\beta\hbar\omega} I(-\omega)$. This prefactor is often referred to as a quantum correction factor. Various corrections exist in the literature, [131, 132] the correction implemented in ChirPy is called the harmonic approximation :

$$Q_{\text{QC}}^{\text{HA}}(\omega) = \frac{\beta\hbar\omega}{1 - e^{-\beta\hbar\omega}} \quad (2.88)$$

Hence the line shape determined by ChirPy has the following expression [7], using Gaussian-centimetre-gram-second (Gaussian-cgs) units:

$$\begin{aligned}
 I_{\text{VCD}}(\omega) &= \frac{8\pi(1 - e^{-\beta\hbar\omega})}{3N\hbar cn(\omega)} \frac{\beta\hbar\omega}{(1 - e^{-\beta\hbar\omega})} \int_{-\infty}^{+\infty} \langle \dot{\vec{\mu}}(0) \cdot \vec{m}(\tau) \rangle e^{-i\omega\tau} d\tau \quad (2.89) \\
 &= \frac{8\pi\beta\omega}{3Ncn(\omega)} \int_{-\infty}^{+\infty} \langle \dot{\vec{\mu}}(0) \cdot \vec{m}(\tau) \rangle e^{-i\omega\tau} d\tau
 \end{aligned}$$

with N being the number of units, $n(\omega)$ the refraction index of the medium and c the velocity of light. The prefactor is valid for isotropic systems at thermal equilibrium [8] and is constructed through Fermi's golden rule for a transition probability between two states. The time correlation function is taken at the classical limit and Kubo-transformed [7, 8] to give equation (2.90).

It is now possible to obtain the line shape of a VCD spectrum from the electric and magnetic dipole moments determined during a force field molecular dynamics trajectory. The next step is to assess the conditions of the MD simulations needed to obtain correct spectra.

2.8.3. Assignment of simulation parameters: the case of trans-1-amino-2 indanol

AMOEBA force field parameters for trans-1-amino-2 indanol (trans-AI) had already been determined and tested previously [133]. It is thus a good molecule to try and implement the present methodology.

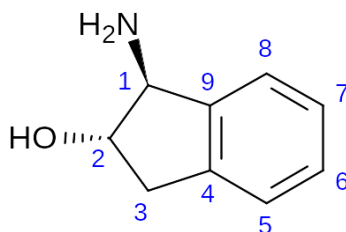


Figure 2.19: Structure of (1S,2S)-trans-1-amino-2 indanol and atom numbering.

The first test was to compare an IR spectrum obtained from a trajectory with a long sampling time and one obtained with an average over trajectories with a shorter sampling time. The spectra from a simulation of 1 ns were compared to one from 10 simulations of 100 ps. The MD trajectories were performed with a time step of 0.2 fs with a temperature of 150 K. The IR spectra obtained from these trajectories are shown in figure 2.20.

We find that the bands are generally defined with a higher number of peaks when averaging over multiple simulations especially just below 1600 cm^{-1} . This is because with multiple simulations multiple starting points can be chosen and also velocities are chosen at random in the beginning which improves sampling. So the simulation can explore a wider variety of configurations.

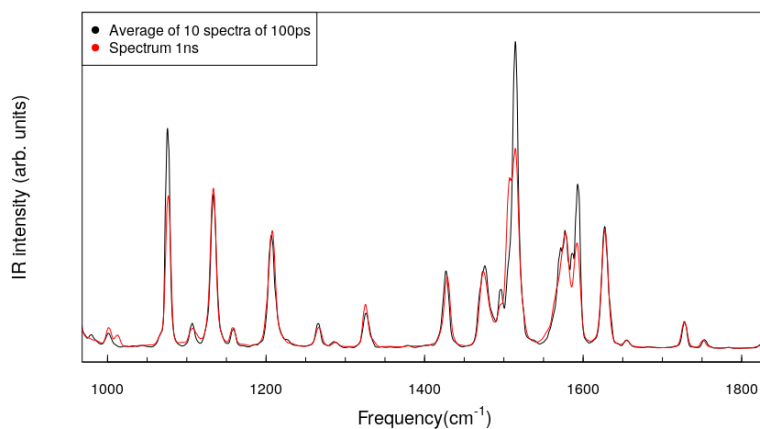


Figure 2.20: IR spectra of the trans-Al monomer in the gas phase from a FFMD trajectory at 300 K of 1 ns and from the average of 10 MD trajectories of 100 ps.

The goal of the second test was to determine the best time step to use for the MD simulation. For this, 5 simulations of 100 ps were performed at 150 K for the trans-Al monomer only changing the time step from 0.2 fs to 1.6 fs. The corresponding IR spectra are depicted in figure 2.21.

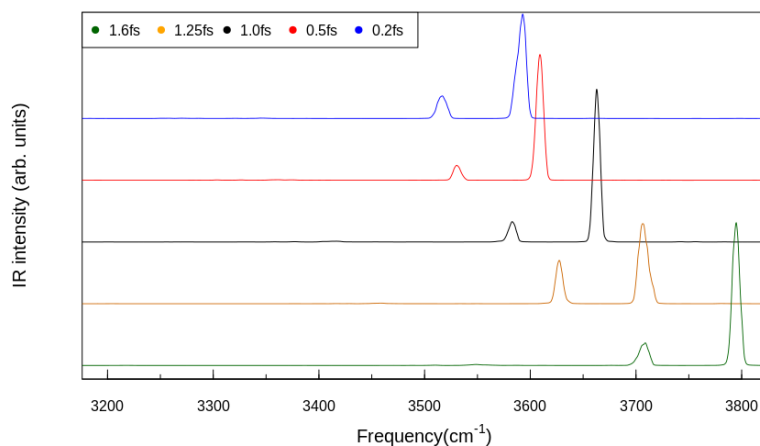


Figure 2.21: IR spectra of the trans-Al monomer from FFMD simulations at 300 K with different time steps.

The peak of the OH stretch, which is the vibration with the highest frequency for this molecule, is found experimentally at 3616 cm^{-1} . The blue spectrum using a time step of 0.2 fs reproduces this value. However as the value of the time step increases, a larger

shift from this value appears. We chose the value of 0.5 fs as it is a good compromise between accuracy and computational time. Horníček et al [9] reached the same conclusion, explaining that the error between the expected and actual values of the peak corresponding to the OH stretch for long time steps, originates from the Beeman integration during the molecular dynamics trajectory.

The molecular dynamics simulations and thus the dipoles moments are calculated with a time step of 0.5 fs however the time correlation functions sample the dipole moments with a time step of 4 fs. Using a time step of 4 fs for the time correlation functions gives the same line shape as with a time step of 0.5 fs but with a reduced computational cost.

Such a short time step is only needed for determining spectra. For exploring the potential energy surface a longer time step (1 fs) is used as it permits a longer simulation time and thus a better exploration of the surface.

The next goal is to determine the vibrational modes that correspond to the bands in the spectrum.

2.9. Assignment of spectral modes

In section 2.3.2, the assignment of the frequencies in the harmonic approximation through the diagonalisation of the mass-weighted Hessian matrix was discussed. However, the interest here is for time-dependent, anharmonic systems. The correlation functions do not provide any information about the assignment of the frequencies, namely spectral modes. This information is essential to compare theoretical and experimental spectra in order for reliable conclusions to be drawn.

2.9.1. Methods to obtain vibrational modes

Within the harmonic approximation, it is straightforward to assign all modes through the diagonalisation of the mass-weighted Hessian matrix of static systems. For dynamical systems, that do not obey the harmonic approximation, various methods exist to determine normal modes that extend the concept of eigenmodes in harmonic systems.

The first method developed by Wheeler et al. [134] is called principal modes analysis (PMA). The displacement of the mass-weighted molecular coordinates from their equilibrium position through a MD simulation is described in a covariance matrix. The eigenvalues and eigenvector of the inverse of this matrix are used to provide the frequencies and the normal modes.

A second method is the instantaneous normal mode analysis (INMA) where the mass-weighted Hessian matrix is diagonalised at each time step. However this can get expensive especially if the system is large as the diagonalisation, which is repeated, scales approximately as $O(N^3)$ with N the number of atoms. Furthermore, INMA must deal with multiple negative eigenvalues. [135, 136]

Bowman et al. [137] proposed an approach called driven molecular dynamics (DMD).

This method becomes competitive compared to INMA for systems with more than 1 000 atoms. In this method a sinusoidal driving force V_{DMD} is added to the Hamiltonian. This term is determined for each step of the dynamics as a function of the distance between the nuclei:

$$V_{\text{DMD}}(t) = \sum_{i,j}^N \lambda_{i,j} r_{i,j} \sin(\omega_n t) \quad (2.90)$$

where $r_{i,j}$ are internuclear distances, ω_n , called the driving frequency, is a fixed exciting frequency and $\lambda_{i,j}$ are the coupling constants. The driving frequency can be varied through a frequency range and energy absorption should maximize locally at the normal mode frequencies. An examination of the molecular motion at a given resonance provides the corresponding normal mode.

DMD was coupled with the determination of the IR spectra through the Fourier transform of the autocorrelation function by Thauay et al. [119]. Hence it was possible to give the corresponding normal modes of a spectrum that includes anharmonicity and temperature effects from the trajectory. This is demonstrated in figure 2.22 where the displacements of C=O and N-H of N-methyl-acetamide are shown for two resonant driving frequencies. At $\omega_n = 1723 \text{ cm}^{-1}$ the normal mode corresponding to the C=O stretch is shown and at $\omega_n = 3518 \text{ cm}^{-1}$ it is the N-H stretching mode.

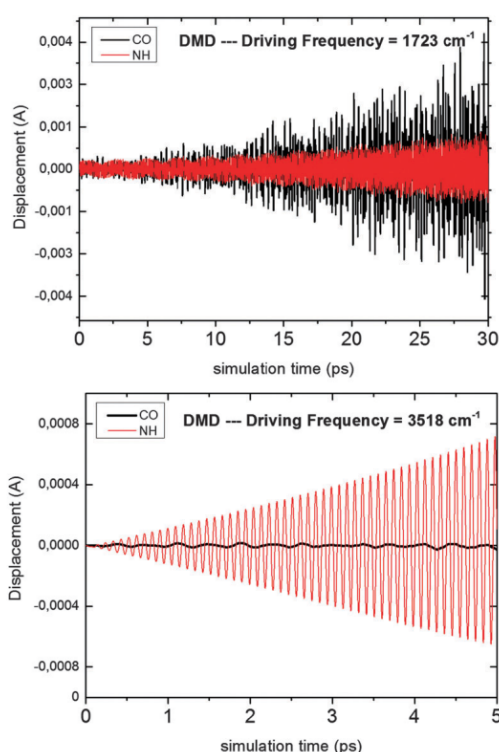


Figure 2.22: Monitoring the CO and N-H N-methyl-acetamide stretch modes by DMD simulations. This figure has been reprinted from Thauay et al. [119].

The Kohanoff procedure [138] extracts vibrational frequencies and eigenvectors from short non-thermally equilibrated molecular dynamics trajectories. The frequencies are estimated, then the eigenvectors are determined through a least squares fit of the trajectory including orthogonality constraints. The trajectory is projected onto each normal mode corresponding to one frequency.

2.9.2. Effective mode analysis

For DMD and the Kohanoff procedure, short MD trajectories need to be performed for each frequency, which is not the case for the last method presented here. It corresponds to an effective mode analysis [139, 140]. In this method, for each mode, an IR spectrum and the corresponding vibrations are determined from a trajectory of internal coordinates and velocities of the system. This method is useful if there is some prior knowledge of the physical chemistry of the system, which can be used to define the internal coordinates.

The total infrared spectrum $I_n(\omega)$ can be written as:

$$I_n(\omega) = \frac{\beta N_A}{6c\epsilon_0} \sum_{i=1}^{3N-6} \left\langle \left| \frac{\partial \vec{\mu}}{\partial q_i}(0) \right|^2 \right\rangle P_{ii}^{(q)}(\omega) \quad (2.91)$$

where, aside from the temperature-dependent prefactor, the index i runs over $3N - 6$ internal degrees of freedom of the molecule, $\vec{\mu}$ is the dipole moment of the whole molecule, q_i is the i -th effective mode (to be determined), and $P_{ii}^{(q)}(\omega)$ is the power spectrum, i.e., the velocity-velocity correlation function of the i -th effective mode

$$P_{ii}^{(q)}(\omega) = \int_{-\infty}^{+\infty} dt e^{-i\omega t} \langle \dot{q}_i(t) \dot{q}_i(0) \rangle \quad (2.92)$$

In equation (2.91) the IR spectrum is decomposed as the sum of independent contributions, each arising from a single effective mode i ; in addition, the effective modes are constructed such that their power spectra are as localized as possible in frequency. This feature is achieved by minimising the functional

$$\Omega^{(n)} = \sum_{i=1}^{3N-6} \left[\frac{\beta}{2\pi} \int_{-\infty}^{+\infty} d\omega \omega^{2n} P_{ii}^{(q)}(\omega) - \left(\frac{\beta}{2\pi} \int_{-\infty}^{+\infty} d\omega \omega^n P_{ii}^{(q)}(\omega) \right)^2 \right] \quad (2.93)$$

with respect to the elements of the transformation matrix \mathbf{Z}^{-1} from internal coordinates ζ to effective modes \mathbf{q} , namely

$$\mathbf{q} = \mathbf{Z}^{-1} \zeta \quad \text{or} \quad q_i = \sum_{k=1}^{3N-6} Z_{ik}^{-1} \zeta_k \quad (2.94)$$

The functional $\Omega^{(n)}$ represents the variance of the n -th moment of $P_{ii}^{(q)}(\omega)$ when interpreting $P_{ii}^{(q)}(\omega)$ as a probability distribution.

In order to derive equation (2.91) for the IR spectrum starting from the Fourier transform of the dipole-dipole time correlation function, the following hypotheses are made:

1. Positions and velocities are uncorrelated all along the trajectory;
2. Rotations and vibrations are uncorrelated, and the rotational correlation function decays much slower than the vibrational correlation function;
3. The correlation function for the derivatives of the dipole moment decays much slower than the vibrational correlation function, thus we can only consider its value at $t = 0$;
4. The derivatives of the dipole moment with respect to q_i can be determined from the atomic polar tensor in Cartesian coordinates, then transformed to internal coordinates, then transformed to effective modes;
5. The off-diagonal elements of the q -power spectrum are negligible if compared to the diagonal elements.

The VCD spectrum has peaks present at the same frequencies as the IR spectrum as there can be no VCD activity if there is no IR absorption. Thus by determining the modes for IR spectroscopy, the information can be transferred to VCD.

2.10. Concluding remarks

Throughout this chapter, various methods to compute vibrational circular dichroism spectra have been presented as well as procedures to assign molecular vibrations to spectral peaks.

After giving the general theory of VCD, first a static approach was discussed using DFT and ways to include anharmonicity were presented. Then molecular dynamics, which naturally accounts for anharmonic and temperature effects, was presented and how to achieve VCD spectroscopy from this method was shown. Subsequently two approaches using MD were examined. The first approach using a quantum chemical description of the electrons, and the other describing the system with a force field. The implementation of VCD spectroscopy into force field MD was then explained, including the difficulties in periodic boundary conditions, how the spectra were retrieved from electric and magnetic dipole moments, and the importance of the time step of the MD simulation. It is this magnetic dipole moment that was implemented in a force field MD program during this work and that will be assessed in chapters 4 and 5, dealing with systems in the gas and condensed phases, respectively. Finally methods to determine spectral modes from MD simulations were presented.

Bibliography

- [1] S. S. Andrews and J. Tretton, "Physical principles of circular dichroism," *Journal of Chemical Education*, vol. 97, no. 12, pp. 4370–4376, 2020.
- [2] D. J. Griffiths, *Introduction to electrodynamics*. 1900 E. Lake Ave., Glenview, IL 60025: Pearson, 4th ed., 2014.
- [3] J. J. Sakurai, *Modern Quantum Mechanics*. University Printing House, Cambridge: Cambridge University Press, 3rd ed., 2021.
- [4] S. S. Andrews, "Using rotational averaging to calculate the bulk response of isotropic and anisotropic samples from molecular parameters," *Journal of chemical education*, vol. 81, no. 6, p. 877, 2004.
- [5] L. Rosenfeld, "Quantum-mechanical theory of the natural optical activity of liquids and gases," *Zeitschrift für Physik*, vol. 52, pp. 161–174, 1928.
- [6] S. Yang and M. Cho, "Direct calculations of vibrational absorption and circular dichroism spectra of alanine dipeptide analog in water: Quantum mechanical-molecular mechanical molecular dynamics simulations," *The Journal of chemical physics*, vol. 131, no. 13, p. 10B605, 2009.
- [7] S. Jähnigen, A. Zehnacker, and R. Vuilleumier, "Computation of solid-state vibrational circular dichroism in the periodic gauge," *The Journal of Physical Chemistry Letters*, vol. 12, pp. 7213–7220, 2021.
- [8] A. Scherrer, R. Vuilleumier, and D. Sebastiani, "Vibrational circular dichroism from ab initio molecular dynamics and nuclear velocity perturbation theory in the liquid phase," *The Journal of Chemical Physics*, vol. 145, no. 8, p. 084101, 2016.
- [9] J. Horníček, P. Kaprálová, and P. Bouř, "Simulations of vibrational spectra from classical trajectories: Calibration with ab initio force fields," *The Journal of Chemical Physics*, vol. 127, no. 8, p. 084502, 2007.
- [10] F. Thaunay, *Développement de champs de forces polarisables et applications à la spectroscopie vibrationnelle*. PhD thesis, Ecole Polytechnique, 2016.
- [11] P. Bayley, "The analysis of circular dichroism of biomolecules," *Progress in Biophysics and Molecular Biology*, vol. 27, pp. 1–76, 1973.
- [12] L. A. Nafie, *Vibrational optical activity: principles and applications*. The Atrium, Southern Gate, Chichester, West Sussex: John Wiley & Sons, 2011.
- [13] T. eChem team, *Property gradients*, 2022. https://kthpanor.github.io/echem/docs/property_grad.html#dipole-mom-gradient-label (accessed June 28th, 2022).
- [14] Y. Cornaton, M. Ringholm, O. Louant, and K. Ruud, "Analytic calculations of anharmonic infrared and raman vibrational spectra," *Physical Chemistry Chemical Physics*, vol. 18, no. 5, pp. 4201–4215, 2016.

- [15] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Physical Review*, vol. 136, no. 3B, p. B864, 1964.
- [16] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Physical Review*, vol. 140, no. 4A, p. A1133, 1965.
- [17] A. D. Becke, "Density-functional thermochemistry. III. the role of exact exchange," *Journal of Chemical Physics*, vol. 98, pp. 5648–52, 1993.
- [18] A. J. Cohen and N. C. Handy, "Dynamic correlation," *Molecular Physics*, vol. 99, pp. 607–15, 2001.
- [19] A. Austin, G. Petersson, M. J. Frisch, F. J. Dobek, G. Scalmani, and K. Throssell, "A density functional with spherical atom dispersion terms," *Journal of Chemical Theory and Computation*, vol. 8, p. 4989, 2012.
- [20] T. M. Henderson, A. F. Izmaylov, G. Scalmani, and G. E. Scuseria, "Can short-range hybrids describe long-range-dependent properties?," *Journal of Chemical Physics*, vol. 131, p. 044108, 2009.
- [21] T. Yanai, D. Tew, and N. Handy, "A new hybrid exchange-correlation functional using the coulomb-attenuating method (CAM-B3LYP)," *Chemical Physics Letters*, vol. 393, pp. 51–57, 2004.
- [22] J.-D. Chai and M. Head-Gordon, "Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections," *Chemical physics and chemical physics*, vol. 10, pp. 6615–20, 2008.
- [23] Y. Zhao and D. 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*, vol. 120, pp. 215–41, 2008.
- [24] C. Adamo and V. Barone, "Toward reliable density functional methods without adjustable parameters: The PBE0 model," *Journal of Chemical Physics*, vol. 110, pp. 6158–69, 1999.
- [25] J. P. Perdew and K. Schmidt, "Jacob's ladder of density functional approximations for the exchange-correlation energy," in *AIP Conference Proceedings*, vol. 577, pp. 1–20, American Institute of Physics, 2001.
- [26] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, "Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields," *The Journal of Physical Chemistry*, vol. 98, no. 45, pp. 11623–11627, 1994.
- [27] I. Alata, A. Pérez-Mellor, F. Ben Nasr, D. Scuderi, V. Steinmetz, F. Gobert, N.-E. Jaïdane, and A. Zehnacker-Rentien, "Does the residues chirality modify the conformation of a cyclo-dipeptide? vibrational spectroscopy of protonated cyclo-diphenylalanine in the gas phase," *The Journal of Physical Chemistry A*, vol. 121, no. 38, pp. 7130–7138, 2017.

- [28] K. Le Barbu-Debus, A. Scherrer, A. Bouchet, D. Sebastiani, R. Vuilleumier, and A. Zehnacker, "Effect of puckering motion and hydrogen bond formation on the vibrational circular dichroism spectrum of a flexible molecule: the case of (S)-1-indanol," *Physical Chemistry Chemical Physics*, vol. 20, no. 21, pp. 14635–14646, 2018.
- [29] S. Grimme, S. Ehrlich, and L. Goerigk, "Effect of the damping function in dispersion corrected density functional theory," *Journal of Computational Chemistry*, vol. 32, no. 7, pp. 1456–1465, 2011.
- [30] T. Schwabe, R. Huenerbein, and S. Grimme, "Large molecules-small energies: Challenges for contemporary quantum chemistry," *Synlett*, vol. 2010, no. 10, pp. 1431–1441, 2010.
- [31] T. Schwabe and S. Grimme, "Theoretical thermodynamics for large molecules: walking the thin line between accuracy and computational cost," *Accounts of Chemical Research*, vol. 41, no. 4, pp. 569–579, 2008.
- [32] J. Cheeseman, M. Frisch, F. Devlin, and P. Stephens, "Ab initio calculation of atomic axial tensors and vibrational rotational strengths using density functional theory," *Chemical Physics Letters*, vol. 252, no. 3-4, pp. 211–220, 1996.
- [33] T. B. Freedman, K. M. Spencer, N. Ragunathan, L. A. Nafie, J. A. Moore, and J. M. Schwab, "Vibrational circular dichroism of (S, S)-[2,3-²H₂] oxirane in the gas phase and in solution," *Canadian Journal of Chemistry*, vol. 69, no. 11, pp. 1619–1629, 1991.
- [34] J. Simons, J. Nichols, *et al.*, *Quantum mechanics in chemistry*. Oxford University Press New York, 1997.
- [35] T. Barakat, K. Abodayeh, and O. Al-Dossary, "Exact solutions for vibrational levels of the morse potential via the asymptotic iteration method," *Czechoslovak Journal of Physics*, vol. 56, no. 6, pp. 583–590, 2006.
- [36] R. Kiselev, *Towards clinical translation of raman spectroscopy for tumor cell identification*. PhD thesis, Friedrich-Schiller-Universität Jena, Jena, 2019.
- [37] M. P. Andersson and P. Uvdal, "New scale factors for harmonic vibrational frequencies using the B3LYP density functional method with the triple- ζ basis set 6-311+G(d, p)," *The Journal of Physical Chemistry A*, vol. 109, no. 12, pp. 2937–2941, 2005.
- [38] C. A. Jiménez-Hoyos, B. G. Janesko, and G. E. Scuseria, "Evaluation of range-separated hybrid density functionals for the prediction of vibrational frequencies, infrared intensities, and Raman activities," *Physical Chemistry Chemical Physics*, vol. 10, no. 44, pp. 6621–6629, 2008.
- [39] I. Alecu, J. Zheng, Y. Zhao, and D. G. Truhlar, "Computational thermochemistry: scale factor databases and scale factors for vibrational frequencies obtained from electronic model chemistries," *Journal of Chemical Theory and Computation*, vol. 6, no. 9, pp. 2872–2887, 2010.

- [40] J. P. Merrick, D. Moran, and L. Radom, "An evaluation of harmonic vibrational frequency scale factors," *The Journal of Physical Chemistry A*, vol. 111, no. 45, pp. 11683–11700, 2007.
- [41] National Institute of Standards and Technology, *Precomputed vibrational scaling factors*, 2022. <http://cccbdb.nist.gov/vibscalejust.asp> (accessed June 15th, 2022).
- [42] V. Barone, "Anharmonic vibrational properties by a fully automated second-order perturbative approach," *The Journal of Chemical Physics*, vol. 122, no. 1, p. 014108, 2005.
- [43] A. S. Perera, J. Thomas, M. R. Poopari, and Y. Xu, "The clusters-in-a-liquid approach for solvation: new insights from the conformer specific gas phase spectroscopy and vibrational optical activity spectroscopy," *Frontiers in Chemistry*, vol. 4, p. 9, 2016.
- [44] W. H. Miller, R. Hernandez, N. C. Handy, D. Jayatilaka, and A. Willetts, "Ab initio calculation of anharmonic constants for a transition state, with application to semi-classical transition state tunneling probabilities," *Chemical Physics Letters*, vol. 172, no. 1, pp. 62–68, 1990.
- [45] V. Barone, M. Biczysko, J. Bloino, M. Borkowska-Panek, I. Carnimeo, and P. Panek, "Toward anharmonic computations of vibrational spectra for large molecular systems," *International Journal of Quantum Chemistry*, vol. 112, no. 9, pp. 2185–2200, 2012.
- [46] A. Foldes and C. Sandorfy, "Anharmonicity and hydrogen bonding: Part III. examples of strong bonds. General discussion," *Journal of Molecular Spectroscopy*, vol. 20, no. 3, pp. 262–275, 1966.
- [47] J. Tomasi, B. Mennucci, and R. Cammi, "Quantum mechanical continuum solvation models," *Chemical Reviews*, vol. 105, no. 8, pp. 2999–3094, 2005.
- [48] R. Cammi, *The PCM Model*, pp. 1–11. Cham: Springer International Publishing, 2013.
- [49] B. Mennucci, E. Cancès, and J. Tomasi, "Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: theoretical bases, computational implementation, and numerical applications," *The Journal of Physical Chemistry B*, vol. 101, no. 49, pp. 10506–10517, 1997.
- [50] C. Merten, C. J. Berger, R. McDonald, and Y. Xu, "Evidence of dihydrogen bonding of a chiral amine–borane complex in solution by VCD spectroscopy," *Angewandte Chemie International Edition*, vol. 53, no. 37, pp. 9940–9943, 2014.
- [51] T. Richardson, S. de Gala, R. H. Crabtree, and P. E. Siegbahn, "Unconventional hydrogen bonds: intermolecular BH ··· HN interactions," *Journal of the American Chemical Society*, vol. 117, no. 51, pp. 12875–12876, 1995.
- [52] L. Weirich and C. Merten, "Solvation and self-aggregation of chiral alcohols: how hydrogen bonding affects their VCD spectral signatures," *Physical Chemistry Chemical Physics*, vol. 21, no. 25, pp. 13494–13503, 2019.

- [53] A. S. Perera, J. Cheramy, C. Merten, J. Thomas, and Y. Xu, "IR, Raman, and vibrational optical activity spectra of methyl glycidate in chloroform and water: The clusters-in-a-liquid solvation model," *ChemPhysChem*, vol. 19, no. 17, pp. 2234–2242, 2018.
- [54] S. Ghidinelli, S. Abbate, J. Koshoubu, Y. Araki, T. Wada, and G. Longhi, "Solvent effects and aggregation phenomena studied by vibrational optical activity and molecular dynamics: the case of pantolactone," *The Journal of Physical Chemistry B*, vol. 124, no. 22, pp. 4512–4526, 2020.
- [55] A. R. Leach, *Molecular modelling: principles and applications*. Edinburgh Gate, Harlow, Essex: Pearson education, 2001.
- [56] D. Beeman, "Some multistep methods for use in molecular dynamics calculations," *Journal of Computational Physics*, vol. 20, no. 2, pp. 130–139, 1976.
- [57] L.-V. Woodcock, "Isothermal molecular dynamics calculations for liquid salts," *Chemical Physics Letters*, vol. 10, no. 3, pp. 257–261, 1971.
- [58] H. J. Berendsen, J. v. Postma, W. F. Van Gunsteren, A. DiNola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," *The Journal of chemical Physics*, vol. 81, no. 8, pp. 3684–3690, 1984.
- [59] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *The Journal of Chemical Physics*, vol. 126, no. 1, p. 014101, 2007.
- [60] H. C. Andersen, "Molecular dynamics simulations at constant pressure and/or temperature," *The Journal of chemical physics*, vol. 72, no. 4, pp. 2384–2393, 1980.
- [61] S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods," *The Journal of Chemical Physics*, vol. 81, no. 1, pp. 511–519, 1984.
- [62] W. G. Hoover, "Canonical dynamics: Equilibrium phase-space distributions," *Physical Review A*, vol. 31, no. 3, p. 1695, 1985.
- [63] N. Metropolis, "The beginning of the Monte Carlo method," tech. rep., Los Alamos Science, 1987.
- [64] N. Berova, K. Nakanishi, and R. W. Woody, *Circular dichroism: principles and applications*. 605 Third Avenue, New York, NY: John Wiley & Sons, 2000.
- [65] J. L. McHale, *Molecular spectroscopy*. 6000 Broken Sound Parkway NW, Suite 300 Boca Raton: CRC Press, 2017.
- [66] S. Abbate, G. Longhi, K. Kwon, and A. Moscovitz, "The use of cross-correlation functions in the analysis of circular dichroism spectra," *The Journal of Chemical Physics*, vol. 108, no. 1, pp. 50–62, 1998.
- [67] R. Gordon, "Correlation functions for molecular motion," in *Advances in Magnetic and Optical Resonance*, vol. 3, pp. 1–42, 24–28 Oval road, London: Elsevier, 1968.
- [68] D. A. McQuarrie, *Statistical mechanics*. Harper's chemistry series, New York Evanston San Francisco : Harper and Row, 1976.

- [69] F. Calvo, "Atomistic modeling of IR action spectra under circularly polarized electromagnetic fields: Toward action VCD spectra," *Chirality*, vol. 27, no. 3, pp. 253–261, 2015.
- [70] K. Kwac, K.-K. Lee, J. B. Han, K.-I. Oh, and M. Cho, "Classical and quantum mechanical/molecular mechanical molecular dynamics simulations of alanine dipeptide in water: Comparisons with IR and vibrational circular dichroism spectra," *The Journal of Chemical Physics*, vol. 128, no. 10, p. 03B606, 2008.
- [71] T. Giovannini, M. Olszówka, and C. Cappelli, "Effective fully polarizable QM/MM approach to model vibrational circular dichroism spectra of systems in aqueous solution," *Journal of Chemical Theory and Computation*, vol. 12, no. 11, pp. 5483–5492, 2016.
- [72] J.-H. Choi and M. Cho, "Direct calculations of mid-and near-IR absorption and circular dichroism spectra of chiral molecules using QM/MM molecular dynamics simulation method," *Journal of Chemical Theory and Computation*, vol. 7, no. 12, pp. 4097–4103, 2011.
- [73] A. Scherrer, R. Vuilleumier, and D. Sebastiani, "Nuclear velocity perturbation theory of vibrational circular dichroism," *Journal of Chemical Theory and Computation*, vol. 9, no. 12, pp. 5305–5312, 2013.
- [74] M. Thomas and B. Kirchner, "Classical magnetic dipole moments for the simulation of vibrational circular dichroism by ab initio molecular dynamics," *The Journal of Physical Chemistry Letters*, vol. 7, no. 3, pp. 509–513, 2016.
- [75] E. Ditler, T. Zimmermann, C. Kumar, and S. Lubner, "Implementation of nuclear velocity perturbation and magnetic field perturbation theory in CP2K and their application to vibrational circular dichroism," *Journal of Chemical Theory and Computation*, vol. 18, no. 4, pp. 2448–2461, 2022.
- [76] R. Car and M. Parrinello, "Unified approach for molecular dynamics and density-functional theory," *Physical Review Letters*, vol. 55, no. 22, p. 2471, 1985.
- [77] R. Iftimie, P. Minary, and M. E. Tuckerman, "Ab initio molecular dynamics: Concepts, recent developments, and future trends," *Proceedings of the National Academy of Sciences*, vol. 102, no. 19, pp. 6654–6659, 2005.
- [78] D. Marx and J. Hutter, *Ab initio molecular dynamics: basic theory and advanced methods*. The Edinburgh Building, Cambridge: Cambridge University Press, 2009.
- [79] L. A. Nafie and T. B. Freedman, "Vibronic coupling theory of infrared vibrational transitions," *The Journal of Chemical Physics*, vol. 78, no. 12, pp. 7108–7116, 1983.
- [80] L. A. Nafie, "Adiabatic molecular properties beyond the Born-Oppenheimer approximation. complete adiabatic wave functions and vibrationally induced electronic current density," *The Journal of Chemical Physics*, vol. 79, no. 10, pp. 4950–4957, 1983.
- [81] A. Scherrer, F. Agostini, D. Sebastiani, E. Gross, and R. Vuilleumier, "Nuclear velocity

- perturbation theory for vibrational circular dichroism: An approach based on the exact factorization of the electron-nuclear wave function," *The Journal of Chemical Physics*, vol. 143, no. 7, p. 074106, 2015.
- [82] CPMD, Copyright IBM Corp 1990-2019, Copyright MPI für Festkörperforschung Stuttgart 1997-2001., 2022. <http://www.cpmc.org/> (accessed April 27th, 2022).
- [83] M. Losada, P. Nguyen, and Y. Xu, "Solvation of propylene oxide in water: Vibrational circular dichroism, optical rotation, and computer simulation studies," *The Journal of Physical Chemistry A*, vol. 112, no. 25, pp. 5621–5627, 2008.
- [84] M. Losada and Y. Xu, "Chirality transfer through hydrogen-bonding: experimental and ab initio analyses of vibrational circular dichroism spectra of methyl lactate in water," *Physical Chemistry Chemical Physics*, vol. 9, no. 24, pp. 3127–3135, 2007.
- [85] M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith, and D. Marx, "Dissecting the THz spectrum of liquid water from first principles via correlations in time and space," *Proceedings of the National Academy of Sciences*, vol. 107, no. 27, pp. 12068–12073, 2010.
- [86] J. E. Bertie, S. L. Zhang, and C. D. Keefe, "Infrared intensities of liquids XVI. accurate determination of molecular band intensities from infrared refractive index and dielectric constant spectra," *Journal of Molecular Structure*, vol. 324, no. 1-2, pp. 157–176, 1994.
- [87] N. Foloppe and A. D. MacKerell, Jr, "All-atom empirical force field for nucleic acids: I. parameter optimization based on small molecule and condensed phase macromolecular target data," *Journal of Computational Chemistry*, vol. 21, no. 2, pp. 86–104, 2000.
- [88] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules," *Journal of the American Chemical Society*, vol. 117, no. 19, pp. 5179–5197, 1995.
- [89] W. L. Jorgensen, D. S. Maxwell, and J. 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*, vol. 118, no. 45, pp. 11225–11236, 1996.
- [90] W. R. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, and W. F. van Gunsteren, "The GROMOS biomolecular simulation program package," *The Journal of Physical Chemistry A*, vol. 103, no. 19, pp. 3596–3607, 1999.
- [91] N. Allinger, "Calculation of molecular structure and energy by force-field methods," in *Advances in Physical Organic Chemistry*, vol. 13, pp. 1–82, Elsevier, 1976.
- [92] M. R. Frierson, M. R. Imam, V. B. Zalkow, and N. L. Allinger, "The MM2 force field for silanes and polysilanes," *The Journal of Organic Chemistry*, vol. 53, no. 22, pp. 5248–5258, 1988.

- [93] N. L. Allinger, Y. H. Yuh, and J. H. Lii, "Molecular mechanics. the MM3 force field for hydrocarbons.," *Journal of the American Chemical Society*, vol. 111, no. 23, pp. 8551–8566, 1989.
- [94] C. H. Langley and N. L. Allinger, "Molecular mechanics (MM4) and ab initio study of amide- amide and amide- water dimers," *The Journal of Physical Chemistry A*, vol. 107, no. 26, pp. 5208–5216, 2003.
- [95] J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio Jr, *et al.*, "Current status of the AMOEBA polarizable force field," *The Journal of Physical Chemistry B*, vol. 114, no. 8, pp. 2549–2564, 2010.
- [96] M. J. Lépinay, *Impact des chimies de nettoyage et des traitements plasma sur les matériaux diélectriques à basse permittivité*. PhD thesis, Université Montpellier II-Sciences et Techniques du Languedoc, 2014.
- [97] Dna-Dennis, *Coulombs Law*, 2008. <https://commons.wikimedia.org/wiki/File:CoulombsLaw.svg> (accessed June 3rd, 2022).
- [98] Z. Jing, C. Liu, S. Y. Cheng, R. Qi, B. D. Walker, J.-P. Piquemal, and P. Ren, "Polarizable force fields for biomolecular simulations: Recent advances and applications," *Annual Review of Biophysics*, vol. 48, pp. 371–394, 2019.
- [99] J. Huang, A. C. Simmonett, F. C. Pickard IV, A. D. MacKerell Jr, and B. R. Brooks, "Mapping the drude polarizable force field onto a multipole and induced dipole model," *The Journal of Chemical Physics*, vol. 147, no. 16, p. 161702, 2017.
- [100] P. E. Lopes, O. Guvench, and A. D. MacKerell, "Current status of protein force fields for molecular dynamics simulations," in *Molecular modeling of proteins*, pp. 47–71, Springer, 2015.
- [101] D. Semrouni, A. Sharma, J.-P. Dognon, G. Ohanessian, and C. Clavaguéra, "Finite temperature infrared spectra from polarizable molecular dynamics simulations," *Journal of Chemical Theory and Computation*, vol. 10, no. 8, pp. 3190–3199, 2014.
- [102] F. Thauay, C. Jana, C. Clavaguéra, and G. Ohanessian, "Strategy for modeling the infrared spectra of ion-containing water drops," *The Journal of Physical Chemistry A*, vol. 122, no. 3, pp. 832–842, 2018.
- [103] S. W. Rick, S. J. Stuart, and B. J. Berne, "Dynamical fluctuating charge force fields: Application to liquid water," *The Journal of Chemical Physics*, vol. 101, no. 7, pp. 6141–6156, 1994.
- [104] S. W. Rick and S. Stuart, "Potentials and algorithms for incorporating polarizability in computer simulations," *Reviews in Computational Chemistry*, vol. 18, pp. 89–146, 2003.
- [105] P. Cieplak, F.-Y. Dupradeau, Y. Duan, and J. Wang, "Polarization effects in molecular mechanical force fields," *Journal of Physics: Condensed Matter*, vol. 21, no. 33, p. 333102, 2009.

- [106] C. J. Illingworth and C. Domene, "Many-body effects and simulations of potassium channels," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 465, no. 2106, pp. 1701–1716, 2009.
- [107] C. Bret, M. J. Field, and L. Hemmingsen, "A chemical potential equalization model for treating polarization in molecular mechanical force fields," *Molecular Physics*, vol. 98, no. 11, pp. 751–763, 2000.
- [108] J. W. Caldwell and P. A. Kollman, "Structure and properties of neat liquids using nonadditive molecular dynamics: water, methanol, and n-methylacetamide," *The Journal of Physical Chemistry*, vol. 99, no. 16, pp. 6208–6219, 1995.
- [109] P. Cieplak, P. Kollman, and T. Lybrand, "A new water potential including polarization: Application to gas-phase, liquid, and crystal properties of water," *The Journal of Chemical Physics*, vol. 92, no. 11, pp. 6755–6760, 1990.
- [110] L. X. Dang, "Importance of polarization effects in modeling the hydrogen bond in water using classical molecular dynamics techniques," *The Journal of Physical Chemistry B*, vol. 102, no. 3, pp. 620–624, 1998.
- [111] J. W. Ponder and D. A. Case, "Force fields for protein simulations," *Advances in protein chemistry*, vol. 66, pp. 27–85, 2003.
- [112] P. Ren and J. W. Ponder, "Polarizable atomic multipole water model for molecular mechanics simulation," *The Journal of Physical Chemistry B*, vol. 107, no. 24, pp. 5933–5947, 2003.
- [113] T. Walsh and T. Liang, "A multipole-based water potential with implicit polarization for biomolecular simulations," *Journal of computational chemistry*, vol. 30, no. 6, pp. 893–899, 2009.
- [114] Central Michigan university, *Model Box Periodic Boundary Conditions*, 2022. <http://isaacs.sourceforge.net/phys/psc.html> (accessed July 11th, 2022).
- [115] A. Y. Toukmaji and J. A. Board Jr, "Ewald summation techniques in perspective: a survey," *Computer Physics Communications*, vol. 95, no. 2-3, pp. 73–92, 1996.
- [116] C. M. Roth, B. L. Neal, and A. M. Lenhoff, "Van der waals interactions involving proteins," *Biophysical Journal*, vol. 70, no. 2, pp. 977–987, 1996.
- [117] T. D. Rasmussen, P. Ren, J. W. Ponder, and F. Jensen, "Force field modeling of conformational energies: importance of multipole moments and intramolecular polarization," *International Journal of Quantum Chemistry*, vol. 107, no. 6, pp. 1390–1395, 2007.
- [118] J. Kaminsky and F. Jensen, "Force field modeling of amino acid conformational energies," *Journal of Chemical Theory and Computation*, vol. 3, no. 5, pp. 1774–1788, 2007.
- [119] F. Thauay, J.-P. Dognon, G. Ohanessian, and C. Clavaguéra, "Vibrational mode assignment of finite temperature infrared spectra using the amoeba polarizable force field," *Physical Chemistry Chemical Physics*, vol. 17, no. 39, pp. 25968–25977, 2015.

- [120] Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren, "Polarizable atomic multipole-based amoeba force field for proteins," *Journal of Chemical Theory and Computation*, vol. 9, no. 9, pp. 4046–4063, 2013.
- [121] A. J. Stone, "Distributed multipole analysis, or how to describe a molecular charge distribution," *Chemical Physics Letters*, vol. 83, no. 2, pp. 233–239, 1981.
- [122] E. B. Wilson, J. C. Decius, and P. C. Cross, *Molecular vibrations: the theory of infrared and Raman vibrational spectra*. Courier Corporation, 1980.
- [123] T. A. Halgren, "The representation of van der waals (vdW) interactions in molecular mechanics force fields: potential form, combination rules, and vdW parameters," *Journal of the American Chemical Society*, vol. 114, no. 20, pp. 7827–7843, 1992.
- [124] D. M. Young, *Iterative solution of large linear systems*. 24-28 Oval road, London: Elsevier, 2014.
- [125] J. Applequist, J. R. Carl, and K.-K. Fung, "Atom dipole interaction model for molecular polarizability. Application to polyatomic molecules and determination of atom polarizabilities," *Journal of the American Chemical Society*, vol. 94, no. 9, pp. 2952–2960, 1972.
- [126] B. T. Thole, "Molecular polarizabilities calculated with a modified dipole interaction," *Chemical Physics*, vol. 59, no. 3, pp. 341–350, 1981.
- [127] P. Ren and J. W. Ponder, "Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations," *Journal of Computational Chemistry*, vol. 23, no. 16, pp. 1497–1506, 2002.
- [128] J. W. Ponder, *TINKER - Software Tools for Molecular Design (version 8)*, 2019. <http://dasher.wustl.edu/tinker> (accessed July 22th, 2020).
- [129] V. Hnizdo, "Magnetic dipole moment of a moving electric dipole," *American Journal of Physics*, vol. 80, no. 7, pp. 645–647, 2012.
- [130] S. Jähnigen, "Chirpy - A python package for analysing supramolecular and electronic structure, chirality and dynamics," 2021. <https://doi.org/10.5281/zenodo.4775330>.
- [131] S. Egorov and J. Skinner, "Semiclassical approximations to quantum time correlation functions," *Chemical Physics Letters*, vol. 293, no. 5-6, pp. 469–476, 1998.
- [132] R. Ramirez, T. López-Ciudad, P. Kumar P, and D. Marx, "Quantum corrections to classical time-correlation functions: Hydrogen bonding and anharmonic floppy modes," *The Journal of Chemical Physics*, vol. 121, no. 9, pp. 3973–3983, 2004.
- [133] K. Le Barbu-Debus, J. Bowles, S. Jähnigen, C. Clavaguéra, F. Calvo, R. Vuilleumier, and A. Zehnacker, "Assessing cluster models of solvation for the description of vibrational circular dichroism spectra: synergy between static and dynamic approaches," *Physical Chemistry Chemical Physics*, vol. 22, no. 45, pp. 26047–26068, 2020.

- [134] R. A. Wheeler, H. Dong, and S. E. Boesch, "Quasiharmonic vibrations of water, water dimer, and liquid water from principal component analysis of quantum or QM/MM trajectories," *ChemPhysChem*, vol. 4, no. 4, pp. 382–384, 2003.
- [135] M. Nonella, G. Mathias, M. Eichinger, and P. Tavan, "Structures and vibrational frequencies of the quinones in Rb. sphaeroides derived by a combined density functional/molecular mechanics approach," *The Journal of Physical Chemistry B*, vol. 107, no. 1, pp. 316–322, 2003.
- [136] L. Pejov, D. Spångberg, and K. Hermansson, "Using MD snapshots in ab initio and DFT calculations: OH vibrations in the first hydration shell around Li^+ (aq)," *The Journal of Physical Chemistry A*, vol. 109, no. 23, pp. 5144–5152, 2005.
- [137] J. M. Bowman, X. Zhang, and A. Brown, "Normal-mode analysis without the hessian: A driven molecular-dynamics approach," *The Journal of Chemical Physics*, vol. 119, no. 2, pp. 646–650, 2003.
- [138] J. Kohanoff, "Phonon spectra from short non-thermally equilibrated molecular dynamics simulations," *Computational Materials Science*, vol. 2, no. 2, pp. 221–232, 1994.
- [139] M. Martinez, M.-P. Gaigeot, D. Borgis, and R. Vuilleumier, "Extracting effective normal modes from equilibrium dynamics at finite temperature," *The Journal of Chemical Physics*, vol. 125, no. 14, p. 144106, 2006.
- [140] F. Agostini, R. Vuilleumier, and G. Ciccotti, "Infrared spectroscopy of small protonated water clusters at room temperature: an effective modes analysis," *The Journal of Chemical Physics*, vol. 134, no. 8, p. 084302, 2011.

Chapter 3

Trans-1-amino-2-indanol: a case study for VCD and fluxionality

3.1. Introduction

In experimental VCD, the molecules of interest are usually solvated and one fundamental issue is to assess the role of the solvent on the molecular conformation and on the spectrum. Successful assignment of VCD signals based on computed spectra therefore requires that the solvent is correctly described, and that the spectra are calculated with sufficient accuracy.

Sufficient sampling of the conformational space is also important to capture the effects on the spectra from the various conformations. These conformations can be found by chemical intuition, a random-search method for finding low-energy conformations like the Monte Carlo Multiple Minimum method [1] or through molecular dynamics simulations.

3.1.1. Systems of interest

Only little attention has been paid so far to the influence of solvation on the VCD spectrum of 1,2-amino-alcohols. Tarczay and coworkers demonstrated the suitability of matrix isolation VCD spectroscopy for determining the absolute configuration of fluxional molecules that have strong intermolecular interactions, especially hydrogen bonds [2]. Matrix isolation involves condensing the substance to be studied with a large excess of inert gas (usually argon or nitrogen) at low temperature to form a rigid solid (the matrix) [3].

The aim of this chapter is to unravel the interplay between the intramolecular structure and the solvation network, and explore how the solvation dynamics affects the VCD spectrum of 1-amino-2-indanol, depicted in figure 3.1.

The cis isomer (cis-AI) was explored employing a Monte Carlo Multiple Minimum method and the cluster-in-a-liquid model. This isomer was studied in reference [4]. For the trans isomer (trans-AI), polarisable force field molecular dynamics (FFMD) was used to extensively explore the potential energy surface of the monomer as well as that of its complexes with one and two DMSO molecules. Being designed for bulk DMSO, the AMOEBA FF used is expected to reproduce satisfactorily the solvation trends observed in the system. This isomer was studied during the present work.

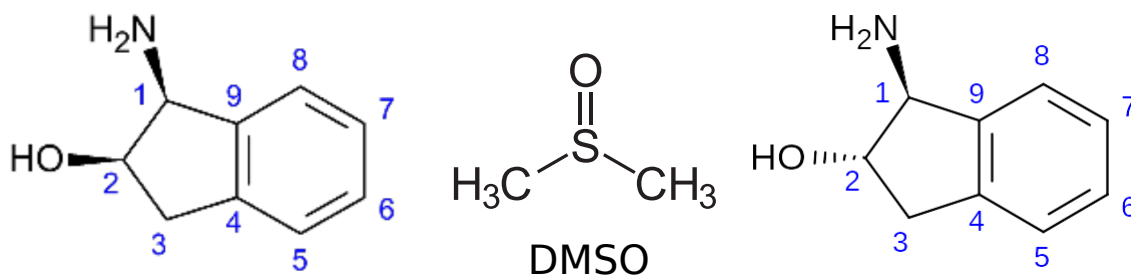


Figure 3.1: Structure of (1 *S*, 2 *R*)-(-)-cis-1-amino-2-indanol (cis-AI), of the solvent, dimethyl sulfoxide (DMSO) and of (1*S*,2*S*)-trans-1-amino-2-indanol (trans-AI).

Representative structures were extracted and further optimised at the density-functional theory level to compute IR absorption and VCD harmonic spectra, in the frame of the cluster-in-a-liquid model, with the contribution from our collaborators Anne Zehnacker and Katia Le Barbu-Debus. Special attention was paid to the factors that influence the reconstructed spectra, in particular through Boltzmann weights associated with each cluster.

Furthermore, FFMD at fixed temperature was used to explore the potential energy surfaces of the trans-AI molecule solvated by one, two and five DMSO molecules, and compare them to the bulk limit, as part of this thesis. A picture of the solvent fluxionality is provided, which aims at understanding full solvation from the interaction of trans-AI with a limited number of DMSO molecules. The most stable trans-AI monomer and its 1:1 solvent-solute clusters with a single DMSO molecule were used as starting points for ab initio molecular dynamics simulations and calculations of the VCD spectrum by means of the nuclear velocity perturbation theory approach.

3.1.2. An example of the study of VCD spectra calculated using a cluster-in-a-liquid model: cis-1-amino-2-indanol

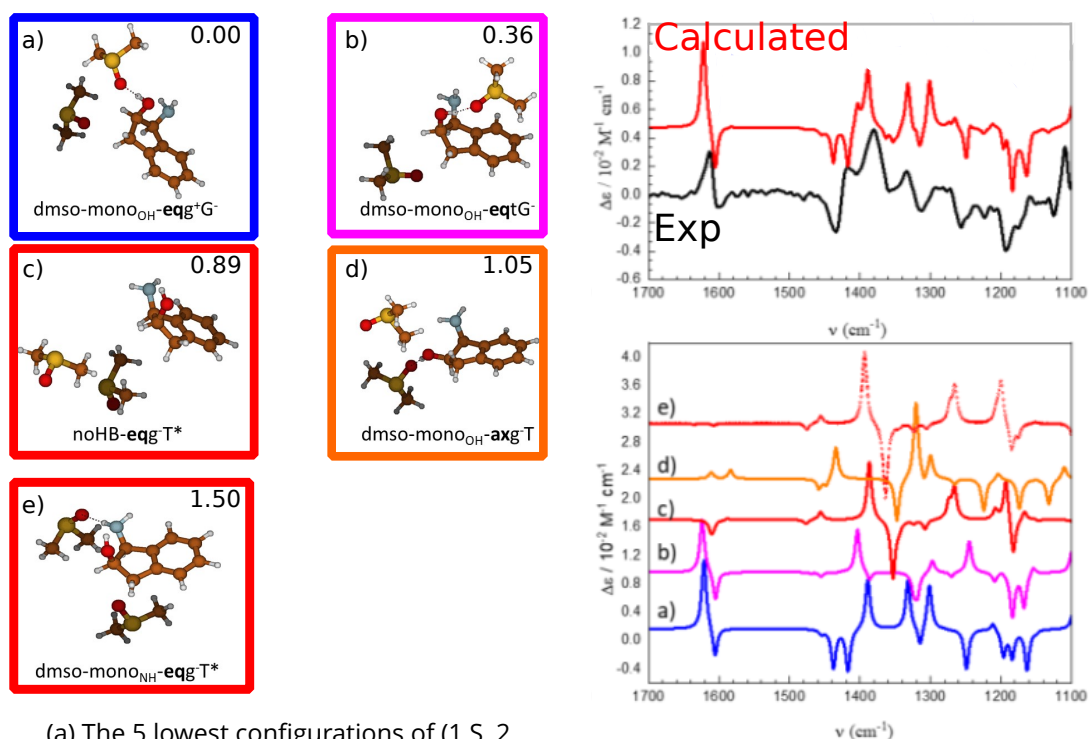
In the following example, the influence of the solvent on the spectra is studied for (1 *S*, 2 *R*)-(-)-cis-1-amino-2-indanol (cis-AI) in DMSO [4]. The solvent is described with the cluster-in-a-liquid model, as described in the previous chapter.

The structures of cis-AI with two molecules of explicit solvent, namely the 1:2 complex, were optimised at the B3LYP-D3BJ/6-311++G(d,p) level of theory and VCD spectra were computed at the same level of theory. Both the structure and the corresponding VCD spectra are represented in figure 3.2. Two molecules of explicit solvent were added to explore their interaction with the two functional groups of cis-AI. All structures were calculated in a DMSO continuum solvent with a polarisable continuum model (PCM).

Special attention was paid to the modes that are possibly involved in hydrogen bonding interactions with the solvent. These modes are represented with β and, in parenthesis, the group in the solute that is hydrogen bonded to the solvent. For example, the mode $\beta(\text{OH})$ is found at $\sim 1400 \text{ cm}^{-1}$ and the mode $\beta(\text{NH}_2)$ is represented at $\sim 1600 \text{ cm}^{-1}$. The individual spectra, which are the same color as the frames around the corresponding structure, show varying spectral shapes at 1400 cm^{-1} and 1600 cm^{-1} depending on the

solvent-solute interaction. The blue individual spectrum bears a strong resemblance to the experimental spectrum shown above. Its structure is also the lowest in energy. Thus it can be confirmed that the most likely structure of cis-AI corresponds to the structure shown surrounded by a blue frame in figure 3.2. This structure has an hydrogen bond between the OH group of cis-AI and the oxygen of a DMSO molecule.

However, the cluster-in-a-liquid model involves a limited number of explicit solvent molecules and the anharmonic and temperature effects are only included through Boltzmann averaging. The next system, trans-1-amino-2-indanol, will be explored with the cluster-in-a-liquid method, as previously, though it will also be analysed with molecular dynamics methods.



(a) The 5 lowest configurations of (1 S ,2 R)-(-)-cis-1-amino-2-indanol (cis-AI) with 2 molecules of explicit DMSO solvent. The values in the top right corners are the Gibbs free energies in kcal·mol⁻¹.

(b) Top: Experimental VCD spectrum compared with the simulation resulting from Boltzmann-averaged contributions of the 1:2 complexes. Bottom: Individual VCD contributions.

Figure 3.2: Structures and corresponding spectra of the 1:2 complex composed of 1 molecule of (1S,2R)-(-)-cis-1-amino-2-indanol (cis-AI) and 2 molecules of explicit DMSO solvent. This figure was adapted from Le Barbu-Debus et al. [4].

3.2. Trans-1-amino-2-indanol: cluster-in-a-liquid approach to VCD spectroscopy

3.2.1. Nomenclature

Three parameters are important for the description of the trans-AI geometry. The first parameter is related to the alicyclic ring puckering motion, represented in figure 3.3a. Due to the stereochemistry of trans-AI, the substituents are both either in axial or equatorial positions, resulting in two kinds of geometry: axial (dihedral angle $C_2C_3C_4C_9 > 0$) denoted hereafter **ax**, or equatorial denoted **eq** (dihedral angle $C_2C_3C_4C_9 < 0$), as shown in figure 3.3a. The two other parameters are the rotation of the OH and the NH₂ groups, represented in figure 3.3b. The orientation of the OH group (see figure 3.3b) will be denoted by g^+ , g^- and t when the HC_2OH dihedral angle is close to 60° , -60° and 180° , respectively. For the orientation of the NH₂ group, we will consider the position of the lone pair (lp) relative to C_1H ; it will be denoted G^+ , G^- and T when the HC_1Nlp dihedral angle is around 60° , -60° and 180° , respectively.

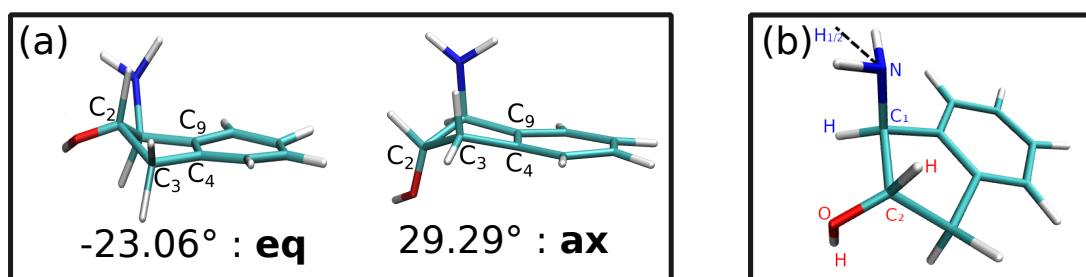


Figure 3.3: (a) Two examples of the dihedral angle $C_2C_3C_4C_9$ of the 5-membered ring of trans-AI showing one axial and one equatorial position; (b) Dihedral angles describing the geometries of the hydroxyl (red) and amino groups (blue) of trans-AI.

3.2.2. Cluster-in-a-liquid approach

In a first step, the system is studied within the cluster-in-a-liquid model, using the same level of theory as the previous system. This level of theory has been shown to reproduce the VCD spectrum of similar molecules in solution, as well as their structure and vibrational spectrum in the gas phase [5, 6, 7]. It has also been used successfully for calculating the VCD spectrum of amino-acids in aqueous solution [8]. Solvent effects were taken into account by the IEFPM implicit polarisable continuum model [9]. Vibrational frequencies were computed at the same level of theory and scaled by 0.98 to correct for anharmonicity and basis set incompleteness. This value is close to that used for similar theoretical methods [10, 11]. The final vibrational spectra were obtained by convoluting the harmonic intensities with a Lorentzian line shape (FWHM 4 cm^{-1}).

The individual spectra of the most relevant 1:1 complexes are shown in figure 3.4, where one configuration from each family is also depicted. The structures resulting from the exploration of the potential energy surface can be classified into three fami-

lies, namely, bidentate, monodentate and non-hydrogen-bonded complexes. Each family encompasses several conformations of trans-AI that will be included in the Boltzmann-averaged contributions when simulating the IR absorption and VCD spectra. In the bidentate family, the DMSO oxygen atom interacts with both OH and NH₂ groups of trans-AI. In the monodentate families, the DMSO oxygen interacts with either the OH (mono_{OH}) or the NH₂ (mono_{NH}) group of trans-AI. The last family is called non-hydrogen bonded (noHB) and accordingly contains structures lacking any intermolecular hydrogen bond. However, this type of structure with a loose DMSO molecule is not likely to be realistic.

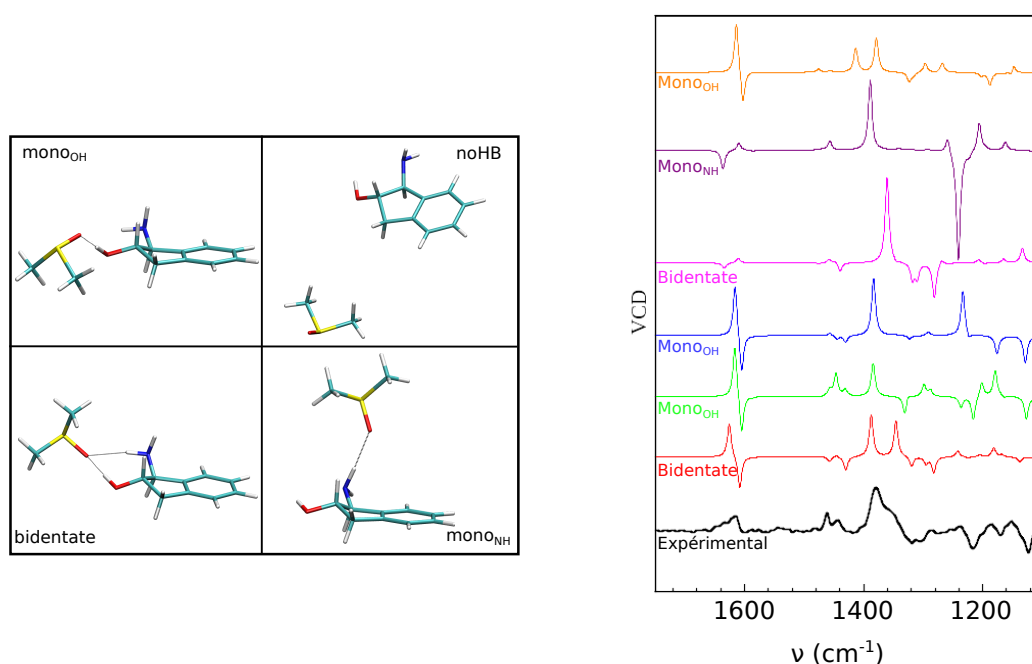


Figure 3.4: (a) Four families of configurations of the (S,S)-trans-AI (b) Experimental VCD spectrum of (S,S)-trans-AI and simulated spectra for different configurations with the lowest-energy configuration at the bottom and the highest-energy configuration at the top. Adapted from Le Barbu-Debus et al. [12].

At first sight, inspection of figure 3.4 indicates that the experimental spectrum can be accounted for by the contribution of a limited number of structures. The best match is obtained for mono_{OH}-**eqg**⁻G⁻, which is the second most stable complex corresponding to the light green spectrum. All features except the positive band at ~1180 cm⁻¹ find their counterpart in the experimental spectrum. The weak positive doublet calculated at 1447/1431 cm⁻¹ corresponds to that observed at 1462/1443 cm⁻¹. An intense band is predicted at 1384 cm⁻¹, which meets the position of the intense experimental feature. The negative band calculated at 1331 cm⁻¹ may correspond to that observed at 1318 cm⁻¹. The last two calculated negative bands at 1216 and 1126 cm⁻¹ are in particularly good agreement with those observed at 1217 and 1123 cm⁻¹. Other complexes show partial overlap with the experimental spectrum and can contribute to it as well. The VCD spectrum of the most stable bi-**eqg**⁺G⁻ complex displays, like the **eqg**⁺G⁻ monomer,

the bisignate transition centred at 1616 cm^{-1} . Two positive features calculated at 1387 and 1345 cm^{-1} are close to the intense and broad experimental band at 1380 cm^{-1} and its shoulder at 1355 cm^{-1} . The negative signal at 1320 cm^{-1} is in good agreement with that observed at 1318 cm^{-1} . However, the negative features observed at 1228 and 1125 cm^{-1} are not predicted for the bidentate complex. They are due to coupled motions of the NH_2 and OH groups, which are strongly modified by the intermolecular hydrogen bond network and lose intensity in the complex. The $\text{mono}_{\text{OH}}\text{-ax}$ complexes also display satisfactory overlap with the experimental spectrum; in particular, $\text{mono}_{\text{OH}}\text{-axtG}^-$ shows the bisignate signature of the G^- forms and it becomes one of the most stable complexes when dispersion corrections are considered. The axial complexes all have negative VCD signals in the region of $\sim 1200\text{ cm}^{-1}$, as experimentally observed. While the experimental features find at least partial counterpart in the spectra of the mono OH or the bidentate families, this is not the case for mono_{NH} , which lacks in particular the strong positive band between 1340 and 1390 cm^{-1} . This poor match, together with very high relative energy, allows discarding the mono_{NH} complex. $\text{bi-egq}^+\text{T}$ cannot contribute to the spectrum due to the wrong signs of the bisignate at 1620 cm^{-1} and of all the bands between 1100 and 1250 cm^{-1} . Lastly, $\text{mono}_{\text{OH}}\text{egqG}^-$ exhibits a wrong sign for the doublet at $\sim 1450\text{ cm}^{-1}$.

The Boltzmann-averaged VCD spectra of the monomer and the 1:1 complex are compared to the experimental spectra in figure 3.5. The calculated VCD spectrum of the 1:1 complex is in good agreement with the experimental spectrum, in particular in the 1350 cm^{-1} range. The calculated positive band at 1624 cm^{-1} corresponds to that at 1614 cm^{-1} in the experimental spectrum. The 1462 and 1443 cm^{-1} bands and the intense positive band at 1380 cm^{-1} with the shoulder at 1355 cm^{-1} are well reproduced. The negative features at 1217 and 1123 cm^{-1} also have their counterpart in the calculated spectrum. However, the band calculated at 1281 cm^{-1} has no equivalent in the experimental spectrum. The influence of the solvent can be seen around 1380 cm^{-1} with a negative then positive signal present in the experimental spectrum and the 1:1 complex spectrum but not in the monomer spectrum. This VCD band contains contributions of $\beta(\text{OH})$ and $\beta(\text{NH})$ bends.

Although the hydrogen bond pattern does not change upon inclusion of a second DMSO molecule, the energetic ordering is strongly modified compared to the 1:1 complexes. This observation points to the difficulty of defining the relative contribution of the different structures by their Boltzmann weights alone and the intrinsic limitation of describing solvation by a finite number of molecules.

In summary for the 1:1 complexes, an important role of the solvent is to stabilise the structures selectively, therefore changing the contributions of the different conformers in the spectra. In this respect, it is worth noting that axial forms are stabilised in DMSO complexes, in particular $\text{mono}_{\text{OH}}\text{-axtG}^-$ can contribute to the final spectrum. Although almost no intra-molecular structural modification happens upon solvation, hydrogen bonding induces some band shifting and changes in the intensity or sign of the VCD spectrum in the corresponding regions. This is especially significant in the region of the $\beta(\text{NH})$ and $\beta(\text{OH})$ modes, just below 1400 cm^{-1} . Not all interaction types perturb the spectrum to the same extent. Non-hydrogen bonded complexes show the same spectrum as the corresponding monomer. $\text{bi-egq}^+\text{G}^-$ shows strong modification of the VCD

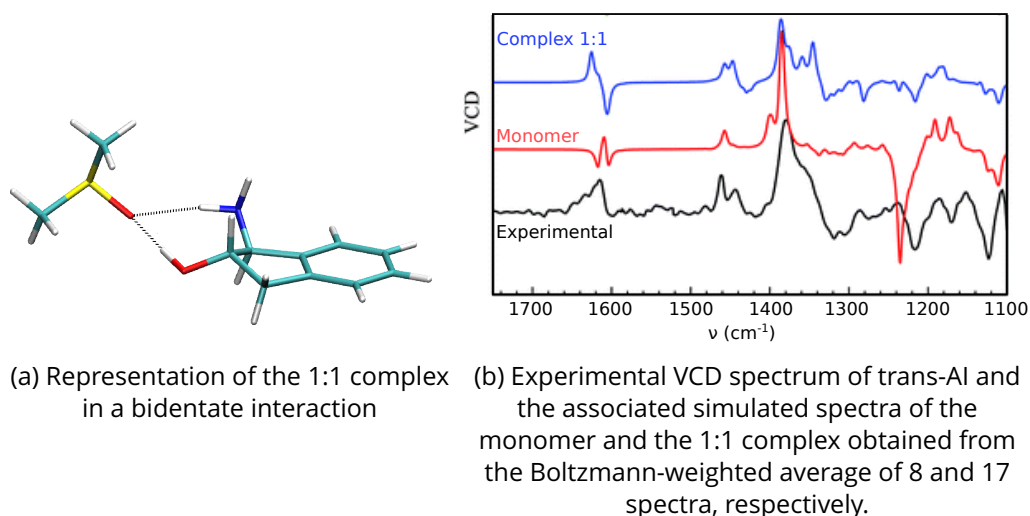


Figure 3.5: Structures of the 1:1 complex, experimental VCD spectrum and calculated spectra of the monomer and the 1:1 complex. Adapted from Le Barbu-Debus et al. [12].

spectrum compared to eqg^+G^- in the 1200 cm^{-1} region, as mentioned above. The interaction with NH_2 perturbs the spectrum much less than the interaction with OH does: the VCD spectrum of $\text{mono}_{\text{NH}}\text{-eqg}^+\text{G}^+$ is almost identical to that of the corresponding monomer, while that of mono_{OH} complexes is more affected by solvation.

3.3. Beyond cluster-in-a-liquid: exploration of the potential energy surface with molecular dynamics

MD simulations were performed with the AMOEBA force field and with AIMD to explore the potential energy surface for trans-AI with a varying number of DMSO molecules.

To characterise the position of the DMSO molecules with respect to trans-AI, the distances between the nitrogen or oxygen of trans-AI and the oxygen of DMSO were evaluated throughout the trajectory and drawn as 2D contour plots using the Plotly library in the R software [13, 14]. The distance between the nitrogen atom of trans-AI and the oxygen atom of the DMSO is represented on the y-axis and the distance between the oxygen atom of trans-AI and the oxygen atom of the DMSO is on the x-axis. Then, the z-axis represents the occurrence of the distance observed over the trajectory in a false colour scale. Each graph is divided in four regions: (i) a spot at ($x \approx 3\text{ \AA}$, $y \approx 3\text{ \AA}$) corresponds to the bidentate family; (ii) a spot around ($x \approx 3\text{ \AA}$, $y \approx 5\text{ \AA}$) represents a monodentate configuration with the DMSO hydrogen-bonded to the OH group of trans-AI; (iii) a spot at ($x \approx 4\text{ \AA}$, $y \approx 3\text{ \AA}$) corresponds to a monodentate configuration with the DMSO hydrogen bonded to the NH_2 group of trans-AI; (iv) finally, for a spot at ($x \geq 5\text{ \AA}$, $y \approx 8\text{ \AA}$), the DMSO is in the vicinity of trans-AI without being hydrogen bonded.

3.3.1. AIMD approach

For isolated trans-AI, the 1:1 as well as the 1:2 complexes with DMSO, one, three, and one starting structures were created, respectively, based on the optimised geometries found from the static calculations. AIMD calculations based on DFT were carried out, by our collaborator S. Jähnigen, with the Quickstep module of the CP2K software package [15, 16].

The simulations were of Born-Oppenheimer MD type with a time step of 0.5 fs, using the BLYP or B3LYP exchange-correlation functional [17, 18], Grimme's dispersion correction (D3) [19], GTH pseudopotentials [20, 21, 22], and the Gaussian and plane wave basis TZVP-MOLOPT-GTH with an energy cutoff of 400 Ry [23]. All simulations were performed in the canonical ensemble using the CSVR thermostat [24].

The simulations were performed at a slightly elevated temperature of 340 K to counterbalance the underestimation of temperature by the chosen functional [25]. However, the effect of this temperature on the intermolecular sampling is marginal, as shown in figure 3.6.

The AIMD simulations were carried out in a vacuum supercell, that is each setup, either the isolated molecule or the complexes (1:1 or 1:2) with DMSO, was placed in a box of predefined size: 16^3 \AA^3 and 20^3 \AA^3 for the isolated molecule and the complexes, respectively. Each sample underwent geometry optimisation followed by a 5 ps equilibration performed under massive thermostating with a coupling constant in the 10-500 fs range. The production trajectory of 30 ps was carried out under global thermostating with a coupling constant of 500 fs.

AIMD trajectories were out of reach for 1:5 complexes, but could be carried out at 340 K for three 1:1 (trajectories b1 to b3) and one 1:2 complexes using the BLYP functional and at 320 K for two 1:1 complexes (trajectories d1 to d2) using the B3LYP functional. The starting structures contained DMSO interacting either in a bidentate manner or via OH...O interaction. The BLYP-AIMD trajectories are short and produce results that depend more strongly on the starting structure, their analysis indicates that trans-AI spends 28% and 30% in the axial position for the 1:1 and 1:2 systems, respectively.

The influence of the functional on the trajectories is illustrated by comparing the results described above with similar MD trajectories obtained with the B3LYP functional (figure 3.7). For the isolated molecule, the B3LYP trajectories also evidence contributions from both axial and equatorial configurations. The main difference compared to BLYP results is an increased contribution of G^- structures. For the 1:1 complex also, a balanced contribution of axial (41%) and equatorial configurations (59%) is obtained, as well as a larger contribution of G^- structures relative to BLYP trajectories.

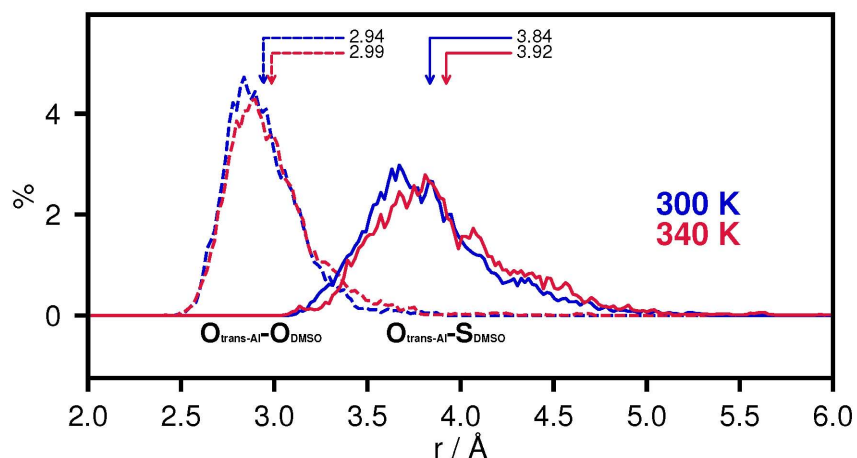


Figure 3.6: Distribution of distance between the oxygen atom of trans-Al and the oxygen/sulfur atom of DMSO, obtained from the AIMD trajectories at different temperatures (BLYP-D3).

In the 1:1 complex, the contour plot on figure 3.7 shows that both the bidentate and monodentate (OH and NH₂) conformations are visited. The DMSO molecule remains close to the OH group of trans-Al.

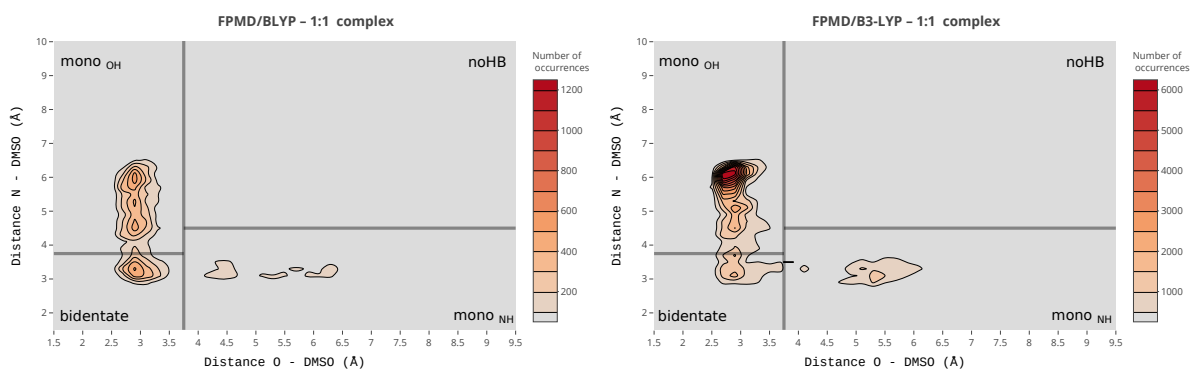


Figure 3.7: Contour plots of occurrence of O(trans-Al)-DMSO and N(trans-Al)-DMSO distances for the 1:1 complex, with comparison between the BLYP (left) and B3LYP (right) functionals in the AIMD simulations.

With two DMSO molecules, conformational exploration is very limited, as manifested by the contour plots of figure 3.8 where the simulation mainly explores one type of interaction with DMSO. However, even under this short time scale the dynamics shows a partial mobility away from the initial conformation of trans-Al.

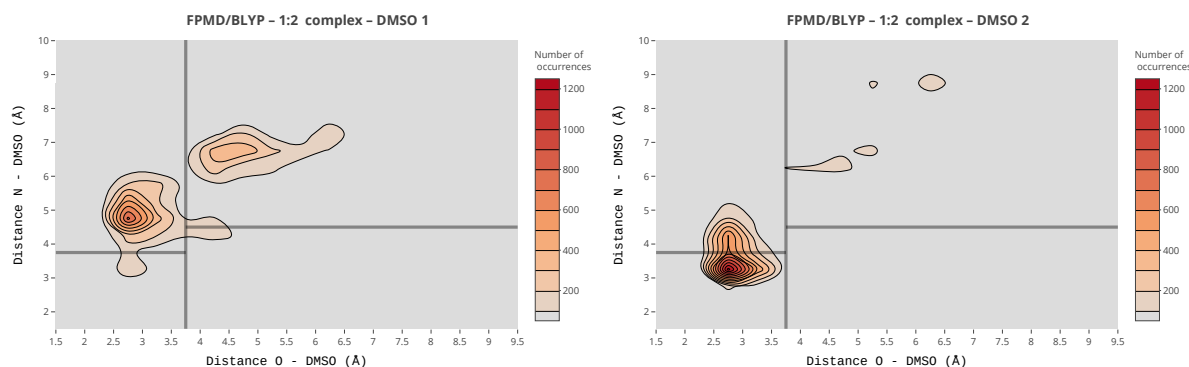


Figure 3.8: Contour plots of occurrence of O(trans-AI)-DMSO and N(trans-AI)-DMSO distances for the 1:2 complex, using the BLYP functional in the AIMD simulations. The left plot represents DMSO 1 and the right plot DMSO 2.

3.3.2. Force field based MD

Polarisable FFMD simulations resting on the AMOEBA force field were carried out to model trans-AI in interaction with one, two and five DMSO molecules on time scales much longer than affordable with AIMD. A set of multipoles was generated for the trans-AI isolated monomer using the distributed multipole analysis of the MP2/ cc-pVTZ electron density [26, 27].

The temperature was set to 300 K or 150 K and was controlled using the Nosé-Hoover thermostat. The additional trajectories at 150 K were performed to constrain the dynamics of the system to remain around a selected local potential well of trans-AI. At 300 K, the duration of the trajectories was 1 ns and a time step of 0.5 fs was used. For the clusters with 2 and 5 DMSO molecules, a spherical van der Waals potential of 17 Å of diameter was used to avoid evaporation. At 150 K, the duration of the simulations was increased up to 3 ns to improve potential energy surface exploration but a time step of 1 fs was used to reduce the computational time. MD simulations of bulk systems were carried out using a cubic DMSO box with an edge length of 19.20 Å that contains 56 DMSO molecules. This box was pre-equilibrated before soaking the trans-AI solute. Simulations were performed at constant volume and 300 K using periodic boundary conditions, the Nosé-Hoover thermostat, and a 0.5 fs time step. Ewald summation was used for the long-range electrostatic interactions. The Ewald real-space cutoff was 7 Å, the van der Waals cutoff was 7 Å. Several simulations were carried out for a total simulation time of 4 ns.

3.3.2.1 Trajectories at 150 K

A local and detailed description of the conformational landscape around a given structure in a potential depth was obtained using 150 K FFMD trajectories.

Under such conditions, the temperature is low enough to catch the trans-AI molecule in individual potential wells, allowing the intrinsic stability of its different conformations to be explored. Four trajectories were carried out for trans-AI in the equatorial position:

one for the 1:1 bidentate conformation that corresponds to the most stable structure in the static sampling, two for the 1:2 complex (one with one bidentate DMSO and the second DMSO in the vicinity, one as mono-OH and mono-NH conformation) and one with five DMSO molecules spread around trans-Al with two molecules hydrogen-bonded to the OH and NH₂ groups. With trans-Al in the axial position, six trajectories were simulated, including two with one DMSO (mono-OH and mono-NH, respectively), three with two DMSO (mono-OH and mono-NH, mono-OH and in the vicinity, mono-NH and in the vicinity, respectively) and one with five DMSO (with similar configuration of the solvent molecules to that of the equatorial system).

The contour plots (figure 3.9 and 3.10) show that the complexes remain near their starting points, keeping the hydrogen-bond interactions with one or two DMSO molecules. One exception is the simulation starting from trans-Al in the equatorial position with one DMSO molecule. In this trajectory, trans-Al switches to the axial position after 1.5 ns and remains axial during the rest of the simulation. This result contrasts with the static DFT results, which give only a few stable structures in axial configuration for the 1:1 complex. It could be explained by an artifact of the force field, which may excessively favour the axial configuration.

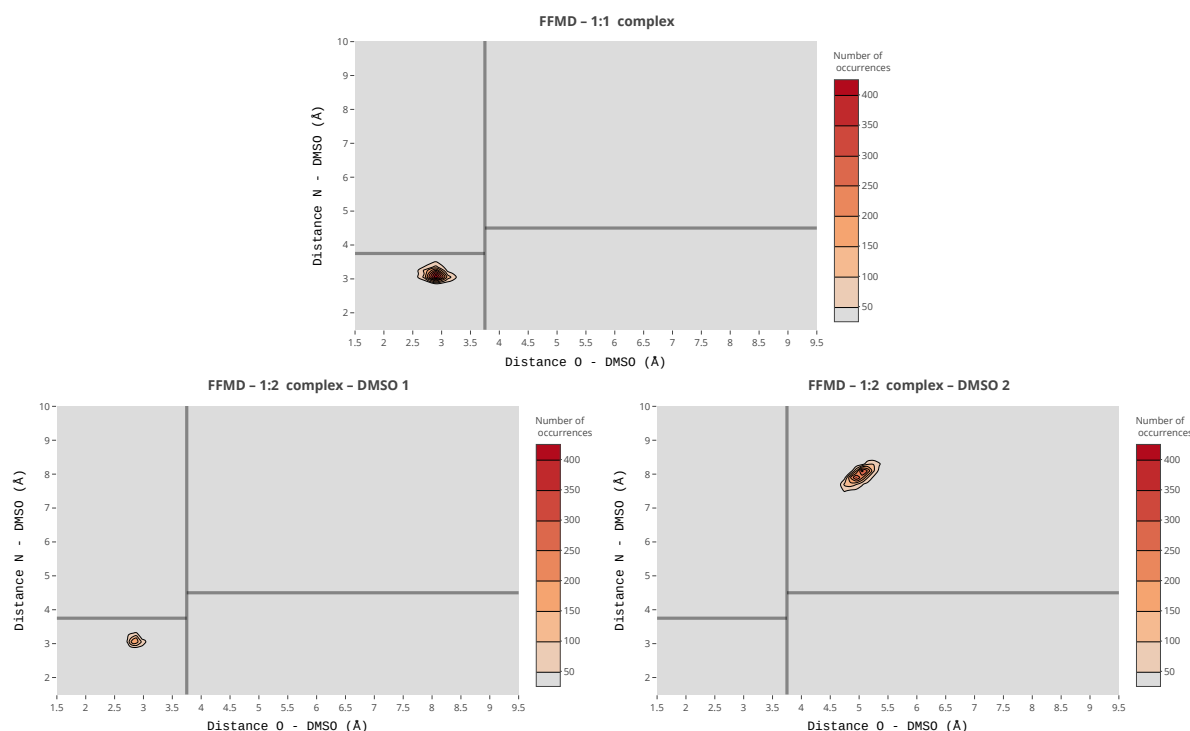


Figure 3.9: Contour plots of occurrence of O(trans-Al)-DMSO and N(trans-Al)-DMSO distances in the trans-Al:(DMSO) 1:1 and the trans-Al:(DMSO)₂ 1:2 equatorial complexes at 150 K in FFMD simulations.

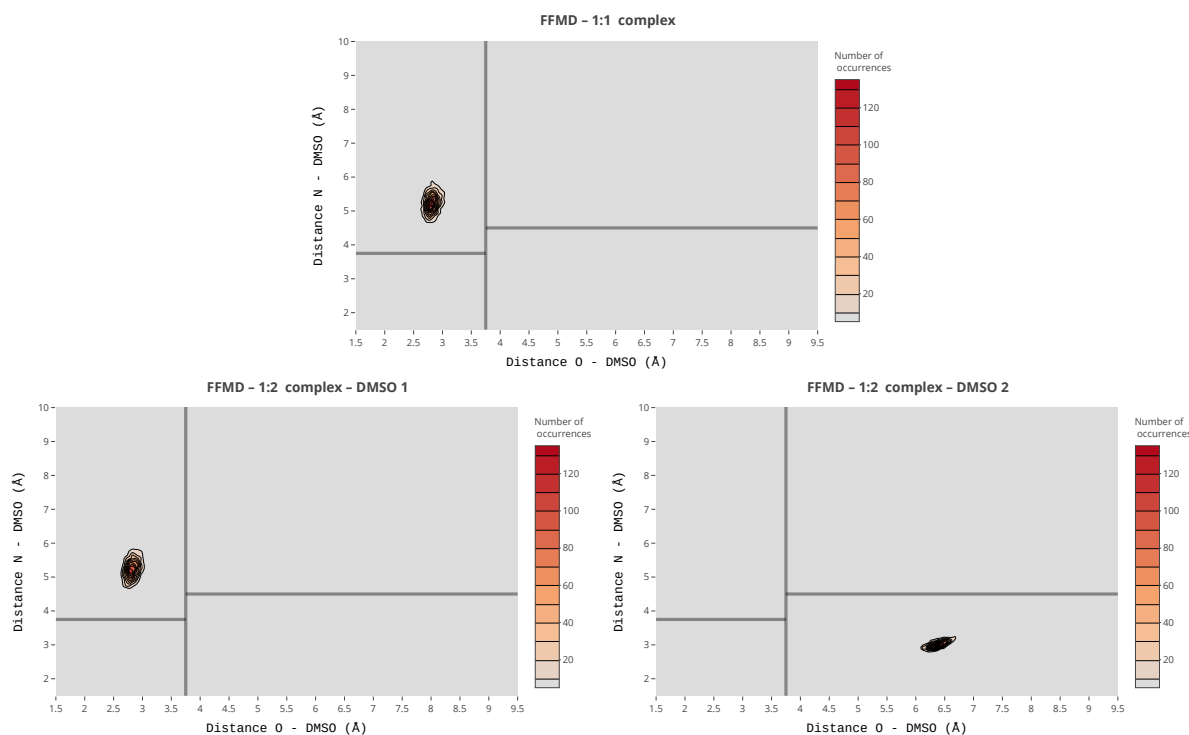


Figure 3.10: Contour plots of occurrence of O(trans-Al)–DMSO and N(trans-Al)–DMSO distances in the trans-Al:(DMSO) 1:1 and the trans-Al:(DMSO)₂ 1:2 axial complexes at 150 K in FFMD simulations.

In figure 3.11, the contour plots of the 1:5 equatorial complex show the five DMSO molecules staying mainly around their initial position. One molecule stays hydrogen bonded to the OH group, two are in interaction with the NH₂ group with one closer to the OH, and the remaining two molecules are in the vicinity of the trans-Al solute. These results are consistent with the main intermolecular interactions expected in this complex, i.e. OH ··· O and NH ··· O hydrogen bonds, and solvent-solvent interactions. The dynamics of the third DMSO molecule is more diverse. Lastly, the bidentate conformation is present but in a smaller amount and does not correspond to a well-defined spot in figure 3.11. Still, the superimposed contour plots for the five DMSO molecules show that all the expected interaction sites of trans-Al (protic groups and aromatic ring) interact with DMSO, which demonstrates that 1:5 complexes give a realistic description of solvation for this system.

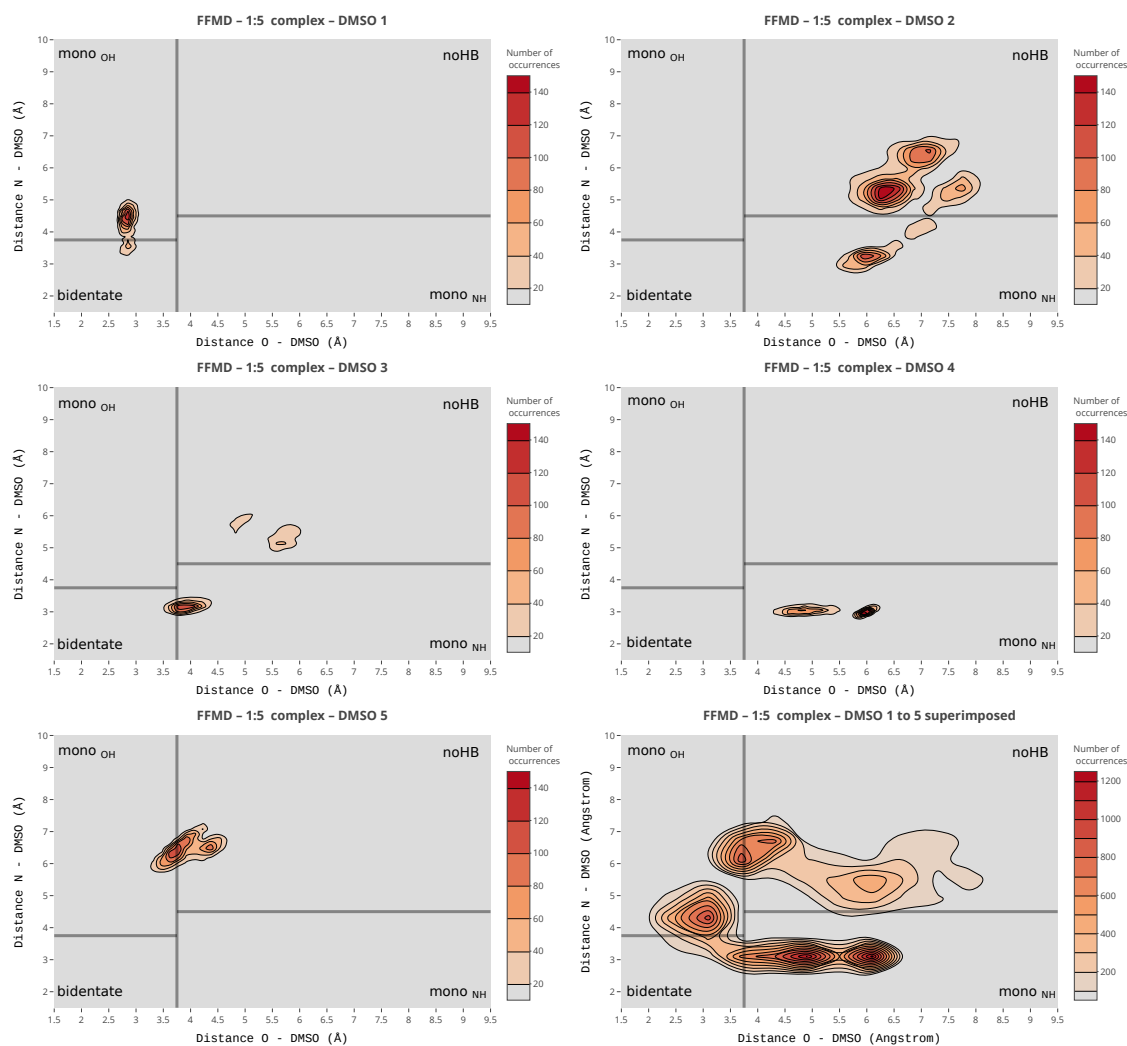


Figure 3.11: Contour plots of occurrence of O(trans-AI)-DMSO and N(trans-AI)-DMSO distances for each of the five DMSO molecules in the 1:5 equatorial complex at 150 K, together with their superposition. The contour plots are divided in four regions labelled with the corresponding static 1:1 complexes.

3.3.2.2 Trajectories at 300 K

FFMD simulations at 300 K were performed to assess the sampling of a few solvent molecules around trans-AI. At 300 K, the FFMD trajectories are less impacted by the starting configuration. Four trajectories were produced, namely two with one DMSO, and two with either two or five DMSO molecules, respectively. The dihedral angle of the 5-membered ring was monitored to characterise the equatorial and axial conformations (see figure 3.12). In the FFMD trajectory for the 1:1 complex, trans-AI spent most the time in axial position (93%). However the results are more balanced for the system with two (80%) and even more so for five DMSO (45%).

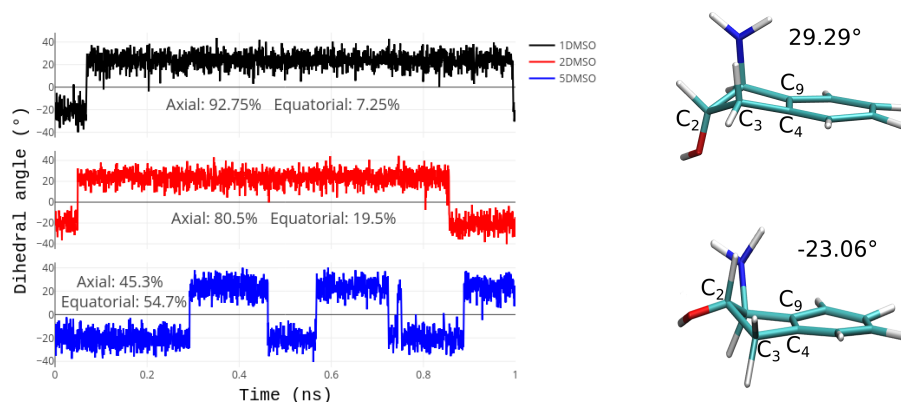


Figure 3.12: Dihedral angle $C_2C_3C_4C_9$ of the 5-membered ring of trans-AI during the FFMD simulations at 300 K with one, two and five DMSO molecules.

The time evolution of the dihedral angles of the hydroxyl and amino groups of trans-AI are shown in figure 3.13. Each plot represents a trajectory with 1, 2 or 5 DMSO molecules. For all three FFMD trajectories, the dihedral angle HC_2OH stayed mainly at around 180° or 60° (t and g^+ conformations), which is consistent with the static picture at the DFT level. The residence time of one value of the dihedral angle increases with the number of DMSO molecules as a confirmation of the stability of these t and g^+ conformations in solution.

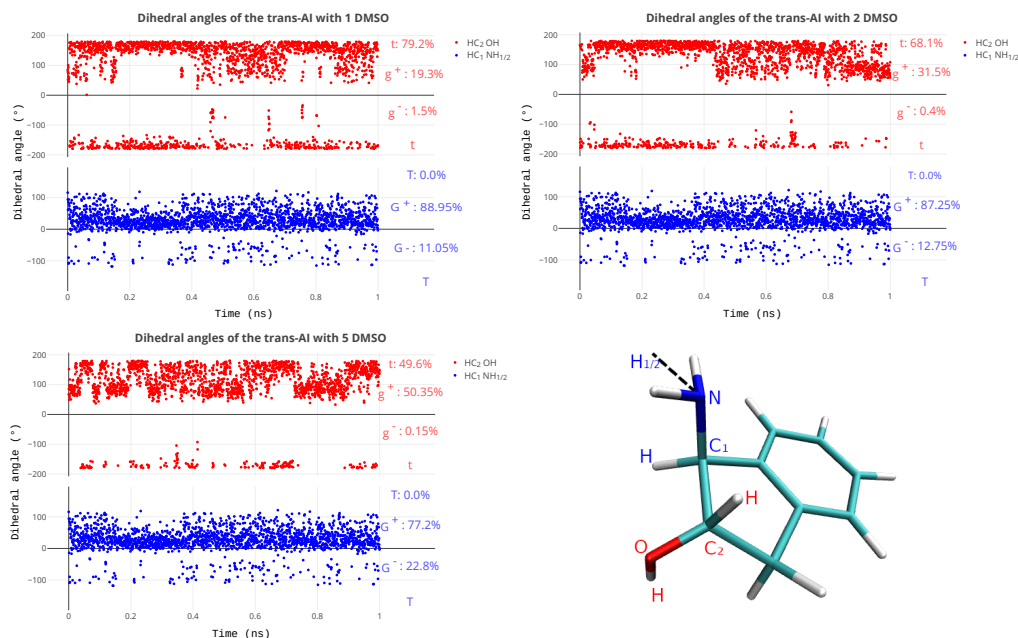


Figure 3.13: Dihedral angles describing the geometries of the hydroxyl (red) and amino (blue) groups of trans-AI during the FFMD simulations at 300 K with one, two and five DMSO molecules.

For the amino group, the structure with the G^+ orientation is clearly favoured with FFMD, however at the DFT level it is the one with the G^- orientation that has the lowest energy with explicit solvation.

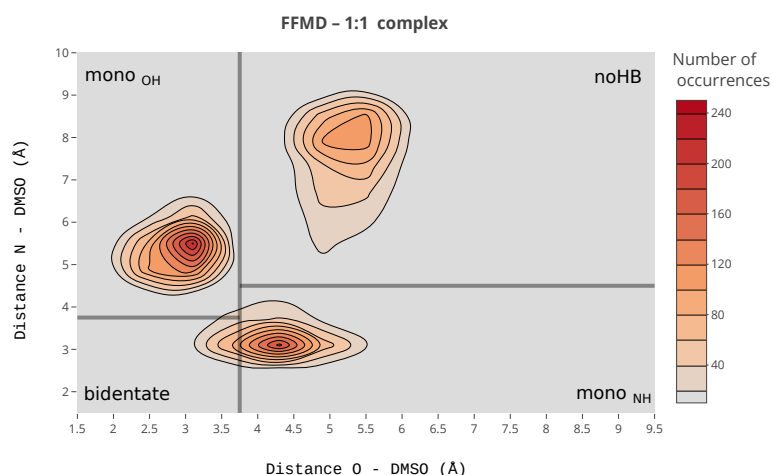


Figure 3.14: Contour plots of occurrence of O(trans-Al)-DMSO and N(trans-Al)-DMSO distances in the 1:1 complex at 300 K in the FFMD simulations.

In the 1:1 complex (figure 3.14), the DMSO molecules mainly takes three positions, near the OH and NH₂ groups of trans-Al (red spot) and non-hydrogen bonded to trans-Al. Contrary to the static DFT calculations, the bidentate interaction is not favoured because trans-Al is mainly in the axial position during the simulation. Two starting points were exploited, one with the solvent hydrogen-bonded to the OH group and the other with the solvent on the NH₂ group. Yet, the two trajectories give similar contour plots.

With two DMSO molecules as explicit solvent, one DMSO molecule stays mainly hydrogen-bonded to the OH or the NH₂ group of trans-Al (figure 3.15, right). The second DMSO molecule takes mainly three positions, hydrogen-bonded to the OH or the NH₂ group of the trans-Al, or non-hydrogen bonded in the OH region (figure 3.15, left).

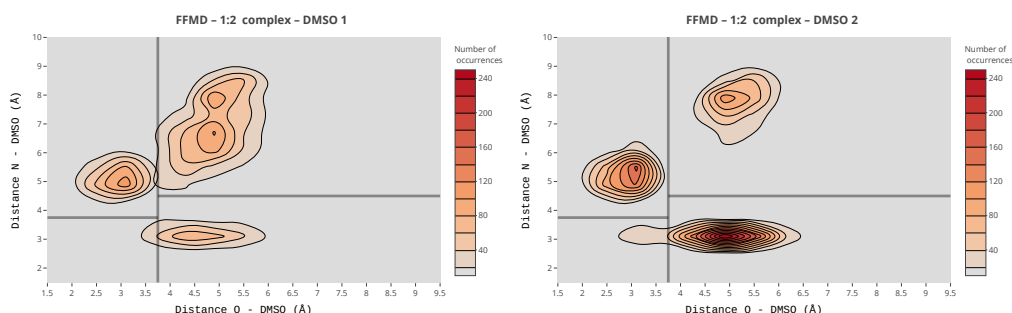
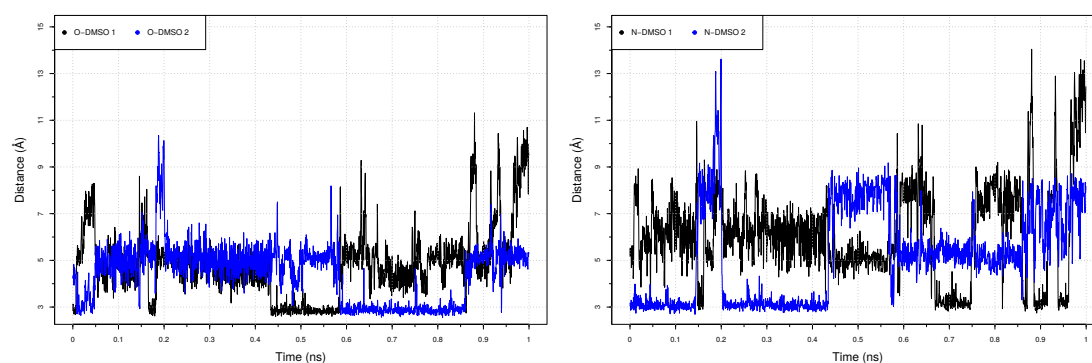


Figure 3.15: Contour plots of occurrence of O(trans-Al)-DMSO and N(trans-Al)-DMSO distances for both DMSO molecules (left and right) in the 1:2 complex at 300 K in the FFMD simulations.

At 300 K, the DMSO molecules are much more mobile than at 150 K, moving around the oxygen and the nitrogen atoms of trans-Al. Figure 3.16 provides the distances between the OH and NH₂ groups of trans-Al and DMSO as a function of time, showing the fluctuations of the two DMSO molecules moving close, then far away from trans-Al. Exchanges between the two molecules happen several times on the NH₂ group with an approximate residence time, i.e. lifetime of the hydrogen bond, between 100 and 200 ps. Nevertheless, the contour plots indicate that there is a tendency of each DMSO molecule to remain in certain spots throughout the simulation.



(a) Distance between the oxygen of trans-Al and the oxygen of DMSO

(b) Distance between the nitrogen of trans-Al and the oxygen of DMSO

Figure 3.16: Distances evolution as a function of time in the trans-Al:(DMSO)₂ 1:2 complex at 300 K, as obtained from FFMD simulations.

FFMD simulations were also performed with five DMSO molecules as explicit solvent. The solvent moves around the trans-Al with an extensive sampling for all the molecules. The contour plots show similar preferred spots as for the 1:1 and 1:2 complexes: near the OH and NH₂ groups of trans-Al and non-hydrogen bonded in the OH region (figure 3.17). The bidentate interaction for one DMSO molecule is also found and can be correlated to a larger proportion of the equatorial conformation.

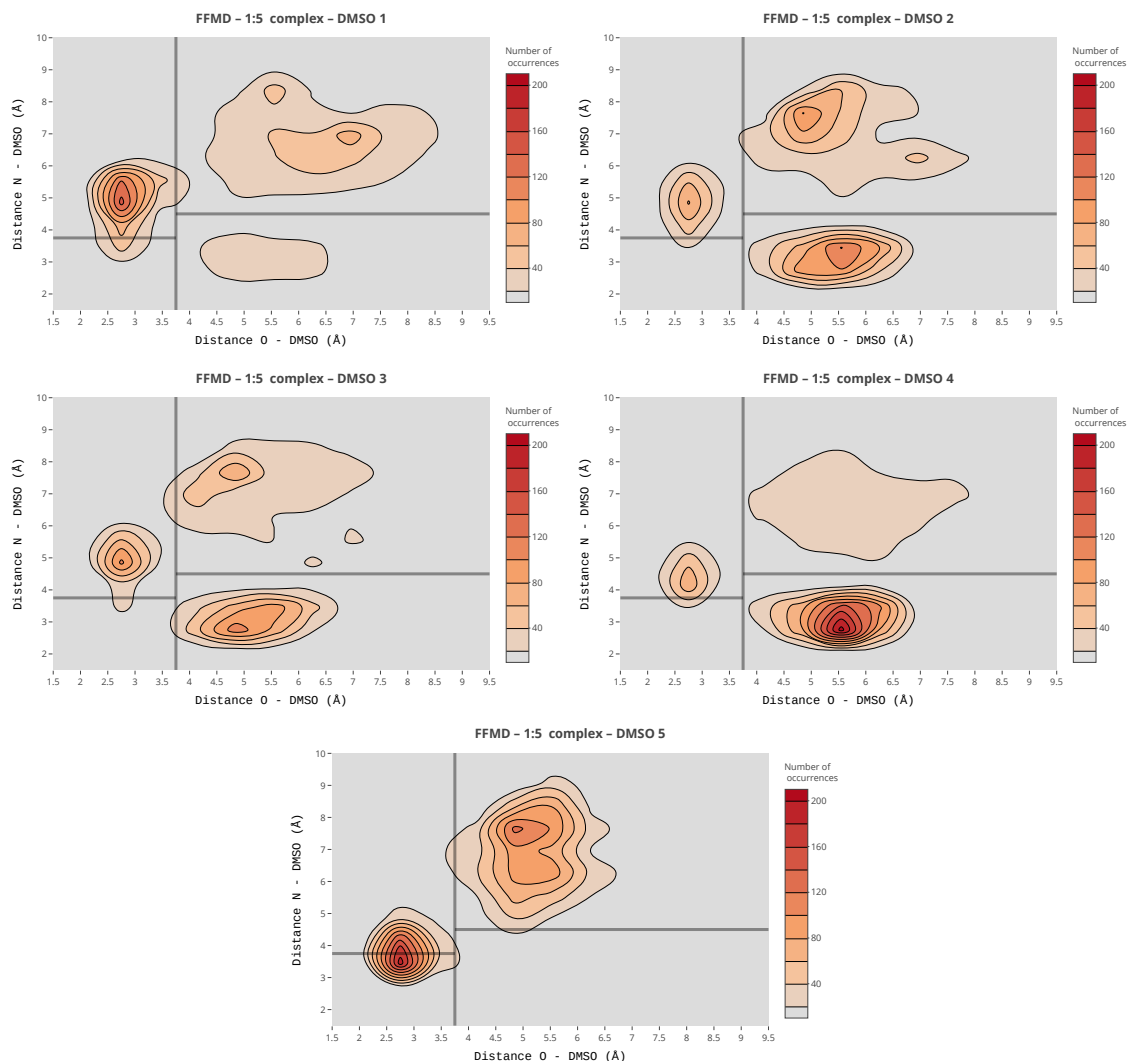
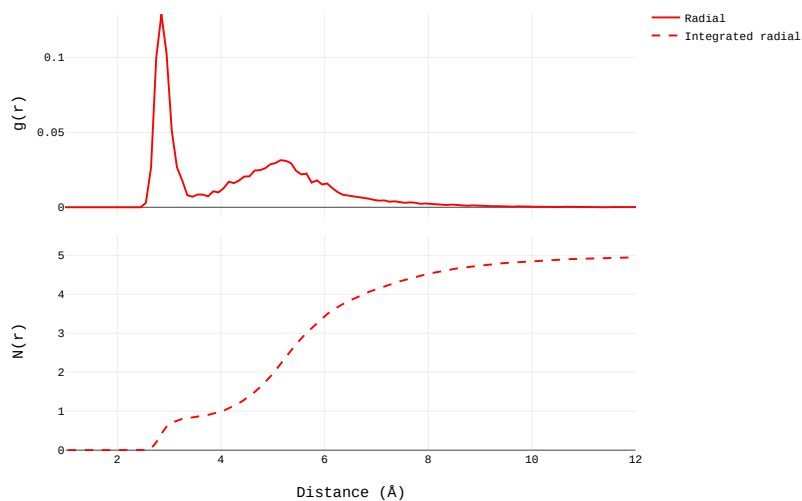
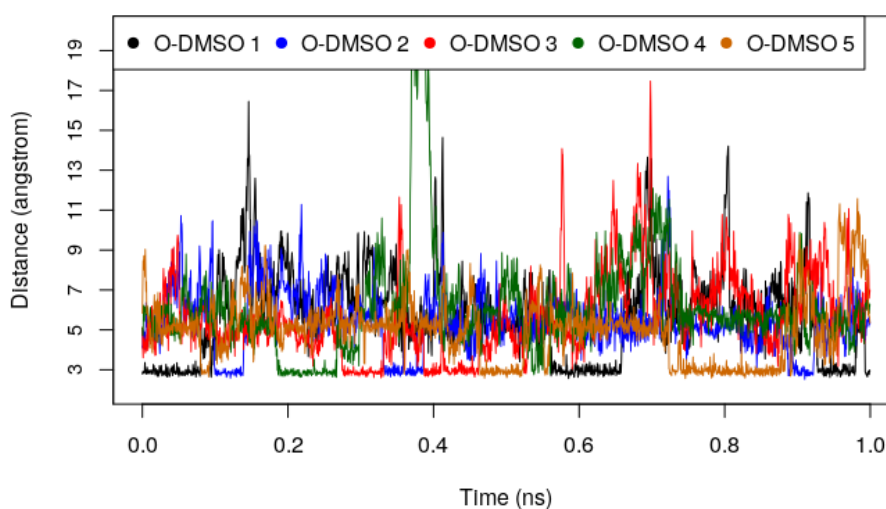


Figure 3.17: Contour plots of occurrence of O(trans-Al)–DMSO and N(trans-Al)–DMSO distances for the five DMSO molecules in the trans-Al:(DMSO)₅ complex during the FFMD simulations at 300 K.

Figure 3.18 provides the evolution with time of the O(trans-Al)–O(DMSO) distance and figure 3.19 the evolution of the N(trans-Al)–O(DSMO) distance, as well as the corresponding radial distribution functions (RDF). Two main conclusions can be drawn from these graphs: (i) at short distances, one DMSO molecule is always hydrogen-bonded to the OH and NH₂ groups, although this interaction is dynamical with frequent exchanges, leading to a residence time of ca. 70 ps and 100 ps, respectively; (ii) the other molecules (three or four, depending on the observed conformation) solvate the trans-Al solute and are also in its vicinity. These observations point to the mobility of the solvent at room temperature.

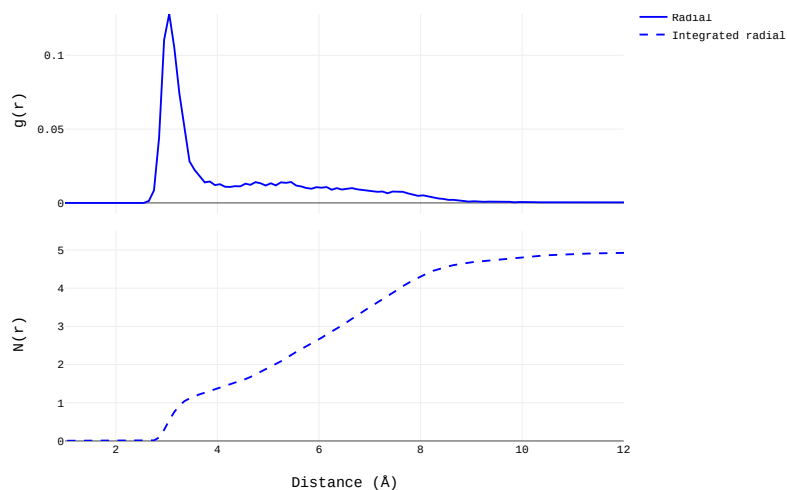


(a) Radial distribution function, and its integral, between the oxygen atom of trans-Al and the oxygens atoms of 5 DMSO molecules.

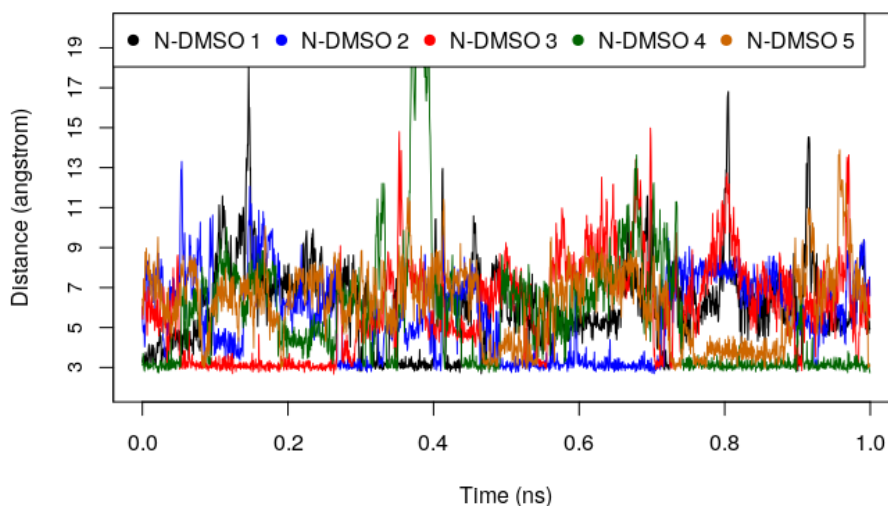


(b) Distance between the oxygen atom of trans-Al and the oxygen atoms of DMSO molecules.

Figure 3.18: Radial distribution functions and distance evolution as a function of time in the trans-Al:(DMSO)₅ 1:5 complex, as obtained from the FFMD simulations at 300 K. Here the O(trans-Al)–O(DMSO) hydrogen bonds are examined.



(a) Radial distribution function, and its integral, between the nitrogen atom of trans-Al and the oxygen atoms of 5 DMSO molecules.



(b) Distance between the nitrogen atom of trans-Al and the oxygen atoms of DMSO molecules.

Figure 3.19: Radial distribution functions and distance evolution as a function of time in the trans-Al:(DMSO)₅ 1:5 complex at 300 K. Here the N(trans-Al)–O(DSMO) hydrogen bonds are examined.

A simulation of 1 ns with a time step of 0.5 fs was computed in a periodic box of 19.20 Å containing trans-Al molecule and 56 DMSO molecules. The trans-Al molecule spends 76.75% of the time in the axial position and the HC₂OH dihedral angle stays mainly around 180° or 60°.

The RDF between trans-Al and the oxygens of 56 DMSO molecules are shown in figure 3.20. The RDF between the oxygen of trans-Al and the DMSO molecules (red) show an intense peak at 2 Å. The RDF between the nitrogen of trans-Al and the DMSO molecules (blue) show an intense peak at 3 Å. The first plateaux of the integration of the RDF are

at 0.9 and 1.7 for the oxygen and the nitrogen of trans-AI, respectively. This confirms the hydroxyl group having nearly always an hydrogen bond with a DMSO molecule at 2 Å and the amino group mostly having two hydrogen bonds with a DMSO molecule at 3 Å. It also suggests the bidentate family being in a small minority. The second solvation shells are around 6 Å from the oxygen of trans-AI and 5.75 Å from the nitrogen of trans-AI.

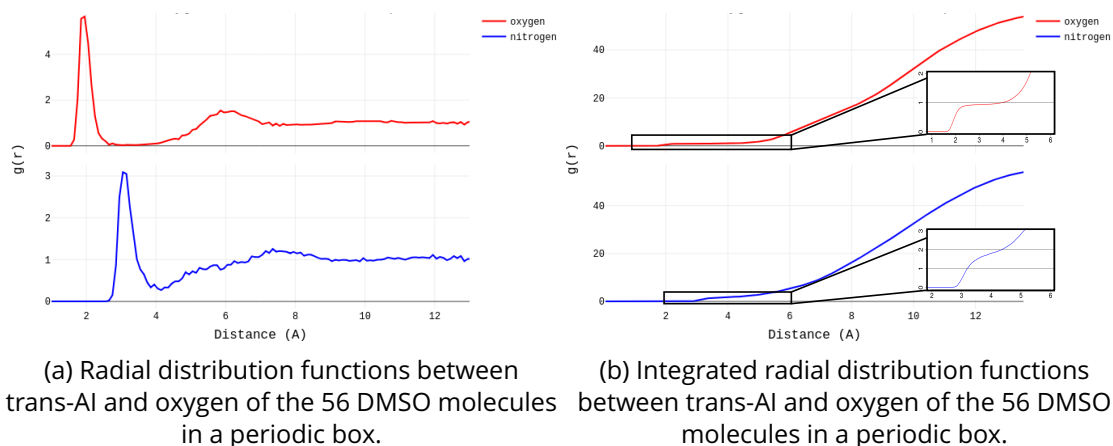


Figure 3.20: Radial distribution functions and their integrations for trans-AI in a box of 56 DMSO molecules at 300 K. The RDF between the oxygen of trans-AI and the DMSO molecules (red) shows an intense peak at 2 Å while the RDF between the nitrogen of trans-AI and the DMSO molecules (blue) shows an intense peak at 3 Å.

In terms of solvation, the results of the AIMD simulations for the 1:1 and 1:2 complexes parallel those of FFMD. They also highlight the mobility of the DMSO molecules (figure 3.7) and the preference of the DMSO for the OH group of trans-AI, which parallels the conclusion of static calculations where the mono OH or bidentate complexes are the most stable structures.

3.4. Locality of vibrational absorption and VCD and the role of the solvent

To further explore the effect of the solvent, a radial cutoff function for the moments was applied as it has been used and introduced in the previous chapter and used in earlier studies [5, 28, 29]. The electric and magnetic dipole moments entering the TCF were thus evaluated according to their originating position in space. The spatially resolved IR and VCD spectra derived from AIMD-NVPT calculations based on the B3LYP trajectories are shown in figure 3.21. This study of the dipole moment was carried out by our collaborator S. Jähnigen.

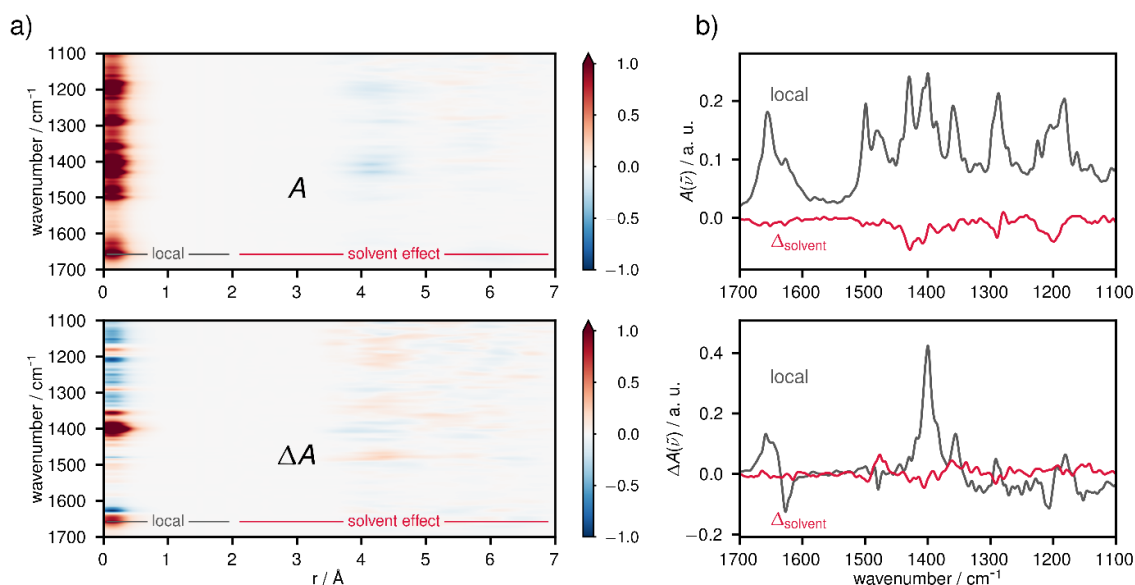


Figure 3.21: AIMD study of non-local effects on the IR and VCD spectrum of trans-AI in DMSO solution by means of 1:1 clusters at the B3LYP-D3 level of theory; (a) radially resolved spectra (top: IR, bottom: VCD) of the chiral solute with spatial regions marked as "local" and "solvent effect", respectively; (b) IR (top) and VCD (bottom) spectrum of trans-AI without consideration of the solvent's polarisation (grey lines) and the difference spectrum as a consequence of the solvent (red lines).

In figure 3.21a, the radius r corresponds to the distance from the centre of mass of trans-AI as it is scanned starting from 0, where only local contributions to the IR/VCD spectrum appear, i.e. those contributions that stem from trans-AI itself. By moving farther from the chiral solute, non-local signatures can be captured; for trans-AI this corresponds to contributions stemming from DMSO. Hence, if a non-local signal can be found, the polarisation of the centre (trans-AI) is coupled with the environment (DMSO), which can be due to either coupled oscillations, or induced polarisation [30]. It becomes clear that the spatial region above 2 Å, denoted as "solvent effect", hardly returns any signal that may account for intermolecular IR/VCD; the main contributions are indicated as being of local origin, that is, stemming from trans-AI itself. Consequently, DMSO as a solvent does not significantly contribute to neither the IR nor the VCD spectrum of trans-AI in solution. There are negligible traces at about 1360 cm⁻¹ and 1400 cm⁻¹ in the IR spectrum and, additionally, at 1425 cm⁻¹ in the VCD spectrum, where the large dipole moment of the solvent slightly polarises the solute. This does not come as a surprise since these frequencies correspond to modes localised at the OH and NH₂ groups, respectively, which bind the molecule to the solvent. Although in these regions the non-local contributions sum up to 20 %, they do not change qualitatively or quantitatively the overall shape of the spectra (figure 3.21b). Consequently, what matters is the conformational sampling of trans-AI and this is where solvent effects are discernible.

Such a conclusion is important as it means that in the case of DMSO electrostatics do

barely influence the VCD response of trans-AI directly. This is at least true for the case of a 1:1 cluster of trans-AI and DMSO, on which the AIMD results rely. Yet, as the solvation study suggests, it can be assumed that this small-scale image generalises to full solvation. The number of solvent molecules surrounding the solute may ultimately lead to a rise of the small contributions shown in figure 3.21, but since the nature of the interaction does not change significantly with the number of DMSO molecules, it is unlikely to alter the locality of the VCD spectrum. Consequently, only the electronic wavefunction of the chiral solute is needed, which encourages the use of theoretical models like QM/MM approaches for IR and VCD determination.

3.5. Concluding remarks

The chosen system, (1S,2S)-trans-1-amino-2-indanol in DMSO, is especially challenging for VCD spectroscopy. Because of its flexibility, many structures coexist within tiny energy differences that strongly depend on the method used. Static calculations based on the cluster-in-a-liquid model capture the main features of the experimental VCD spectra for the 1:1 complex. Conformations higher in energy that are explored in FFMD or AIMD, like axial conformations, may contribute to the spectrum more than their static energy suggests.

The solvent influences only indirectly the calculated spectra through structural effects because of the locality of the VCD signal. This locality is an important finding that further supports the interpretation of vibrational spectra from the sole perspective of the solute conformational space and justifies the use of QM/MM calculations, with only a cluster of limited size being treated quantum mechanically. As with any MD method, ergodicity is an important issue. The limited duration of the AIMD trajectories, due to their computational cost, makes it important to define properly the starting structure. In this respect, combining FFMD for an exhaustive exploration of the PES with AIMD for the calculation of the VCD spectra proved to be efficient. The results obtained here stress the importance of solvent mobility onto the results, which is hardly taken into account by small finite-size clusters. Therefore, a thorough FFMD study seems essential to fully understand VCD for solvated compounds, as a complementary tool to the static cluster picture. This is why development of polarisable force fields amenable to very large sizes, close to full solvation, is necessary. No spectrum was obtained here with such a method as it was still in development when this study was in progress. It is the object of the next chapters where the determination of such spectra is assessed and demonstrated.

Bibliography

- [1] G. Chang, W. C. Guida, and W. C. Still, "An internal-coordinate monte carlo method for searching conformational space," *Journal of the American Chemical Society*, vol. 111, no. 12, pp. 4379–4386, 1989.
- [2] G. Tarczay, G. Magyarfalvi, and E. Vass, "Towards the determination of the absolute configuration of complex molecular systems: Matrix isolation vibrational circular dichroism study of (r)-2-amino-1-propanol," *Angewandte Chemie International Edition*, vol. 45, pp. 1775–1777, mar 2006.
- [3] A. J. Barnes, "Matrix isolation spectroscopy: Technique and applications," in *Low-Temperature Chemistry of the Atmosphere*, pp. 351–372, Springer, 1994.
- [4] K. Le Barbu-Debus and A. Zehnacker, "Competition between inter and intramolecular hydrogen bond evidenced by vibrational circular dichroism spectroscopy: The case of (1s, 2r)-(-)-cis-1-amino-2-indanol," *Chirality*, vol. 33, no. 12, pp. 858–874, 2021.
- [5] K. L. Barbu-Debus, A. Scherrer, A. Bouchet, D. Sebastiani, R. Vuilleumier, and A. Zehnacker, "Effect of puckering motion and hydrogen bond formation on the vibrational circular dichroism spectrum of a flexible molecule: the case of (S)-1-indanol," *Physical Chemistry Chemical Physics*, vol. 20, no. 21, pp. 14635–14646, 2018.
- [6] K. L. Barbu-Debus, N. Guchhait, and A. Zehnacker-Rentien, "Electronic and infrared spectroscopy of jet-cooled (\pm)-cis-1-amino-indan-2-ol hydrates," *Physical Chemistry Chemical Physics*, vol. 9, no. 32, p. 4465, 2007.
- [7] K. L. Barbu-Debus, F. Lahmani, A. Zehnacker-Rentien, and N. Guchhait, "Electronic and infrared spectroscopy of chiral (\pm)-cis-1-amino-indan-2-ol in a supersonic jet," *Chemical Physics Letters*, vol. 422, pp. 218–225, apr 2006.
- [8] E. Orestes, C. Bistafa, R. Rivelino, and S. Canuto, "Including thermal disorder of hydrogen bonding to describe the vibrational circular dichroism spectrum of zwitterionic L-alanine in water," *The Journal of Physical Chemistry A*, vol. 119, pp. 5099–5106, dec 2014.
- [9] J. Tomasi, B. Mennucci, and R. Cammi, "Quantum mechanical continuum solvation models," *Chemical Reviews*, vol. 105, pp. 2999–3094, jul 2005.
- [10] C. Fábri, T. Szidarovszky, G. Magyarfalvi, and G. Tarczay, "Gas-phase and ar-matrix SQM scaling factors for various DFT functionals with basis sets including polarization and diffuse functions," *The Journal of Physical Chemistry A*, vol. 115, pp. 4640–4649, apr 2011.
- [11] National Institute of Standards and Technology, *Precomputed vibrational scaling factors*, 2022. <http://cccbdb.nist.gov/vibscalejust.asp> (accessed June 15th, 2022).
- [12] K. Le Barbu-Debus, J. Bowles, S. Jähnigen, C. Clavaguéra, F. Calvo, R. Vuilleumier, and A. Zehnacker, "Assessing cluster models of solvation for the description of vibra-

- tional circular dichroism spectra: synergy between static and dynamic approaches," *Physical Chemistry Chemical Physics*, vol. 22, no. 45, pp. 26047–26068, 2020.
- [13] Collaborative data science, *Plotly*, 2015. <https://plot.ly> (accessed June 15th, 2022).
- [14] R Core Team, *R: A language and environment for statistical computing*, 2018. <https://www.R-project.org/> (accessed June 15th, 2022).
- [15] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, "Quickstep: Fast and accurate density functional calculations using a mixed gaussian and plane waves approach," *Computer Physics Communications*, vol. 167, pp. 103–128, apr 2005.
- [16] The CP2K developers group, 2020. www.cp2k.org (accessed June 15th, 2022).
- [17] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Physical Review A*, vol. 38, pp. 3098–3100, sep 1988.
- [18] C. Lee, W. Yang, and R. G. Parr, "Development of the colle-salvetti correlation-energy formula into a functional of the electron density," *Physical Review B*, vol. 37, pp. 785–789, jan 1988.
- [19] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "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*, vol. 132, p. 154104, apr 2010.
- [20] S. Goedecker, M. Teter, and J. Hutter, "Separable dual-space gaussian pseudopotentials," *Physical Review B*, vol. 54, pp. 1703–1710, jul 1996.
- [21] C. Hartwigsen, S. Goedecker, and J. Hutter, "Relativistic separable dual-space gaussian pseudopotentials from h to rn," *Physical Review B*, vol. 58, pp. 3641–3662, aug 1998.
- [22] M. Krack, "Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals," *Theoretical Chemistry Accounts*, vol. 114, pp. 145–152, may 2005.
- [23] J. VandeVondele and J. Hutter, "Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases," *The Journal of Chemical Physics*, vol. 127, p. 114105, sep 2007.
- [24] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *The Journal of Chemical Physics*, vol. 126, p. 014101, jan 2007.
- [25] E. Schwegler, J. C. Grossman, F. Gygi, and G. Galli, "Towards an assessment of the accuracy of density functional theory for first principles simulations of water. II," *The Journal of Chemical Physics*, vol. 121, pp. 5400–5409, sep 2004.
- [26] A. J. Stone and M. Alderton, "Distributed multipole analysis: methods and applications," *Molecular Physics*, vol. 56, no. 5, pp. 1047–1064, 1985.

- [27] A. J. Stone, "Distributed multipole analysis, or how to describe a molecular charge distribution," *Chemical Physics Letters*, vol. 83, no. 2, pp. 233–239, 1981.
- [28] A. Scherrer, R. Vuilleumier, and D. Sebastiani, "Vibrational circular dichroism from *ab initio* molecular dynamics and nuclear velocity perturbation theory in the liquid phase," *The Journal of Chemical Physics*, vol. 145, p. 084101, aug 2016.
- [29] S. Jähnigen, A. Scherrer, R. Vuilleumier, and D. Sebastiani, "Chiral crystal packing induces enhancement of vibrational circular dichroism," *Angewandte Chemie International Edition*, vol. 57, pp. 13344–13348, oct 2018.
- [30] M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith, and D. Marx, "Dissecting the THz spectrum of liquid water from first principles via correlations in time and space," *Proceedings of the National Academy of Sciences*, vol. 107, pp. 12068–12073, jun 2010.

Chapter 4

Validation of the polarisable force field approach for VCD spectroscopy: alanine in the gas phase

Classical molecular dynamics performed with a force field (FFMD) usually achieves a much better sampling than *ab initio* methods (AIMD), because trajectories can be propagated over significantly longer times and a larger number of initial conditions. This ability stems from the full classical description in a force field. This differs from AIMD, in which the electronic structure problem is solved quantum mechanically through DFT.

The present chapter aims to discuss and validate our implementation of the strategy described in Chapter 2 for the computation of VCD spectra with FFMD. The AMOEBA force field is designed to simulate organic molecules and biomolecules. Thus it can describe amino acids which are used as building blocks to characterise proteins. The simplest amino acid which is chiral is alanine. Thus isolated alanine was investigated as our benchmark system in the zwitterionic and neutral forms (see fig. 4.1 a and b). The atom numbering will be used in section 4.4. In the gas phase, alanine has been reported to be more stable in the neutral form though once it is hydrated or in a crystal, the zwitterionic form becomes more stable [1, 2]. Initial tests were performed on the zwitterionic form as the corresponding force field parameters were already available. The calculations in the neutral form were performed later, once the parameters for this form had been determined.

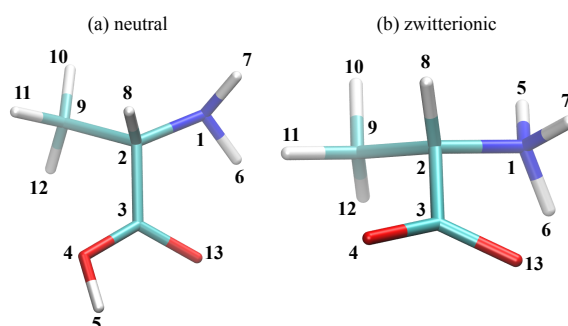


Figure 4.1: Visual depiction of (stable) neutral (a) and zwitterionic (a) forms of alanine in the gas phase.

Various results for the vibrational spectroscopy of isolated alanine have been reported by other groups, from the theoretical [1] and experimental [3] perspectives.

In this chapter, the IR and VCD spectra of isolated alanine were determined with FFMD and the various contributions to the magnetic dipole moment were examined, as well as the time derivative of the electric dipole moment. The quality of sampling was quantified through the determination of static errors and convergence rates of the spectra. Various configurations were determined from an FFMD simulation and reoptimised in DFT to assess the correctness of the FFMD results obtained. The DFT results were also used to evaluate the importance of anharmonicity. Moreover, to validate the dipole moments acquired from the FFMD, they were compared to dipole moments determined with DFT at each step of the same trajectory. Finally effective modes for neutral and zwitterionic alanines were extracted from respective FFMD trajectories. The various contributions from each functional group were also examined.

4.1. Vibrational spectra obtained from force field molecular dynamics (FFMD)

The idea of this section is to obtain IR and VCD spectra from FFMD trajectories. To ensure that the configurations used to determine the electric and magnetic dipole moments are as diverse as possible, different configurations were used to initiate a large number of trajectories. As shown in chapter 2, the resolution of the spectrum is improved by sampling a lot of smaller trajectories rather than one long trajectory. To collect various configurations to be employed as starting points, a long trajectory of 1 ns was achieved in the NVT ensemble. From this trajectory, 40 starting configurations were taken every 10 ps for the last 400 ps. For each starting point, a 5 ps NVT simulation was performed to equilibrate the system at the correct temperature then a 200 ps NVE simulation was carried out, from which the electric and magnetic dipole moments were collected. For each NVE trajectory, an individual spectrum was calculated through the Fourier transform of the time correlation function of the dipole moments, as described in chapter 2. Then the individual spectra were averaged to obtain a spectrum of an hopefully well sampled system.

For neutral alanine, the multipoles were not available in the AMOEBA force field. Thus the atomic multipoles were extracted from ab initio calculations thanks to the distributed multipoles analysis developed by Stone [4]. These multipoles were added to the bonded, van der Waals and polarisation parameters already included in AMOEBA. However the lowest configuration shown in literature [5] was not represented in an initial NVT ensemble FFMD trajectory. The force field parameters were not describing the NH₂ group with sufficient accuracy. Thus the bonded parameters of the NH₂ group were changed to those of similar functional groups already present in the force field. These trials yielded only a slight improvement. Furthermore, the correlation of the relative energy between DFT and FF results was improved by varying the charge on the oxygen of the OH group. The other charges of the system were reequilibrated to maintain the molecular charge at zero.

4.1.1. Calculating the electric and magnetic dipole moments for the IR and VCD spectra

The equations of the electric $\vec{\mu}$ and magnetic \vec{m} dipole moments are recalled below. The first term of electric dipole moment $\vec{\mu}$ is expressed here as a function of the position, thus this electric dipole moment corresponds to the position form. If the electric dipole moment is differentiated with respect to time, $\dot{\vec{\mu}} = \partial\vec{\mu}/\partial t$, the first term of the result becomes a function of the velocity, leading to the so-called velocity form of the electric dipole moment. The magnetic dipole moment is composed of 3 contributions. By ignoring any term with the induced dipole, $\vec{\mu}_i^{\text{ind}}$, the polarisation of the atoms is neglected. The term corresponding to the time varying induced dipoles moving around fixed positions can also be removed to explore the influence of this term.

$$\vec{\mu}(t) = \sum_i^N \left(q_i \vec{r}_i(t) + \vec{\mu}_i^{\text{ind}}(t) \right) \quad (4.1)$$

$$\vec{m}(t) = \sum_i^N \left[\frac{1}{2c} q_i \vec{r}_i(t) \times \vec{v}_i(t) + \frac{1}{c} \vec{\mu}_i^{\text{ind}}(t) \times \vec{v}_i(t) + \frac{1}{2c} \vec{r}_i(t) \times \dot{\vec{\mu}}_i^{\text{ind}}(t) \right] \quad (4.2)$$

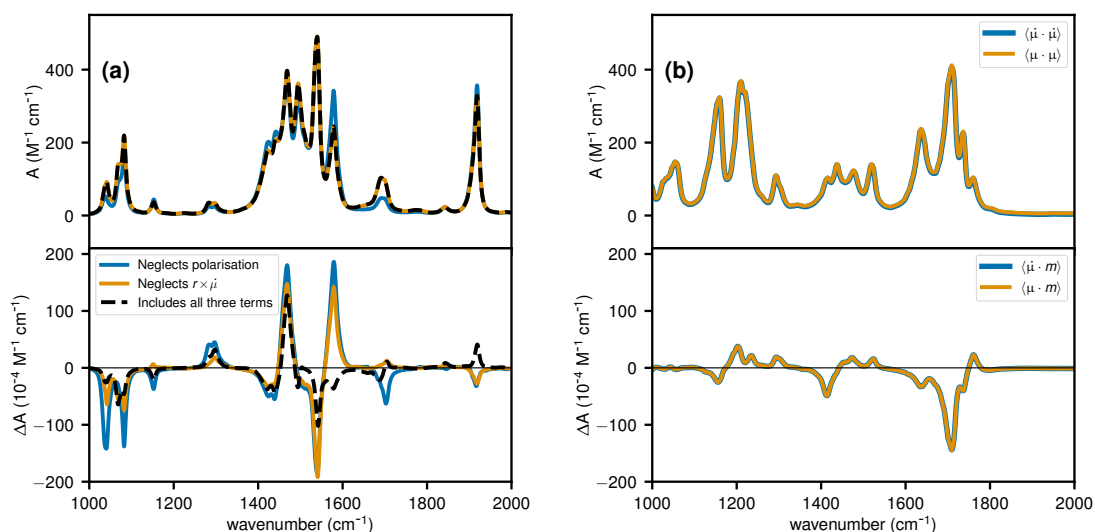


Figure 4.2: IR and VCD spectra of (a) zwitterionic and (b) neutral alanine averaged over 40 FFMD trajectories at T=300 K. Figure (a) emphasizes the influence of the polarisation of the molecule and of the time varying induced dipole moments moving around fixed positions term added to the magnetic dipole moment. Figure (b) shows the spectra using the position and the velocity forms of the electric dipole moment.

In figure 4.2, the IR and VCD spectra of alanine in the zwitterionic and neutral forms are shown. The system was equilibrated at 300 K and the spectra are obtained from the average of 40 trajectories. The influence of polarisation of the molecule through the

induced dipole $\vec{\mu}_i^{\text{ind}}$, of the term $\vec{r}_i \times \vec{\mu}_i^{\text{ind}}$ and of choosing the position ($\vec{\mu}$) or velocity ($\dot{\mu}$) form of the electric dipole moments are also scrutinised. The value of A corresponds to the intensity of the peaks, thus has a unit of $\text{M}^{-1}\text{cm}^{-1}$.

First, the IR spectra of figure 4.2 are examined. For the IR spectrum of figure 4.2a, when the induced dipole moment is neglected, which is only the case for the blue spectrum, a slight change of intensities in the spectrum occurs. No change is found in the IR spectrum of figure 4.2b which arises from the use of the autocorrelation function of the electric dipole moment shown in equation (4.1) and the use of the autocorrelation of the time derivative of the electric dipole moment.

Second, the VCD spectra including a varying number of contributions from the magnetic dipole moment are shown in figure 4.2a. The blue spectrum depicts the simplest expression of the magnetic dipole moment combined with the electric dipole moment which is only composed of its first term. If the induced dipole is added to the electric dipole moment and the second term is added to the magnetic dipole moment, the orange spectrum is obtained. This curve differs from the blue one by peak inversions at 1150 cm^{-1} and 1700 cm^{-1} . The black spectrum arises from the cross-correlation from the full expressions of the electric and magnetic dipole moments. Compared to the orange spectrum, there are peak inversions at 1150 cm^{-1} , 1500 cm^{-1} , 1590 cm^{-1} , 1700 cm^{-1} , 1850 cm^{-1} and 1925 cm^{-1} .

Moreover, the position and velocity forms of the electric dipole moment are compared for the VCD spectra in figure 4.2b. They confirm that there is no influence on the spectrum between the two forms. The velocity form is chosen as the time correlation function (TCF) involving $\dot{\mu}$ decays faster than that involving μ , thus facilitating convergence [6]. In the rest of the thesis, the velocity form of the electric dipole moment and the term $\dot{\mu} \times r_i$ will be used.

These results show the importance of the inclusion of the polarisation, which is performed here through the induced dipole moments within the AMOEBA force field.

By comparing our spectra to the results obtained in the literature for the zwitterionic and neutral forms [1, 7] the peaks above 1600 cm^{-1} could correspond to the C=O stretch of the carboxylate or carboxylic acid and the bend of the amino or amine groups. The bending of the CH_3 is expected around 1450 cm^{-1} and the bending of the CH is predicted at $1300\text{--}1200\text{ cm}^{-1}$. Below 1200 cm^{-1} the presence of the C-N-H bend can be reasonably assumed.

4.1.2. Statistical errors and convergence rates

The spectra of alanine shown above are the result of the average of 40 trajectories. To show the convergence of both forms with the sample size, the averages of an increasing number of trajectories (N_T) are plotted. Convergence is achieved when the intensities of the peaks stabilise at a certain value after a certain sampling time. Starting from the top of figure 4.3, the spectrum arises from the average of 4 trajectories ($N_T = 4$), then the number of trajectories used increases by 4. Each spectrum is compared to the average of 40 trajectories shown in grey.

The spectra for the zwitterionic form converge faster as the charges create stronger electrostatic interactions. Also the IR spectrum converges faster than the VCD spectrum as VCD is more sensitive to a change of configuration of the molecule. Hornivcek et al. [8], found that 100 MD simulations of 20 ps are needed at 300 K to obtain convergence for α -pinene. Similar results are found here with the convergence starting at 10 trajectories of 200 ps.

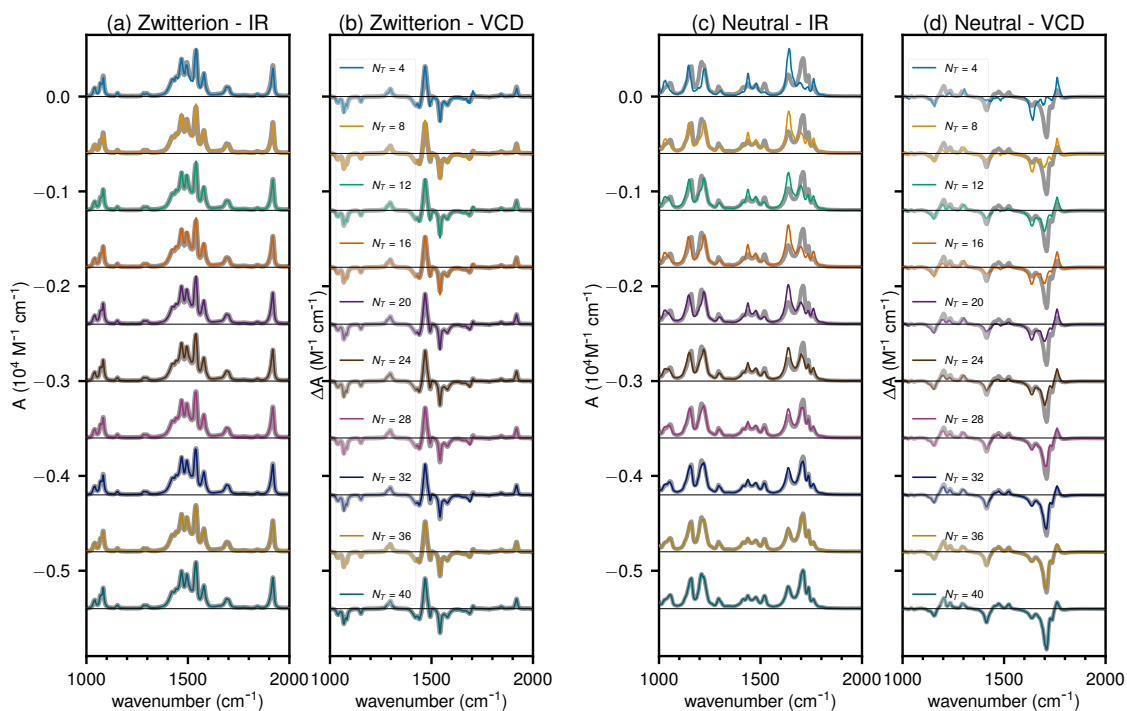


Figure 4.3: IR and VCD spectra from the average of 2 to 40 FFMD trajectories at 300 K of (a,b) the zwitterionic and (c,d) the neutral forms of alanine.

The average of the spectra from 40 FFMD trajectories is considered converged, thus the spectra of $N_T < 40$ are compared to this averaged spectrum so as to determine their normalised error. The idea is to represent quantitatively the convergence with an increasing number of trajectories. For the IR spectra the standard error method is employed to determine the error between the converged spectrum and the others. For the VCD spectra, the SimVCD method [9] is used, which computes a quantitative similarity between two VCD spectra. With these methods the normalised error can be determined with the following equations:

$$E_{\text{IR}} = \frac{A_{12}}{(A_{11} + A_{22})} \quad (4.3)$$

$$E_{\text{VCD}} = 1 - \frac{Z_{12}}{(Z_{11} + Z_{22} - |Z_{12}|)} = 1 - \text{SimVCD} \quad (4.4)$$

where A_{ij} is the quadratic difference between the IR spectra i and j , A_{ii} their quadratic

norm, Z_{ij} the overlap integral between the VCD spectra i and j (also for $i = j$):

$$A_{ij} = \int |I_i(\omega) - I_j(\omega)|^2 d\omega, \quad (4.5)$$

$$A_{ii} = \int I_i(\omega)^2 d\omega, \quad (4.6)$$

$$Z_{ij} = \int \text{VCD}_i(\omega) \times \text{VCD}_j(\omega) d\omega, \quad (4.7)$$

$$Z_{ii} = \int \text{VCD}_i(\omega)^2 d\omega \quad (4.8)$$

Where $I_i(\omega)$ is the IR spectrum intensity at frequency ω and $\text{VCD}_i(\omega)$ the VCD spectrum intensity. Here the averaged spectrum that is considered converged, the so-called complete spectrum and the average of N_T randomly selected trajectories, the so-called partial spectrum are compared. $Y_C(\omega)$ denotes the complete spectrum at frequency ω and $Y_P(\omega)$ denotes the partial spectrum at frequency ω . As a discrete number of frequencies is taken, the integrals are converted into the sums shown in equations (4.9) and (4.10).

As mentioned previously, N_T trajectories are selected randomly for the partial spectrum. This allows us to use a Monte Carlo process to determine the error of the normalised errors. This Monte Carlo process relies on repeated random selections to obtain statistical results. Thus these normalised errors also have their own statistical error calculated with equation (4.11), giving the fluctuations over $m = 10\,000$ trials.

$$E_{\text{IR}} = \frac{\sum_i^\omega (Y_C(i) - Y_P(i))^2}{\sum_i^\omega Y_C^2(i) + \sum_i^\omega Y_P^2(i)} \quad (4.9)$$

$$E_{\text{VCD}} = 1 - \frac{\sum_i^\omega (Y_C(i) \times Y_P(i))}{\sum_i^\omega Y_P^2(i) + \sum_i^\omega Y_C^2(i) - \left| \sum_i^\omega (Y_C(i) \times Y_P(i)) \right|} \quad (4.10)$$

$$\text{Fluc}(E = E_{\text{IR}} \text{ or } E_{\text{VCD}}) = \sqrt{\left| \frac{\sum_i^m E^2}{m} - \left(\frac{\sum_i^m E}{m} \right)^2 \right|} \quad (4.11)$$

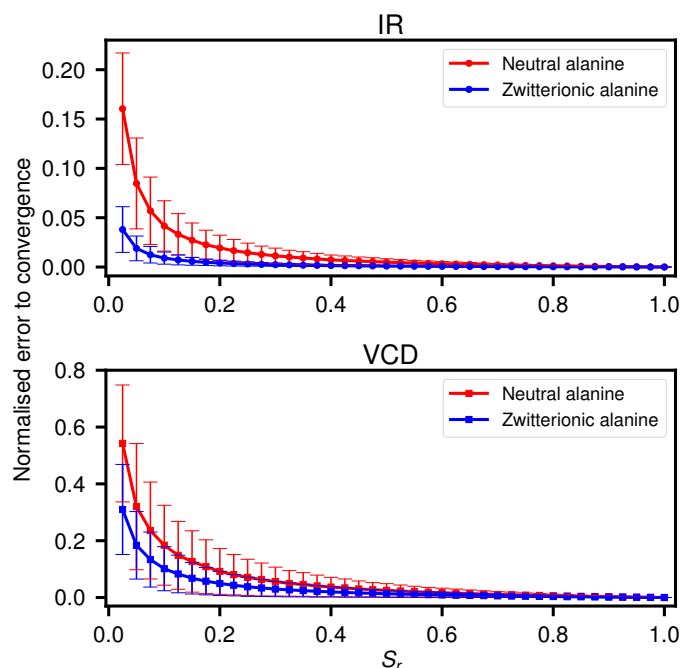


Figure 4.4: Normalised error of the average in the IR and VCD spectra of isolated alanine from N_T FFMD trajectories at 300 K of 200 ps each, compared to the average from 40 trajectories, as a function of the sampling ratio $S_r = N_T/40$.

The variations of the normalised error, shown as a function of the sampling ratio $S_r = N_T/40$, demonstrate that the sampling must be performed for a certain amount of time to reach convergence, see figure 4.4. When the sampling ratio reaches 0.5, hence when the system has been sampled for 20 trajectories of 200 ps which corresponds to 4 000 ps, the spectrum is considered to be converged.

To check whether there is an influence on the spectra depending on which trajectories are chosen, 5 trajectories were selected randomly between 40 trajectories for the zwitterionic and neutral forms of alanine. This was performed 3 times (sample 1-3). The average of 5 trajectories for each sample is shown in figure 4.5. Choosing different trajectories influences the intensity of the peaks especially for alanine in the neutral form. This confirms that the sampling needs to be performed over multiple trajectories to correctly access the spectral variations of the system.

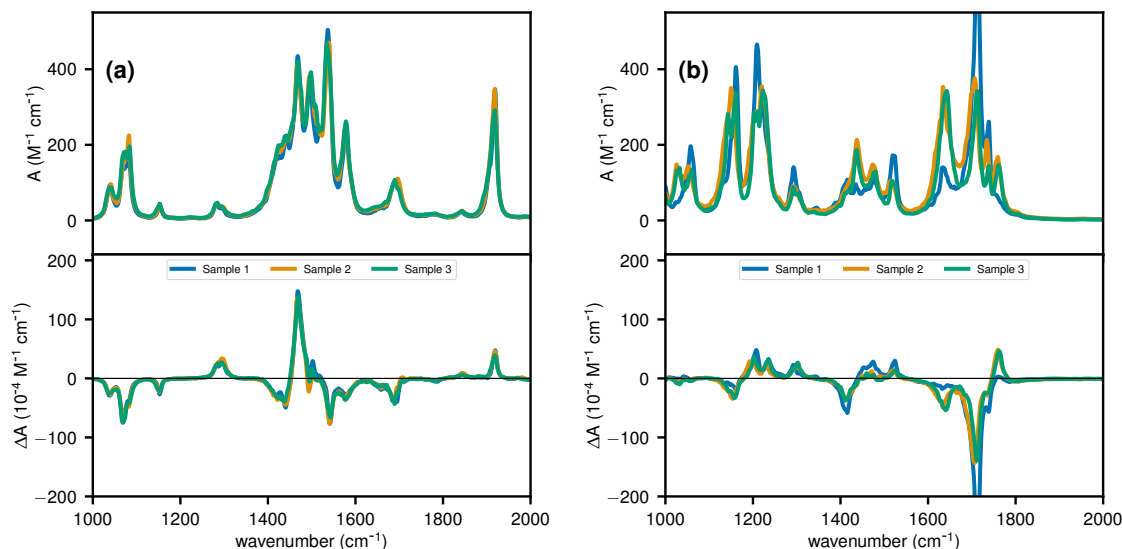


Figure 4.5: IR and VCD spectra from the averages of 5 different trajectories of zwitterionic (left) and neutral forms of alanine (right), as obtained from FFMD trajectories at 300 K.

4.2. Comparison with electronic structure calculations

4.2.1. Structures and relative energies

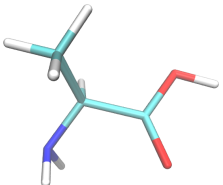
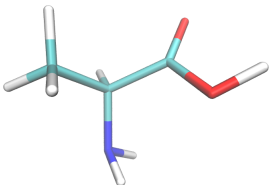
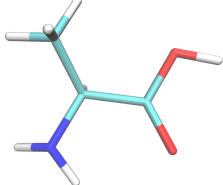
Each configuration of an FFMD simulation of 1 ns of alanine in the neutral form was optimised by the force field and 6 main configurations could eventually be identified from this analysis. A 7th configuration was added, which corresponds to the lowest energy configuration from the article of Blanco et al. [5]. This configuration turns out not to be locally stable with AMOEBA, despite efforts to improve the parameters of the NH₂ and COOH groups. This is probably due to the NH₂ parameters not being fully optimised. Only the neutral form was studied as the zwitterionic form is unstable: during the optimisation, the hydrogen moves back to the oxygen thus the neutral form is recovered. Rahmani et al. [1] argued that the change of structure is due to intramolecular hydrogen bonding between the hydrogen of the amino group and the oxygen of the carboxylic group followed by a rotation of the COO⁻ group.

These configurations were further classified into families based on their structural similarities, see Table 4.1. The configurations within each family differ mainly through the rotation of the CH₃ group. The families differ in the position of the OH group compared to the NH₂ group and the positions of the nitrogen hydrogens compared to the main carbon chain.

DFT calculations at the B3LYP/6-311++G(d,p) level of theory were performed for each of the 7 configurations using the Gaussian 16 software [10]. The energy of each configuration was then recalculated with the force field to compare the values of the relative energy and the electric dipole moments.

For each family, the root-mean-square deviation (RMSD) was calculated taking the configuration of family 1 as a reference. For the other families, the structure used to determine the values of the RMSD, the relative energy and the permanent dipole moment corresponds to the structure with the lowest energy of the family. Families 1 and 3 have a similar orientation of the COOH group, thus their RMSD are closer than for the families 1 and 2. The energy of each family is close, this is mainly due to the hydrogen bond always being present between NH₂ and an oxygen. The charges of the force field are described with a set of point charges, dipoles and quadrupoles centered on each atom, which give the electric dipole. In DFT, the dipole moment is defined through the electronic density. Thus the dipole moments differ as the distribution of charges are not considered in the same way. However the values of the dipole moments follow the same trend for DFT and the force field, which gives a first validation of the dipole moments.

Table 4.1: The 3 families of configurations found for neutral alanine at the DFT/B3LYP/6-311++G(d,p) level of theory. Also included is the RMSD value compared to the 1st configuration, the relative energy (ΔE) calculated with DFT and with the force field (in kcal/mol) and the permanent electric dipole moment (μ^{tot}) calculated by the force field and DFT (in Debye).

		
Family 1 (1 configuration)	Family 2 (2 configurations)	Family 3 (4 configurations)
RMSD = 0.0	RMSD = 1.2257	RMSD = 0.9258
$\Delta E_{\text{FF}} = 0.00$	$\Delta E_{\text{FF}} = 0.61$	$\Delta E_{\text{FF}} = 0.48$
$\Delta E_{\text{DFT}} = 0.00$	$\Delta E_{\text{DFT}} = 1.06$	$\Delta E_{\text{DFT}} = 1.45$
$\mu_{\text{FF}}^{\text{tot}} = 0.450$	$\mu_{\text{FF}}^{\text{tot}} = 1.370$	$\mu_{\text{FF}}^{\text{tot}} = 1.859$
$\mu_{\text{DFT}}^{\text{tot}} = 1.298$	$\mu_{\text{DFT}}^{\text{tot}} = 1.626$	$\mu_{\text{DFT}}^{\text{tot}} = 2.144$

4.2.2. IR and VCD spectra from static DFT calculations

For each of the families shown previously, IR and VCD spectra were determined. The harmonic frequencies were scaled by 0.98 to correct for anharmonicity and basis set incompleteness [11, 12]. The anharmonic spectra were obtained using the second order perturbation theory (VPT2) described in chapter 2. No scaling factors were applied to anharmonic frequencies. The final vibrational spectra were obtained by convoluting the intensities with a Lorentzian line shape (FWHM 10 cm⁻¹).

The comparison of spectra from the harmonic and anharmonic frequencies, in figure 4.6, shows that anharmonicity mainly affects the spectra at 1300 cm⁻¹ with a band inversion in the VCD spectrum for all 3 configurations.

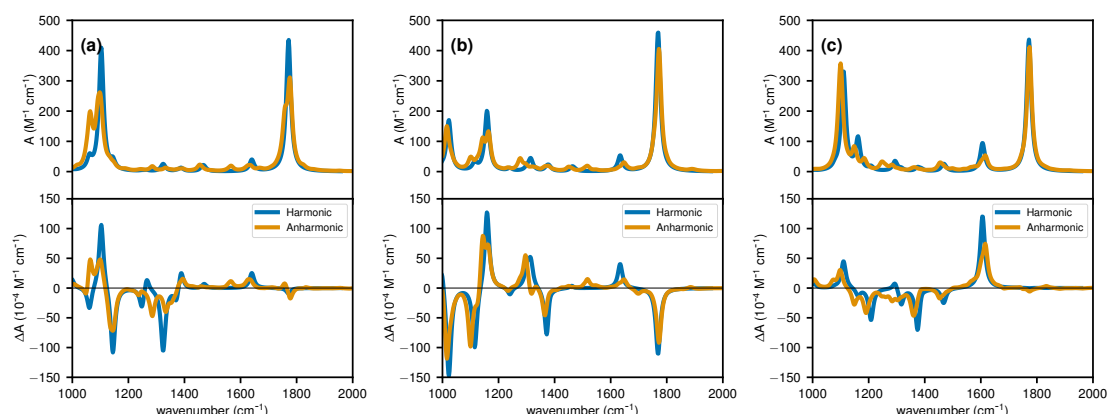


Figure 4.6: Comparison of harmonic and anharmonic spectra obtained from static DFT calculations, for the (a) family 1, (b) family 2 and (c) family 3. The harmonic frequencies are multiplied by a scaling factor of 0.98. No scaling was applied to anharmonic frequencies.

The three VCD spectra show different signals at 1800 cm^{-1} , this is due to the different positions of the C=O group compared to the CH₃ group. The bands of the NH₂ scissor vibration have the same sign but the relative intensity is higher when the NH₂ group is aligned with the C=O group. Families 1 and 3 have the same sign for the bands around 1100 cm^{-1} .

The values of the frequencies of the most intense bands in the IR spectra and their assignments are reported in Table 4.2. The assignments were determined by Rahmani et al. [1]. For each value of frequency reported here, the vibrational modes were examined and similar values of frequencies are found for each type of vibration. This is not surprising as very similar methods are used to determine the frequencies and their assignments.

Table 4.2: A comparison of harmonic scaled frequencies (in cm^{-1}) with different basis sets and experimental frequencies (in cm^{-1}) and their assignments. Rahmani et al. [1] used a 0.97 scaling factor whereas a scaling factor of 0.98 is used here as the level of theory differs.

B3LYP/6-311++G(d,p) family 1	B3LYP/6-311G+(d,p) [1]	Experimental [3]	Assignments [1]
1061	1057	1064	CH ₃ wagg + CH bend + NH ₂ wagg
1103	1103	1110	OH bend + CH bend + CH ₃ wagg + NH bend
1144	1126	1206	OH bend + CH bend + CH ₃ wagg + NH bend
1640	1602	1622	NH ₂ scissor
1771	1764	1774	C=O stretch

bend: bending; wagg: wagging; stretch: stretching.

As the energy difference between each family is low, the assumption is made that they coexist. Thus each configuration needs to be included to form a realistic spectrum. For this purpose an average over the spectra from the 7 conformations was performed. To make the spectra more realistic, weights were used to give more importance to the

spectra from configuration that could exist with a higher probability. These weights can be determined in two ways. The first uses the MD data, where the number of times each configuration was present in the 1 ns trajectory is considered. These will be called the MD weights. The second way relies on Boltzmann weights and requires, for each configuration α , the values of the partition function Z representing the thermal populations at canonical equilibrium. These will be called the DFT weights. In the harmonic approximation, the partition function Z is defined as the following:

$$Z_{\alpha} = e^{-\beta E_{\alpha}} \prod_{i=1}^{\kappa} \frac{e^{-\beta h \nu_{\alpha}^i / 2}}{1 - e^{-\beta h \nu_{\alpha}^i}} \quad (4.12)$$

To account for intrinsic anharmonic effects, in equation (4.12) the (unscaled) anharmonic frequencies, ν_{α}^i , $i = 1, \dots, \kappa = 3N - 6$, determined from our VPT2 calculations is used. β corresponds to $1/k_B T$ and h is Planck's constant.

The average of the 7 spectra using the DFT and MD weights are presented in figure 4.7. The FFMD spectrum shown previously is also represented.

The average from the MD data corresponds mainly to the spectrum of family 3 (94.3%) and family 2 (5.7%) is present as a minority in the FFMD simulation. The weight of family 1 is zero as its structure is not locally stable with AMOEBA. The average from the DFT data has family 1 being the main contributor (49%) followed by family 2 (30.5%) and finally family 3 (20.5%). This implies that the force field and DFT favour the hydrogen from the carboxyl group positioned far from the nitrogen. The amino group is favoured by DFT to be perpendicular to the carboxyl group, whereas the force field favours an aligned configuration. The parameters of the NH_2 group are probably not optimal. Trials were performed to improve these parameters by varying the values corresponding to the angle, torsion and multipole terms. However only a slight improvement was noticed.

In the region 1600 cm^{-1} – 1800 cm^{-1} , the bands from the DFT spectra are assumed to correspond to the peaks at 1700 cm^{-1} from the FFMD results. The sampling from the FFMD leads to broader peaks due to anharmonicity and temperature effects. The bands from the DFT results have been artificially widened through a convolution of the intensities with a Lorentzian line shape. However the FFMD and DFT are still quite different. This originates from the exploration of different configurations. As shown in figure 4.6, the spectra from each of family are very different. Neutral alanine is very flexible and is composed of a large number of configurations that each method explores differently.

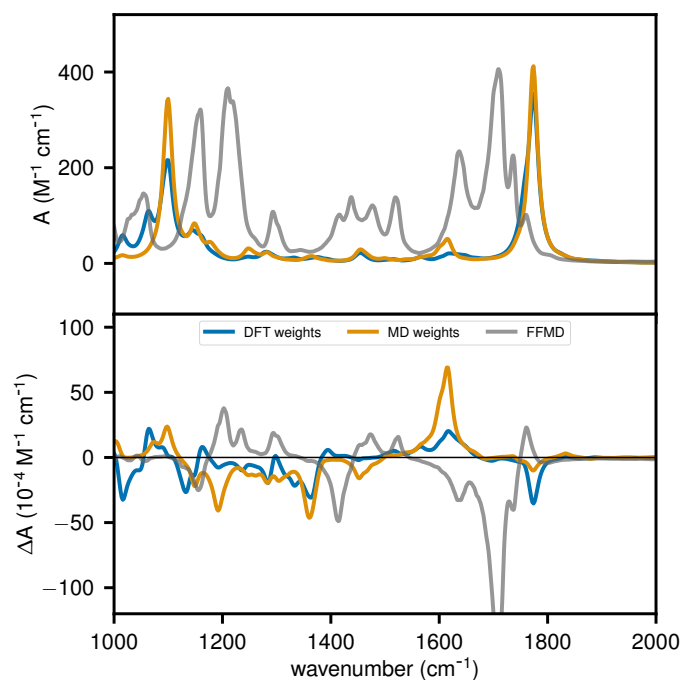


Figure 4.7: IR and VCD spectra obtained by averaging anharmonic DFT spectra from the 3 families using weights from the MD and DFT data. Represented in grey is the FFMD spectrum from the average of 40 trajectories.

4.3. Assessing moments and spectra from one trajectory

The electric and magnetic dipole moments determined with force field molecular dynamics are compared to the predictions of ab initio MD. For a robust comparison, a single MD trajectory (produced by the force field) is exploited. To be able to compare the electric and magnetic dipole moments of the two methods, they need to be determined from the same trajectory. The trajectory includes the positions and velocities throughout the MD simulation. Thus for both the zwitterionic and neutral forms, a trajectory of 200 ps at 300 K was determined with FFMD giving the dipole moments. The dipole moments were then computed again using now DFT within the CPMD software, by our collaborator S. Jähnigen. As the trajectory has already been determined, the DFT method using CPMD is equivalent to performing a DFT calculation for each time step, where the Troullier–Martins pseudopotentials [13] are used and the plane wave basis has a cutoff of 70 Ry. The trajectory was sampled with a time resolution of 4 fs.

From these dipole moments, the IR and VCD spectra can be determined using the Fourier transform of the corresponding time correlation functions. The spectra from both of these methods are shown in figure 4.8. For the zwitterionic form (Fig. 4.8a), the force field gives rise to more activity in the 1410-1450 cm⁻¹ range and a lower intensity at 1925 cm⁻¹. The change in intensity can come from the difference in the description of the charges on the carboxylic acid. In the VCD spectra there is an inversion of sign at

1050 cm^{-1} . For the neutral form (Fig. 4.8b), the VCD signs are all the same, only the intensities vary in the IR and VCD spectra.

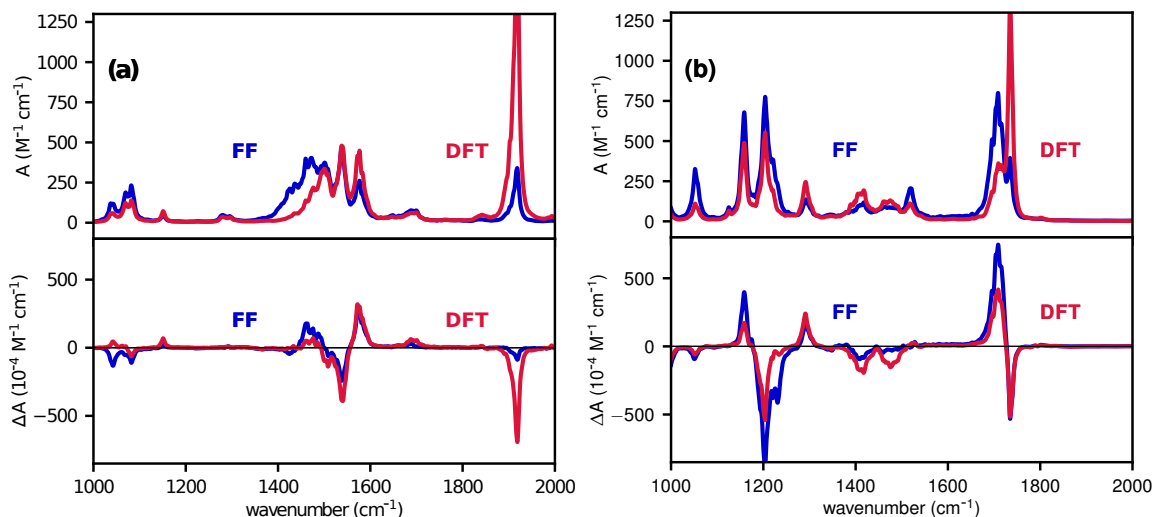


Figure 4.8: Comparison of the IR and VCD spectra of (a) zwitterionic and (b) neutral forms of alanine using the dipole moments calculated with the force field and with DFT from the same FFMD trajectory at 300 K.

In a second step, the dipole moments calculated with both methods using the same trajectory are compared. In anticipation of the results concerning the crystalline phase, the values of the electric dipole moment $\vec{\mu}$ are shifted to the origin. This is to avoid a poorly correlated electric dipole moment due to different molecular orientations in the crystalline phase. The dipole moments are represented by their values on the x , y and z axis and compared using a Pearson correlation coefficient r . The closer the values of the dipole moments from the force field and from DFT are from one another, the closer the value of the correlation coefficient is to 1. The values of the electric dipole moments in the position $\vec{\mu}$ and velocity forms $\dot{\vec{\mu}}$, and the magnetic dipole moment \vec{m} are represented for the the zwitterionic form of alanine in figure 4.9 and for the neutral form of alanine in figure 4.10.

For the electric dipole moment of the zwitterionic form of alanine, the signal is more disperse along the y axis. This is because there is a slight rotation of the molecule in the trajectory. This dispersity does not appear for the time derivative of the electric dipole moment and the magnetic dipole moment as the distribution of these dipole moments are centrosymmetric. No rotation is shown in the dipole moments of the neutral form of alanine.

The value of r is high for the electric dipole moment and satisfactory for its time derivative and the magnetic dipole moment of the zwitterionic form. These values are improved for the neutral form.

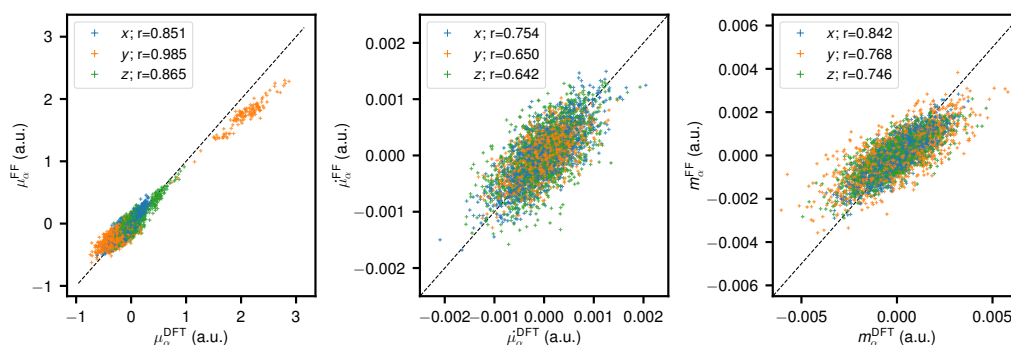


Figure 4.9: Correlation of the projections of the electric dipole moment $\vec{\mu}$ or its time derivative $\dot{\vec{\mu}}$, and the magnetic dipole moment \vec{m} , as obtained for a specific MD trajectory but calculated from the force field (FF) or using DFT for the zwitterionic form of alanine.

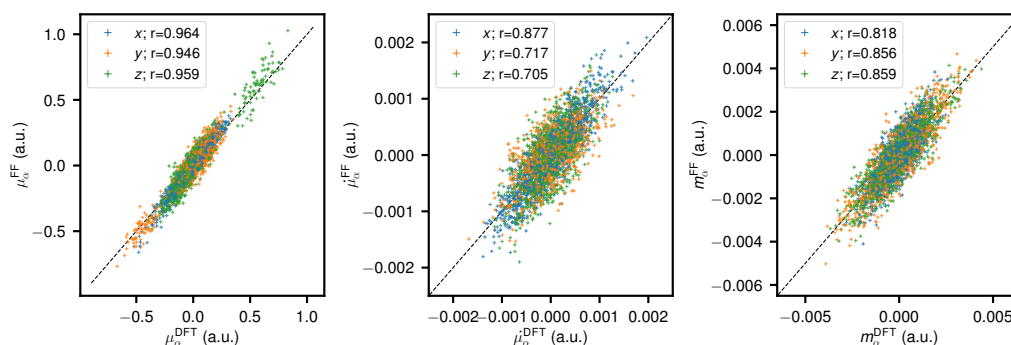


Figure 4.10: Correlation of the projections of the electric dipole moment $\vec{\mu}$ or its time derivative $\dot{\vec{\mu}}$, and the magnetic dipole moment \vec{m} , as obtained for a specific MD trajectory but calculated from the force field (FF) or using DFT for the neutral form of alanine

The results are quite similar between the DFT and FF approaches, which is promising for our new method. The neutral form shows a better correlation with DFT than the zwitterionic form, this is due to the description via point charges of the force field. Despite the charges being expressed by multipoles and by induced dipoles, this description is disadvantaged for the $-\text{COO}^-$ and $-\text{NH}_3^+$ groups. Keeping in mind that the force field was parametrized for solvated molecules, a better description can be expected for the hydrated system compared to the isolated molecule.

4.4. Assignment of the IR spectra into effective modes

The dipole moments and the vibrational spectra were obtained with good accuracy. In this section, the IR spectra are decomposed with the effective modes analysis method explained in chapter 2. The information of the vibrational modes is not included in the time-correlation functions, thus this analysis is needed to attain information about the origin of each peak. The effective modes were determined by analysing one trajectory

(positions and velocities) of 200 ps at 300 K from FFMD. A quick attribution of peaks was fulfilled in sections 4.1.1 and 4.2.2 by comparing our results with the DFT calculations of alanine and with literature [1, 7]. Here a more detailed description is provided.

4.4.1. Analysis of the zwitterionic and neutral forms of alanine

Using the configuration and atom numbering of figure 4.1, the internal coordinates reported in Table 4.3 were used to initialise the effective modes analysis.

bond lengths	valence angles	dihedral angles
N1-C2	C3-C2-N1	O4-C3-C2-N1
C2-C3	O4-C3-C2	H5-N1-C2-C3
C3-O4	C2-N1-H5	H8-C2-N1-H7
N1-H5 (neutral: O4-H5)	C2-N1-H6	C9-C2-N1-H7
N1-H6	C2-N1-H7	H10-C9-C2-N1
N1-H7	N1-C2-H8	(neutral: H5-O4-C3-O13)
C2-H8	C9-C2-N1	
C2-C9	H10-C9-C2	
C9-H10	H11-C9-C2	
C9-H11	H12-C9-C2	
C9-H12	O13-C3-C2	
C3-O13	H5-N1-H6 (neutral: none)	
	H7-N1-H6	
	H12-C9-H11	
	H10-C9-H11	
	O13-C3-O4	

Table 4.3: List of the 33 internal coordinates used to initialise the effective modes analysis of alanine in the zwitterionic form (12 bond lengths, 16 valence angles, 5 dihedral angles). The labels of the atoms refer to Fig. 4.1(b). For the neutral form, the bond N1-H5 is replaced by the bond O4-H5, the valence angle H5-N1-H6 is suppressed, and the dihedral angle H5-O4-C3-O13 is added.

In figure 4.11, the IR spectrum from the previous FFMD simulations is shown and the peak for each effective mode is represented underneath, these are the power spectra of the effective modes. The effective modes analysis gives, in addition to a local spectrum for each mode, a representation of the corresponding vibration using vectors that is shown in figures 4.12 to 4.14. Note that the power spectra are shown without considering the $\langle |\frac{\partial \bar{\mu}}{\partial q_i}(0)|^2 \rangle$ prefactor, from equation (2.91) in chapter 2, which would alter the relative intensity of the peaks.

The functional $\Omega^{(n)}$ was tested with $n = 2$ and $n = 4$ and no major differences were found, thus the results are reported only for $n = 2$ in both zwitterionic and neutral forms of alanine. Note that the minimisation procedure may depend on the choice of initial internal coordinates ζ . Also in this case various options were tested and only one set of results is reported below (no major differences were found among the various tests performed).

4.4.1.1 Analysis of the zwitterionic form

In the low frequency region of Fig. 4.11, four bands can be clearly identified: two overlapping bands in the region $1000\text{--}1100\text{ cm}^{-1}$, one band in the region $1100\text{--}1200\text{ cm}^{-1}$, and one band localised at 1300 cm^{-1} . The four bands see the participation of the CH_3 , NH_3 and CO_2 groups, together with contributions from the vibration of the CH bond. This finding is confirmed in figure 4.12, where the COO motion is seen to participate to modes 1 and 4.

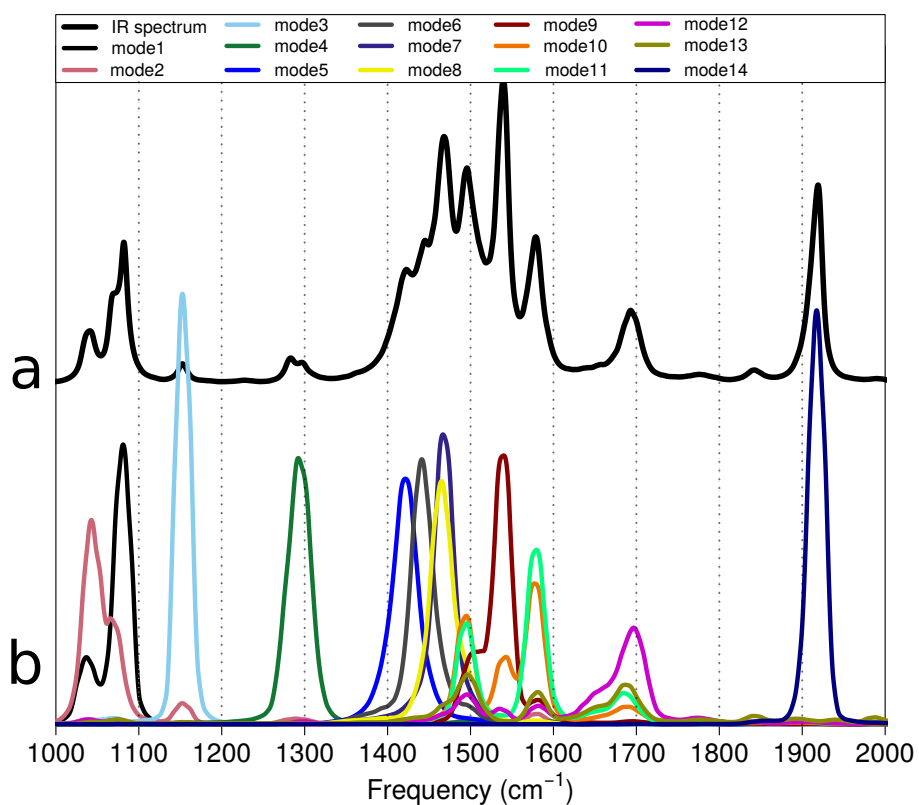


Figure 4.11: (a) IR spectrum of isolated alanine in zwitterionic form, as obtained from FFMD trajectories at 300 K and (b) its decomposition in effective modes, where the different power spectra are shown in color. The power spectra have arbitrary intensities.

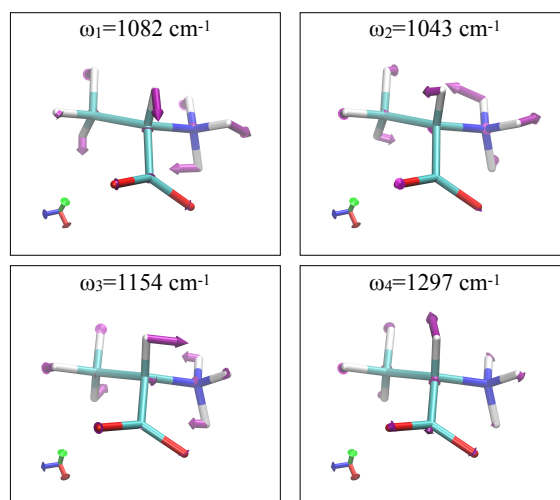


Figure 4.12: Effective modes associated to the spectral region 1000–1350 cm^{-1} of the IR spectrum of zwitterionic alanine in the gas phase. The sticks represent alanine in zwitterionic form. The arrows illustrate the collective atomic displacements along the effective modes.

The region 1350–1500 cm^{-1} of the spectrum in figure 4.11 is populated by four main bands, associated to the 4 modes shown in figure 4.13. This region of the spectrum sees the main participation from the bending and umbrella modes in the two fragments CH_3 and NH_3 . Some distorted COO bending motions contributes as well to the higher frequency modes (modes 7 and 8 in figure 4.13), giving rise to the overlapping band slightly above 1450 cm^{-1} in figure 4.11.

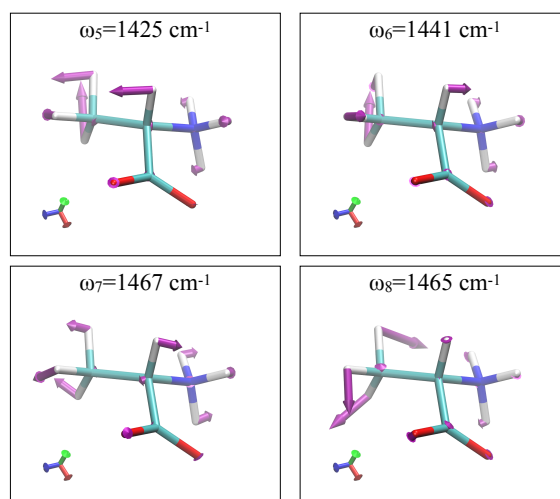


Figure 4.13: Effective modes associated to the spectral region 1350–1500 cm^{-1} of the IR spectrum of zwitterionic alanine in the gas phase.

In the remaining part of the spectrum of zwitterionic alanine in the gas phase, five modes are identified between 1500–1900 cm^{-1} . The high energy band at 1900 cm^{-1} is easily assigned to the asymmetric stretching motion in the COO fragment. This is mode 13, represented in figure 4.14.

At 1700 cm^{-1} , the band corresponding to mode 12 in figure 4.14 is identified. This band is produced by the excitation of the bending and the umbrella motion in NH_3 (see figure 4.18 for comparison), which is now coupled to the bending motion in the CH_3 fragment as well as to the vibration of the CH bond which is not involved in CH_3 .

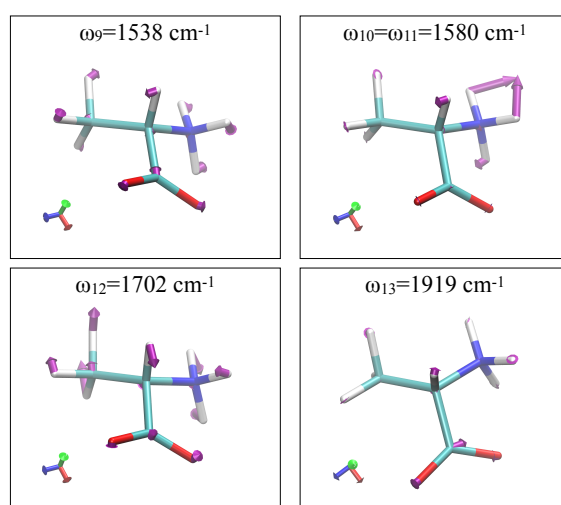


Figure 4.14: Effective modes associated to the spectral region 1500 - 1900 cm^{-1} of the IR spectrum of zwitterionic alanine in the gas phase.

The three remaining bands are located in the region 1500–1600 cm^{-1} , and are mainly produced by the bending and umbrella vibrational modes of the NH_3 fragment. These bands and these effective modes are labelled 9, 10 and 11 in figure 4.14.

The IR data reported in reference [3] on the zwitterionic and neutral forms of alanine is compared to the results of our effective modes analysis in terms of frequencies and bands assignments. Note, however, that in reference [3] the bands assignments were performed based on harmonic normal modes calculations with electronic energies calculated at the Hartree-Fock level. The modes assignment is presented in the central column, where the notation for the modes used in reference [3] is adopted. In Table 4.4, the results for the zwitterionic form of alanine are reported.

Experimental frequencies (cm ⁻¹)	Modes	Theoretical frequencies (cm ⁻¹)
1014	ρCH_3	1043
1027	δCH	
1114	ρCH_3	1082
1149	ρNH_3	1154
1235	δCH	
1307	$\delta_s\text{NH}_3$	1425 1441
1354	νCOO	1297
1408	$\delta_s\text{CH}_3$	
1452	$\delta_{as}\text{CH}_3$	1467
1523	$\delta_{as}\text{CH}_3$	1465
	$\delta_{as}\text{CH}_3, \delta_s\text{NH}_3$	1538
1592	$\delta_{as}\text{NH}_3$	1580
1603	$\nu_{as}\text{COO}$	1919
	$\omega\text{CH}_3, \omega\text{NH}_3$	1702

Table 4.4: Left column: experimental frequencies collected from Ref. [3] in the spectral region 1000–2000 cm⁻¹ for the zwitterionic form of alanine. Central column: bands assignment using the abbreviation from reference [3] (ρ for rocking, δ for bending, ν for stretching, ω for wagging; s for symmetric and as for asymmetric). Right column: theoretical frequencies from this work.

Most of the modes are represented in both experimental and theoretical results with similar frequencies. However for a few modes, the experimental and theoretical frequencies differ from one another. These modes are the NH₃ symmetric bend, the COO symmetric stretch, the CH₃ antisymmetric bend and the COO antisymmetric stretch. The CH₃ and NH₃ antisymmetric bend and wag are not described by theory.

4.4.1.2 Analysis of the neutral form

As expected from the different structures, the IR absorption spectrum of neutral alanine in the gas phase is very different from that calculated for the zwitterionic form. By comparing figure 4.11 and figure 4.15 some major differences can be pointed out: (i) in the region 1000–1300 cm⁻¹, the low intensity band at 1150 cm⁻¹ in zwitterionic alanine becomes more intense and splits into two contributions for the neutral form; (ii) in the region 1350–1500 cm⁻¹, the large absorption band of zwitterionic alanine becomes less intense in the neutral form, and four main peaks appear better resolved; (iii) the region 1500–1900 cm⁻¹ becomes more IR active for the neutral form of alanine than for the zwitterionic form, with the complete disappearance of absorption above 1800 cm⁻¹.

The effective modes of neutral alanine in the gas phase have been determined in order to identify the modes attesting to these differences between the IR spectra of the two forms of the molecule. For the sake of clarity, the 14 modes determined in the spectral region 1000–2000 cm⁻¹ are not described, but the effective modes mainly responsible for the differences (i), (ii) and (iii) discussed above, are commented.

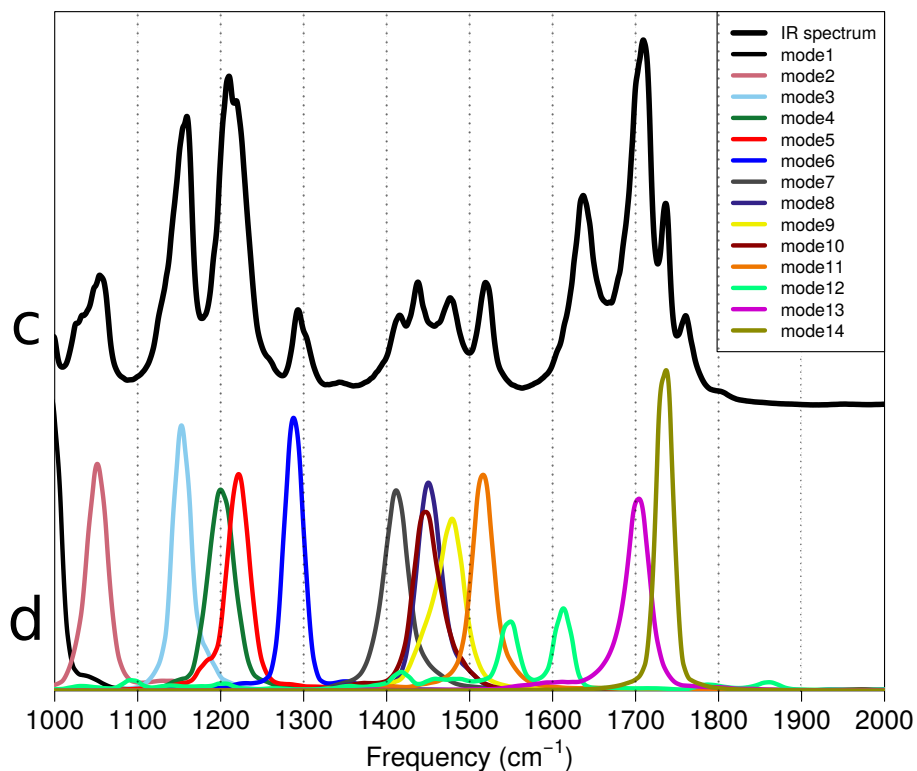


Figure 4.15: IR spectrum of neutral alanine in the gas phase, as obtained from FFMD trajectories at 300 K (black line) along with its decomposition in effective modes.

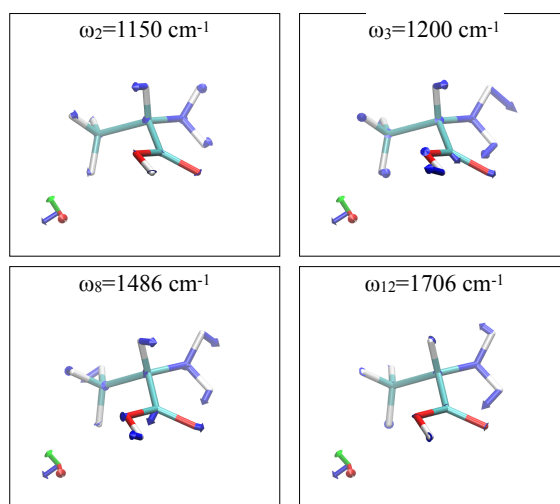


Figure 4.16: Effective modes associated to the spectral region 1000–1800 cm^{-1} of the IR spectrum of neutral alanine in the gas phase, highlighting to the major differences with the IR spectrum of zwitterionic alanine in gas phase.

In figure 4.16, the selected effective modes responsible for the different features of the spectrum of neutral alanine with respect to zwitterionic alanine involve some vibrational motions of the fragments that mainly distinguish the two species: the OH bond in the COOH fragment and the NH₂ fragment itself. Mode 2 of figure 4.16 centered at around 1150 cm⁻¹ is a seemingly wagging motion of the NH₂ moiety, which is absent in the zwitterionic form. Similarly, the wagging motion of NH₂ at 1200 cm⁻¹, which is coupled to the in-plane oscillations of the OH bond (in plane with COO), is totally absent in the absorption spectrum of zwitterionic alanine of figure 4.11. Mode 8, at around 1480 cm⁻¹, is similar to mode 3, just described, but it is coupled to a vibration in the CH₃ fragment that looks like an umbrella mode. Finally, that mode 12, slightly above 1700 cm⁻¹, clearly involves solely the HNH bending motion in NH₂, and was not so well defined in the analysis reported in figure 4.14.

Similarly to the results reported in Table 4.4, Table 4.5 shows the comparison between the calculated frequencies and bands assignment of the effective modes analysis and the experimental results of reference [3].

Experimental frequencies (cm ⁻¹)	Modes	Theoretical frequencies (cm ⁻¹)
1037	ρ CH ₃	1048
1064	ρ CH ₃	
	δ CH, ω NH ₂	1150
1110	ν CN	
1153	δ CO	
1206	δ COH	1200
		1219
	ρ NH ₂ , ω CH ₃ , δ COH	1285
1335	δ CH	
1376	δ_s CH ₃	1416
1408	δ CH	
1454	δ_{as} CH ₃	1454
1460	δ_{as} CH ₃	1443
	ω NH ₂ , ω COO	1517
	ρ NH ₂ , δ OH	1610
1622	δ NH ₂	1706
1774	ν C=O	1732

Table 4.5: Left column: experimental frequencies collected from reference [3] in the spectral region 1000–2000 =cm⁻¹ for the neutral form of alanine. Central column: bands assignment using the abbreviation from reference [3] (ρ for rocking, δ for bending, ν for stretching, ω for wagging; s for symmetric and as for asymmetric). Right column: theoretical frequencies from this work.

The C=O stretch is a lot closer to the experimental value here than the COO antisymmetric stretch for the zwitterionic form. The force field must better represent the C=O bond rather than the CO bond in the COO⁻ group.

Thus the IR spectra of alanine in the zwitterionic and neutral forms have been analysed, and the corresponding molecular vibrations were given. For a finer analysis of these results for the zwitterionic form of alanine, the effective modes analysis will also be performed on the 3 fragments that constitute the molecule.

4.4.2. Fragment analysis

In this section, some of the molecular fragments of alanine in its zwitterionic form are analysed. Since the spectral region of interest, $1000\text{--}2000\text{ cm}^{-1}$, is densely populated by overlapping and not well-defined absorption bands, this secondary study aims to clarify the results of the previous sections on the full molecule. From the entire FFMD trajectory, positions and velocities of three fragments, i.e., CH_3 , NH_3 , and COO are extracted in order to identify their effective modes and the bands of the full IR spectrum arising from their vibrations. Note that, while the analysis is performed on the isolated fragments, their dynamics accounts for the presence of the remaining degrees of freedom of the molecule.

4.4.2.1 IR analysis of the methyl group

To determine the effective-mode power spectrum of the methyl group, the following internal coordinates to determine the effective modes are introduced.

modes	atoms
stretch	C1-H2
stretch	C1-H3
stretch	C1-H4
bend	H2-H3-H4
torsion	H2-C1-H3-H4
torsion	H2-C1-H4-H3

Table 4.6: List of CH_3 internal coordinates used for the specific analysis of the methyl group power spectrum.

The IR absorption bands of the CH_3 fragment are mainly located in the region of $1400\text{--}1500\text{ cm}^{-1}$, as shown by the coloured bands in the lower panel of figure 4.17, and thus contribute to the large absorption region of zwitterionic alanine in gas phase in the region $1300\text{--}1600\text{ cm}^{-1}$ (black line in the upper panel of figure 4.17). The effective modes identified by this analysis are, unsurprisingly, combinations of bending modes (mode 1 and mode 3, depicted as arrows in the insets of figure 4.17) and the umbrella mode (mode 2 in figure 4.17).

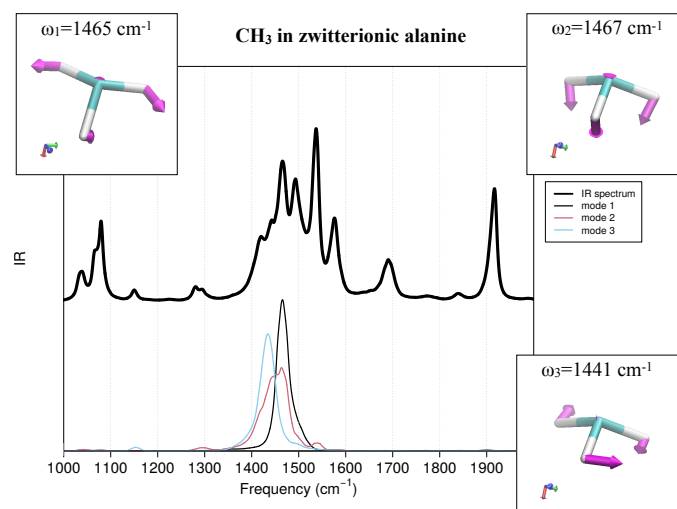


Figure 4.17: Black line: IR spectrum of isolated alanine in its zwitterionic form, as obtained from FFMD trajectories at 300 K. Coloured lines: power spectra determined by the effective modes analysis, corresponding to the 3 modes of the CH₃ fragment. Insets: stick representation of the CH₃ fragment, along with the effective modes, depicted as arrows.

4.4.2.2 IR analysis of the amino group

For the effective modes analysis of the amino group of zwitterionic alanine, the following internal coordinates are used.

modes	atoms
stretch	N1-H2
stretch	N1-H3
stretch	N1-H4
bend	H2-N1-H4
bend	H3-N1-H4
torsion	H2-H3-H4-N1

Table 4.7: List of NH₃ internal coordinates in zwitterionic alanine used for the specific analysis of the amino groups power spectrum.

The IR absorption bands of the NH₃ fragment are slightly blue-shifted with respect to those of CH₃, and are mainly distributed in the region 1500–1700 cm⁻¹, as shown by the coloured bands in the lower panel of figure 4.18. Despite the blue shift, one might expect that the vibrations within the NH₃ fragment are coupled to those of CH₃ since the absorption bands partially overlap. Therefore, NH₃ contributes, similarly to CH₃, to the large absorption region of zwitterionic alanine in gas phase in the region 1300–1600 cm⁻¹ (black line in the upper panel of figure 4.18) as well as to the isolated band at 1700 cm⁻¹. The bands are not well localised in frequency as it was the case for the CH₃ fragment, which might be a manifestation of strong anharmonicity and coupling of NH₃ with other modes that have not been considered in this fragment analysis. The effective modes

identified by the analysis are very similar to those in CH_3 , namely they are a HNH bending mode (mode 1, depicted as arrows in the inset of figure 4.18), the umbrella mode (mode 2 in figure 4.18), and a combination of a HNH bending mode with an out-of-plane vibration of the remaining NH bond (mode 3 in figure 4.18).

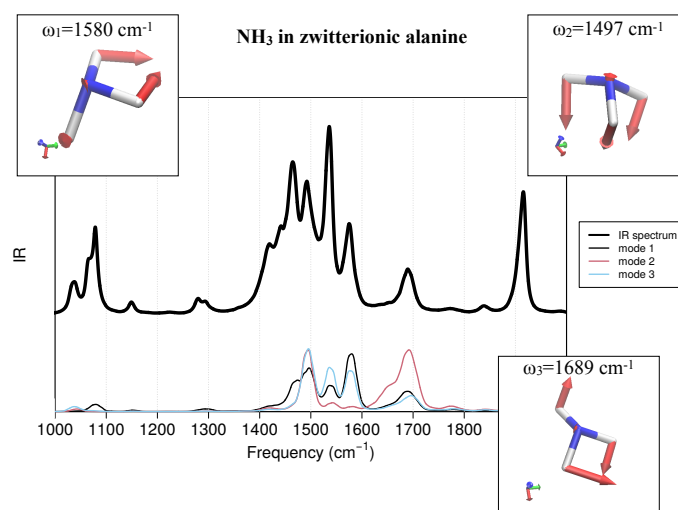


Figure 4.18: Black line: IR spectrum of isolated alanine in its zwitterionic form. Coloured lines: power spectra determined by the effective modes analysis, corresponding to the 3 modes of the NH_3 fragment. Insets: stick representation of the NH_3 fragment, along with the effective modes, depicted as arrows.

4.4.2.3 IR analysis of the carboxyl group

The following internal coordinates are used to determine the effective modes of the carboxyl group.

modes	atoms
stretch	C1-O2
stretch	C1-O3
bend	O2-C1-O3

Table 4.8: List of COO (second column) internal coordinates in zwitterionic alanine used for the specific analysis of the carboxylic acid groups power spectrum.

The IR absorption of the COO fragment is composed of two main bands: one band appears essentially in the region $1500\text{--}1600\text{ cm}^{-1}$, but it is coupled to smaller bands at (a bit less than) 1100 cm^{-1} and at 1300 cm^{-1} ; one band is very well defined and localised around 1900 cm^{-1} . These bands due to the power spectra of the effective modes are shown in the lower panel of figure 4.19. Therefore, also for this fragment, the bending motion in COO (mode 1, depicted as arrows in the inset of figure 4.19) is found to participate to the large absorption region of zwitterionic alanine in the gas phase in the region $1300\text{--}1600\text{ cm}^{-1}$ (black line in the upper panel of figure 4.19), while the asymmet-

ric stretching in COO (mode 2, in the inset of figure 4.19) produces an absorption peak in the high frequency region of the analysed spectrum.

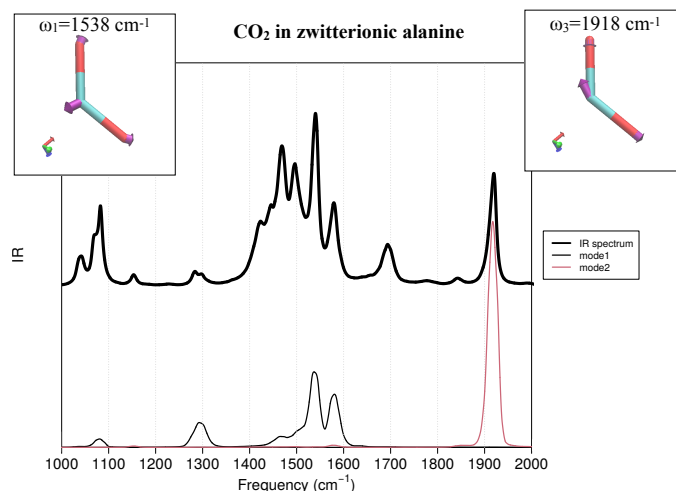


Figure 4.19: Black line: IR spectrum of isolated alanine in its zwitterionic form. Coloured lines: power spectra determined by the effective modes analysis, corresponding to the 2 modes of the COO fragment. Insets: stick representation of the COO fragment, along with the effective modes, depicted as arrows.

The reported decomposition of the spectra of alanine confirms the capability of the effective modes analysis to unravel absorption spectra that can appear at first glance very broad and not well resolved. Since alanine is quite flexible and its dynamics was simulated at room temperature, the normal modes analysis would not be satisfactorily characterise the complex absorption spectra discussed in the present work. The comparison between the two gas phase forms of alanine allowed us to identify the modes responsible for the changes in the IR spectra: even though it is expected that the changes in the neutral form are produced by the fragments of the molecule that are not present in the zwitterionic form, it is interesting to find that those changes are confirmed based on a rigorously-derived numerical procedure. The results from this effective modes analysis can be used to check the value of the frequency corresponding to each vibration. If the values of the frequencies are shifted from their experimental value, the parameters from the force field can be adjusted in consequence.

4.5. Concluding remarks

The results using our implementation of polarisable force field molecular dynamics for vibrational circular dichroism have been scrutinised on the example of the alanine molecule in the gas phase, in its neutral and zwitterionic forms. The importance of accounting for polarisation effects through the induced dipoles moments was demonstrated. The fluctuations of the charges need to be accounted for to represent com-

pletely the VCD spectrum. Through scrutinising the convergence and the error to this convergence, the need to sample over a certain period of time was demonstrated. This sampling time ($20 \times 200 = 4$ ns) is much longer than what is usually achievable with AIMD simulations. Harmonic and anharmonic DFT results were compared, giving a small difference between both spectra, showing the influence of anharmonicity on the IR and VCD spectra.

To confirm the FFMD results, a comparison with ab initio molecular dynamics for one trajectory was accomplished to determine the moments and spectra. This comparison permits the validation of the implementation in the gas phase.

This chapter also discussed the molecular origin of the spectroscopic signals of alanine in the gas phase based on an effective modes analysis aimed to decompose the IR absorption spectra of alanine in zwitterionic and (stable) neutral forms. The decomposition relies on the identification of well-defined, maximally localised in frequency, absorption bands calculated as the power spectra of the effective modes, whose sum yields a good approximation of the full IR spectrum for the molecule at hand. The comparison to experimental frequencies displays a small number of discrepancies which hopefully should be solved in the future by adjusting the force field parameters.

Experimentally, VCD spectra are recorded in dense media where the signals are far more intense. The computational method presently assessed in the gas phase must therefore be extended to the relevant liquid and crystalline cases. This is the subject of the next chapter.

Bibliography

- [1] M. Rahmani and M. E. A. Benmalti, "Theoretical study of the vibrational properties of L-alanine and its zwitterionic form in the gas phase and in solution," *Journal of Biomolecular Structure and Dynamics*, pp. 1–9, 2021.
- [2] S. Jähnigen, A. Scherrer, R. Vuilleumier, and D. Sebastiani, "Chiral crystal packing induces enhancement of vibrational circular dichroism," *Angewandte Chemie International Edition*, vol. 57, no. 40, pp. 13344–13348, 2018.
- [3] M. T. S. Rosado, M. L. R. Duarte, and R. Fausto, "Vibrational spectra (FT-IR, Raman and MI-IR) of α - and β -alanine," *Journal of Molecular Structure*, vol. 410, pp. 343–348, 1997.
- [4] A. J. Stone, "Distributed multipole analysis, or how to describe a molecular charge distribution," *Chemical Physics Letters*, vol. 83, no. 2, pp. 233–239, 1981.
- [5] S. Blanco, A. Lesarri, J. López, and J. Alonso, "The gas-phase structure of alanine," *Journal of the American Chemical Society*, vol. 126, no. 37, pp. 11675–11683, 2004.
- [6] A. Scherrer, R. Vuilleumier, and D. Sebastiani, "Vibrational circular dichroism from ab initio molecular dynamics and nuclear velocity perturbation theory in the liquid phase," *The Journal of Chemical Physics*, vol. 145, no. 8, p. 084101, 2016.
- [7] Merck, *IR Spectrum Table & Chart*, 2022. <https://www.sigmaaldrich.com/FR/fr/technical-documents/technical-article/analytical-chemistry/photometry-and-reflectometry/ir-spectrum-table> (accessed May 10st, 2022).
- [8] J. Horníček, P. Kaprálová, and P. Bouř, "Simulations of vibrational spectra from classical trajectories: Calibration with ab initio force fields," *The Journal of Chemical Physics*, vol. 127, no. 8, p. 084502, 2007.
- [9] J. Shen, C. Zhu, S. Reiling, and R. Vaz, "A novel computational method for comparing vibrational circular dichroism spectra," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 76, no. 3-4, pp. 418–422, 2010.
- [10] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, "Gaussian16 Revision C.01," 2016. Gaussian Inc. Wallingford CT.
- [11] C. Fábri, T. Szidarovszky, G. Magyarfalvi, and G. Tarczay, "Gas-phase and ar-matrix

SQM scaling factors for various DFT functionals with basis sets including polarization and diffuse functions," *The Journal of Physical Chemistry A*, vol. 115, no. 18, pp. 4640–4649, 2011.

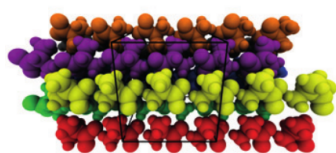
- [12] National Institute of Standards and Technology, *Precomputed vibrational scaling factors*, 2022. <http://cccbdb.nist.gov/vibscalejust.asp> (accessed June 15th, 2022).
- [13] N. Troullier and J. L. Martins, "Efficient pseudopotentials for plane-wave calculations," *Physical Review B*, vol. 43, no. 3, p. 1993, 1991.

Chapter 5

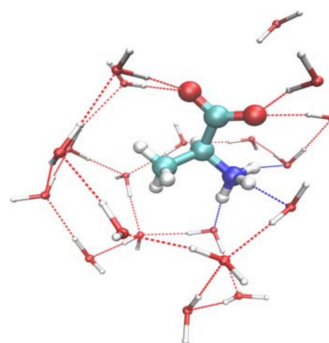
Applications in condensed phases

The previous chapter was aimed at assessing the validity of force field molecular dynamics for VCD spectroscopy, using alanine in the gas phase as our testing ground. Here the scope of the study is extended by focusing on the same molecule but in solid (crystalline) and liquid (solvated) phases. Owing to the usually low intensities of VCD signals, a fair amount of material is needed to actually record them, making condensed phases the natural setting in experiments. As mentioned previously, the correct description of the environment is important in VCD [1].

The VCD spectroscopy of alanine in various condensed phases has been the subject of several earlier works. Jähnigen et al. showed that for crystalline alanine, supramolecular interactions need to be considered to represent properly the VCD spectrum [2]. This work explored the structure of the alanine crystal, depicted in figure 5.1a, where the amino groups are arranged in a helical fashion leading to a chiral supramolecular packing. The supramolecular interactions within a range of 4 to 5 Å give characteristic features of the vibrational spectra.



(a) Unit cell topology of crystalline L-alanine with supramolecular chains along z in the supercell that are collinearly arranged.



(b) L-alanine + 20 water molecules at OPBE0/TZ2P + COSMO level of theory

Figure 5.1: Visual depiction of (a) crystalline [2] and (b) hydrated alanine [3].

In the hydrated phase, various studies examined the number of explicit water molecules needed to embed the alanine solute. With DFT methods, Tajkhorshid et al. demonstrated the need for 4 water molecules [4] whereas Frimand et al. argued the need for 9 explicit water molecules to represent the vibrational spectra [5]. From AIMD simulations Jalkanen et al. found that 20 water molecules are necessary to fully hydrate the zwitteri-

onic species, including 11 solvent molecules on the methyl group, depicted in figure 5.1b [3]. However, they agreed on the importance of an explicit water solvent to capture the effect of the solvent and the chirality transfer effect. In this chirality transfer effect, the chiral properties of the solute are transferred to the solvent [6].

In the present chapter, the assets of FFMD for sampling time and system size are demonstrated for the alanine crystal and for hydrated alanine, both under periodic boundary conditions. An effective modes analysis was also performed to identify the vibrational contributions to the spectrum. To disentangle the chemical effects of hydrogen bonding from those of a physical nature, due to the presence of the embedding medium, alanine in a liquid dinitrogen solvent was also considered. This system was studied under the conditions of temperature and density that are more appropriate to this cryogenic medium. Finally, phenylcyclohexanediol was also studied as a crystal and solvated in dimethyl sulfoxide. This system is of chemical interest owing to the diol group that displays strong hydrogen bonding.

5.1. Crystalline alanine

5.1.1. Vibrational spectra obtained from force field molecular dynamics

In accordance with experimental crystallographic data, an alanine crystal was prepared for simulations. Given the chiral space group of alanine ($P2_12_12_1$), a periodic box of $12.073 \times 12.342 \times 11.57 \text{ \AA}$ was constructed containing 16 alanine molecules, as shown in figure 5.2. It is this system that was used to determine the electric and magnetic dipole moments throughout MD simulations. In the crystal, alanine is in its zwitterionic form so the molecules are closely held together by electrostatic interactions between the amino and carboxylate groups, represented by the black dotted lines for one amino group.

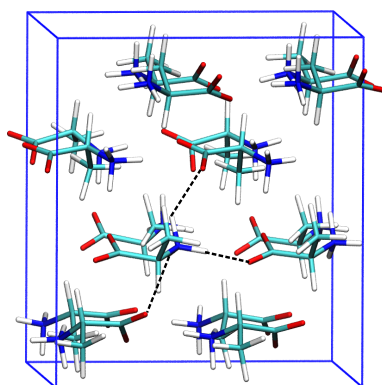


Figure 5.2: Visual depiction of the alanine crystal in its periodic box.

5.1.1.1 Comparing crystalline and gas phase vibrational spectra

Like in the gas phase, a certain number of trajectories was performed to sample the system. Each of these trajectories was initiated from a different starting point, chosen

along a trajectory previously performed in the canonical ensemble. The electric and magnetic dipole moments were retrieved from each trajectory, then the Fourier transform of the corresponding time correlation function was determined. This gave an individual spectrum for each trajectory, all spectra being averaged to obtain a spectrum considered as converged. Under PBC, as explained in chapter 2, the VCD spectrum becomes origin-dependent through the minimum image convention, which changes the value of the position depending on the origin. Thus the nearest periodic image is directly chosen between particles to make the VCD spectrum origin-independent [7]. For the crystalline phase, 30 trajectories of 40 ps for a total simulation time of 1 200 ps at 300 K were simulated. In figure 5.3 the average spectra in both crystalline and gas phases are given.

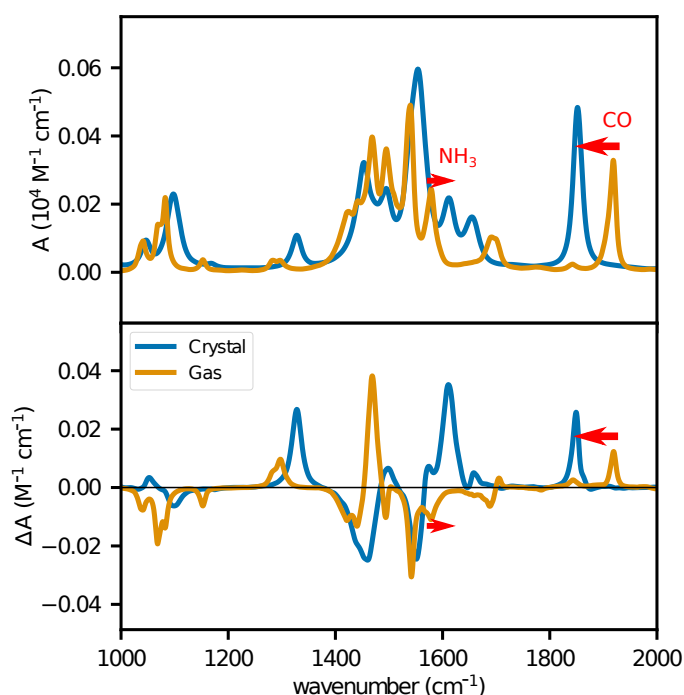


Figure 5.3: IR and VCD spectra averaged over 40 trajectories with a periodic box containing 16 alanine molecules, as obtained from FFMD simulations at 300 K. The VCD spectrum of alanine in the gas phase is multiplied by 3, for visibility. The red arrows highlight the shifts of peaks corresponding to the NH_3 bend and the $\text{C}=\text{O}$ stretch, between the two phases.

The spectra are well resolved and by comparison of the crystal spectra to their gas phase counterparts, the groups of peaks are mainly found around the same positions. However, shifts of the NH_3 bend from 1580 to 1622 cm^{-1} and of the $\text{C}=\text{O}$ stretching mode from 1919 to 1853 cm^{-1} are displayed, represented by the red arrows in figure 5.3. These shifts are due to the electrostatic interactions between the amino and carboxylate groups present in the crystal.

5.1.1.2 Comparison to experimental results

The previous IR and VCD spectra of crystal alanine are compared to experimental data [1] in figure 5.4. The experimental data were obtained from ground crystals as nujol oil mull. These data were only available up to 1800 cm^{-1} .

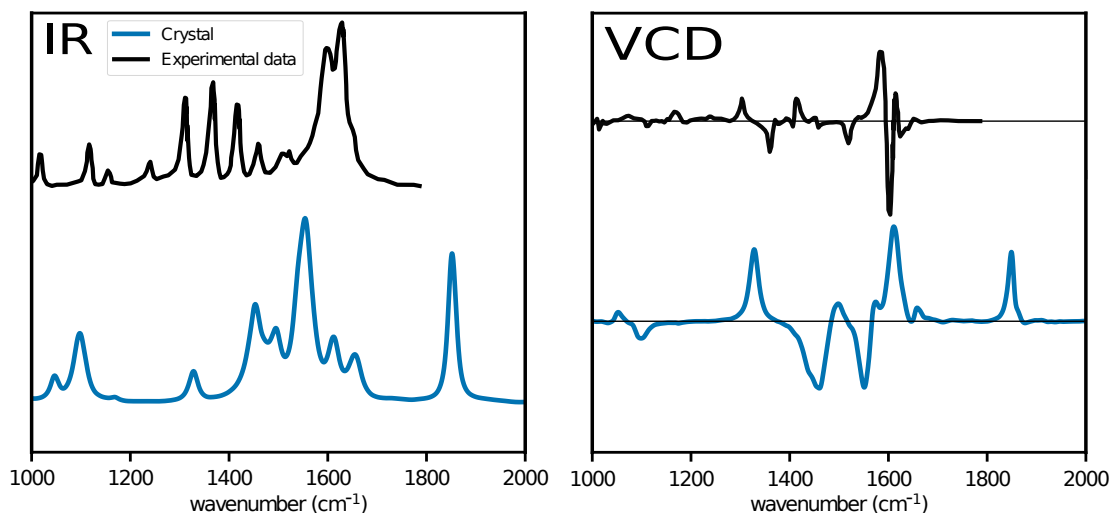


Figure 5.4: (a) IR and (b) VCD spectra of the alanine crystal from experimental, measured as a nujol oil mull [1] and theoretical results from the force field MD approach at 300 K.

In the IR spectra, the peaks match between 1400–1800 cm^{-1} , below this range, the peaks are shifted. For the VCD spectra, a similar pattern of a positive then negative peak is found around 1400 cm^{-1} .

5.1.1.3 Statistical errors and convergence rates

The spectra of the alanine crystal shown above result from the average of 30 trajectories. To determine if the system has been sampled for a sufficient amount of time, the convergence is now examined. To show when the intensities of each peak stabilise at their converged values after a certain sampling time, the averages from an increasing number of trajectories (N_T) are plotted. Starting from the top of the figure 5.5, the spectra arise from the average from 2 trajectories ($N_T = 2$). Then the number of trajectories increases by 3 stepwise. Each spectrum was compared to the reference spectrum obtained from the entire set of 30 trajectories.

The average of the spectra from 30 trajectories was considered converged and was taken as our reference to which other spectra with lesser sampling are compared. Thus the spectra from $1 < N_T < 30$ trajectories were compared to this averaged spectrum to determine their normalised error, with the method and equations shown in chapter 4. This error is shown in figure 5.6 as a function of the sampling ratio $S_r = N_T/30$.

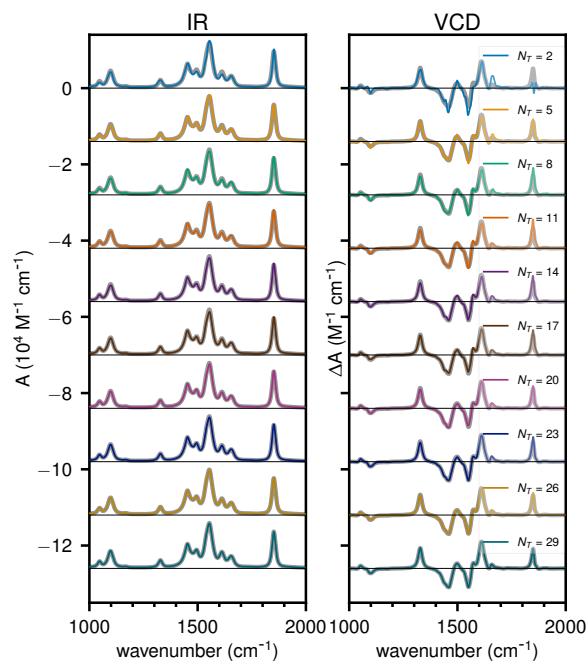


Figure 5.5: IR and VCD spectra of alanine crystal in a periodic box containing 16 molecules, as obtained from FFMD trajectories at 300 K. The averages from 2 to 29 trajectories of 40 ps are displayed. The number of averaged trajectories increases with a step of 3. Represented in grey is the average of 30 trajectories.

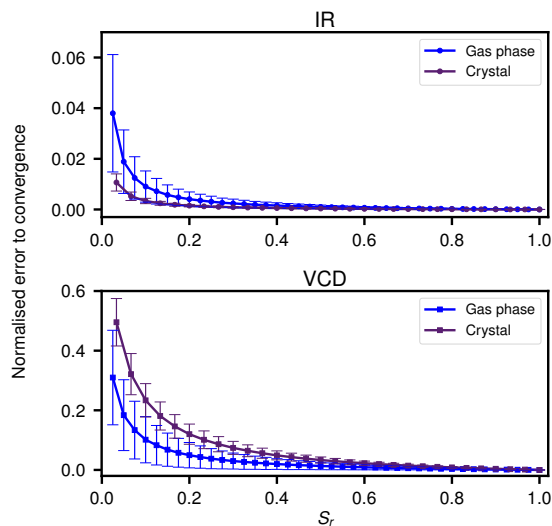


Figure 5.6: Normalised error of the average in the IR and VCD spectra of crystalline alanine from N_T FFMD trajectories of 40 ps at 300 K, compared to the average from 30 trajectories, as a function of the sampling ratio $S_r = N_T/30$. The corresponding data for the gas phase system are superimposed in this figure.

Figure 5.6 shows a slightly faster convergence for IR over VCD. The convergence of the spectrum from the crystal compared to the spectrum from the gas phase is faster for IR but slower for VCD. If only one molecule of alanine is considered for the solid, the convergence of the crystal is faster in both IR and VCD. The supramolecular interactions of the crystal may add instability to the magnetic dipole moment, thus disturbing the convergence.

Furthermore, the crystalline phase shows convergence with a shorter amount of sampling time than the gas phase. In the previous chapter, it was shown that a simulation time of 4 ns was needed for convergence in the gas phase. The crystalline phase can be considered converged at a sampling ratio of 0.5 so after the system has sampled 15 trajectories of 40 ps, thus only 600 ps of sampling time.

5.1.2. Assessing moments and spectra from one trajectory

As in chapter 4, a trajectory (positions and velocities) from classical molecular dynamics based on force fields was exploited to calculate the electric and magnetic dipole moments using DFT within the CPMD software [8]. The DFT dipole moments were computed by our collaborator S. Jähnigen.

The spectra from both methods are shown in figure 5.7. Both IR and VCD spectra from FFMD are very similar to their DFT counterpart. As for zwitterionic alanine in the gas phase, the force field gives rise to more activity in the 1410–1450 cm^{-1} range and a lower intensity at 1925 cm^{-1} in the IR spectrum. The signs of all peaks in the VCD spectrum agree between DFT and FFMD.

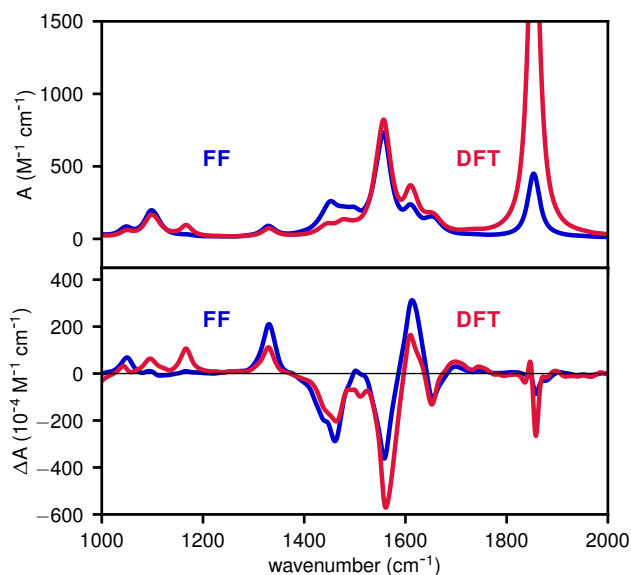


Figure 5.7: Comparison of the IR and VCD spectra of the alanine crystal using the dipole moments calculated with the force field and with DFT from the same FFMD trajectory at 300 K.

For each point along a FFMD trajectory, the electric dipole moment $\vec{\mu}$, its time derivative $\dot{\vec{\mu}}$, and the magnetic dipole moment \vec{m} , were determined from the force field and using DFT. The values of the 3 Cartesian components of the dipole moments obtained at the FF (DFT) level are shown along the $x(y)$ axis. This provides figure 5.8 in which the Pearson correlation coefficient r is also given. The values of r are satisfactory and are generally even improved respect to the values found for the gas phase. This is probably due to the force field describing the crystalline phase with more accuracy than the gas phase. To describe the interactions within the crystal, the force field's treatment of electrostatics via multipoles centered on each atom and the inclusion of polarisation through induced dipole moments prove to be important especially for the good directionality of the hydrogen bond.

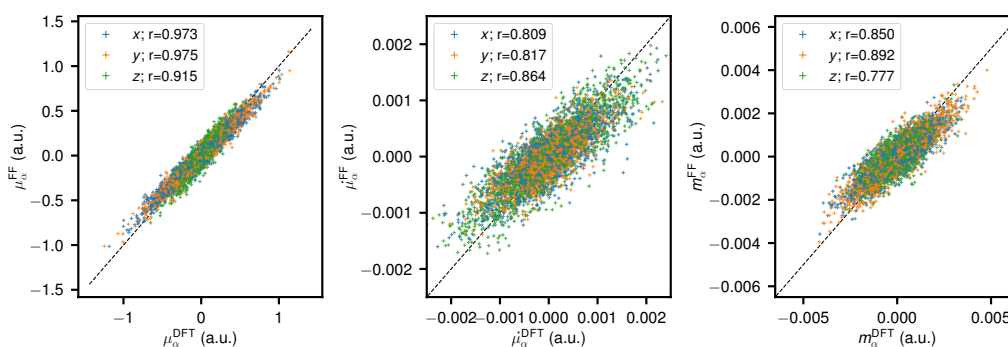


Figure 5.8: Correlation of the projections of the electric dipole moment $\vec{\mu}$ or its time derivative $\dot{\vec{\mu}}$, and the magnetic dipole moment \vec{m} , as obtained for a specific FFMD trajectory at 300 K but calculated from the force field or using density functional theory.

5.2. Hydrated alanine

Amino acids are generally solvated in water thus it is important to examine their behaviour in the hydrated phase. The study of alanine is continued here with a solvent of water. This will permit the analysis of the method using classical molecular dynamics based on force fields to determine VCD spectroscopy in this more realistic phase.

5.2.1. Vibrational spectra obtained from force field molecular dynamics

A molecule of alanine in the zwitterionic form was soaked in a previously equilibrated periodic box of length $L = 18.6 \text{ \AA}$. This procedure gives a periodic box containing 211 water molecules and 1 alanine molecule, see fig. 5.9.

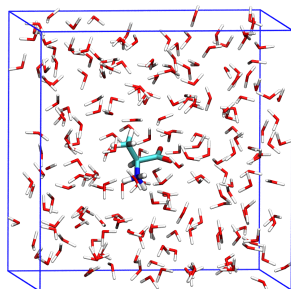


Figure 5.9: Visual depiction of hydrated alanine in its periodic box.

5.2.1.1 Masking the water contribution

Water has a large electric dipole moment, this leads to an IR spectrum showing large intensities ($\sim 20\,000\text{ M}^{-1}\text{cm}^{-1}$) especially compared to the intensities in the gas phase spectrum ($\sim 400\text{ M}^{-1}\text{cm}^{-1}$), as shown in figure 5.10. The IR spectrum is computed from the average of 10 trajectories of 40 ps. Thus, when calculating the IR spectrum of hydrated alanine, the water spectrum hides the contribution of alanine.

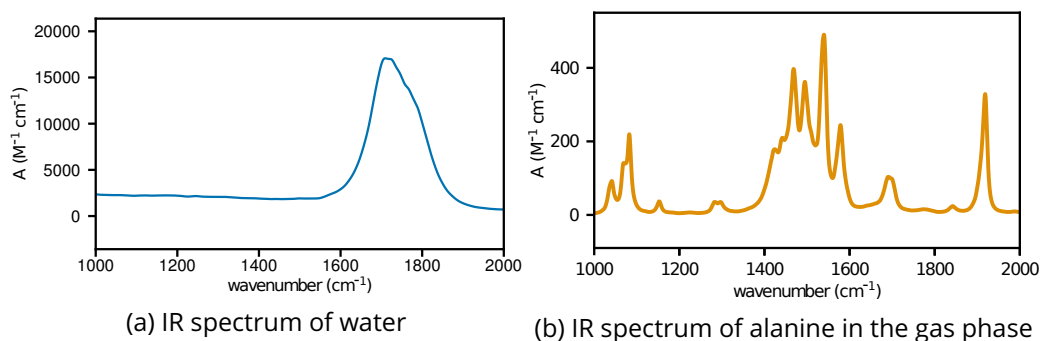


Figure 5.10: Comparison of the intensities of IR spectra of water and of alanine in the gas phase.

Experimentally this is avoided by subtracting the spectrum of pure water from the spectrum of hydrated alanine. Such a heuristic approach would be tricky in simulation because the pure solvent and the hydrated molecule do not share exactly the same periodic box. Thus the spectrum of water alone will not correspond to the spectrum of alanine and water, minus the contribution of alanine.

However if the same molecular dynamics trajectory is exploited and no water molecules are considered when evaluating the dipole moments, it becomes possible to have the local spectrum of alanine influenced by the water molecules. This can be achieved through scaling the dipole moments, as explained in section 2.6.4 of chapter 2 with the following damping function: $P_{KJ}(t) = \left(1 + e^{(|\vec{R}_{KJ}(t)| - R_0)/D}\right)^{-1}$. This scaling employs a cut-off coefficient R_0 . Within the ChirPy program, for a cutoff of 0 \AA , only the solute is considered.

For the VCD spectrum of hydrated alanine, the contribution to the spectrum from the water molecules should be zero as this solvent is achiral. However, the possible effects of chirality transfer in the vicinity of the solute should not be excluded. In this case, the solute could transmit chiral properties, including a VCD signal, to the solvent.

5.2.1.2 Choosing a cutoff around the alanine molecule

The electric and magnetic dipole moments were computed within the periodic box of hydrated alanine. As before, the Fourier transform of the appropriate time correlation functions was performed for each trajectory, then the resulting spectra were averaged. Here 30 trajectories of 40 ps at 300 K were exploited. Again, the nearest periodic image is directly chosen between particles to make the VCD spectrum origin independent. A cutoff of $R_c = 0.0 \text{ \AA}$ includes only the alanine molecule, a cutoff of $R_c = 3.85 \text{ \AA}$ further includes the first hydration shell. The spectra using various values of the cutoff, but all exploiting the very same MD trajectories, are shown in figure 5.11.

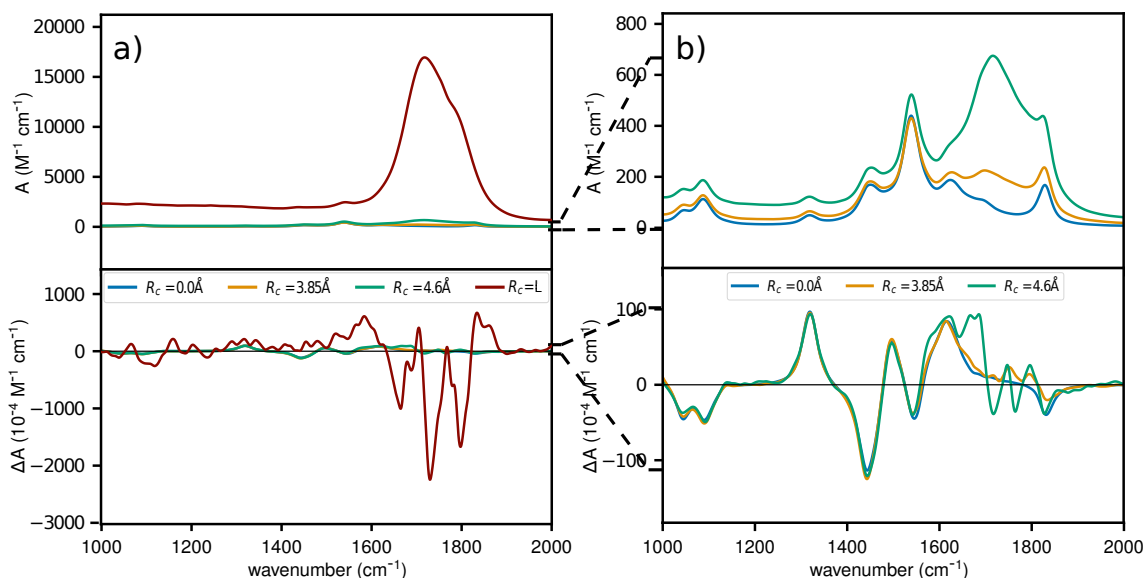


Figure 5.11: IR and VCD spectra of hydrated alanine averaged over 30 trajectories with cutoff values at $R_c = 0.0 \text{ \AA}$, 3.85 \AA and 4.6 \AA , as obtained from FFMD trajectories at 300 K. Figure a) has been enlarged in to produce figure b).

It is shown here that if no cutoff is taken, thus R_c equals the length of the periodic box L , then the intensity of the water peak takes over the whole spectrum (Fig. 5.11a). Depending on the cutoff taken and thus on the number of water molecules included, the influence of water on the alanine spectrum can be varied. This is particularly clear between 1600 and 1900 cm^{-1} in figure 5.11b).

This is the case also for the VCD spectra, though the VCD signal of an achiral molecule is zero, thus there should be no water signal. One reason for this can be that sampling of $30 \times 40 = 1200 \text{ ps}$ is still not long enough for the contributions of the water molecules

to average out to zero. The second reason can be the presence of chirality transfer from the alanine to the solvent, at least in the vicinity of the solute.

In the following, the comparison between the experimental and theoretical spectra is examined. The experimental spectrum [1] has been background subtracted with a water spectrum and the theoretical results correspond to the FFMD spectrum depicted in figure 5.11 with a cutoff value of $R_c = 0.0 \text{ \AA}$, as the water is also subtracted in the experimental spectrum. These two results are compared in figure 5.12 within frequency range of 1000 to 1800 cm^{-1} .

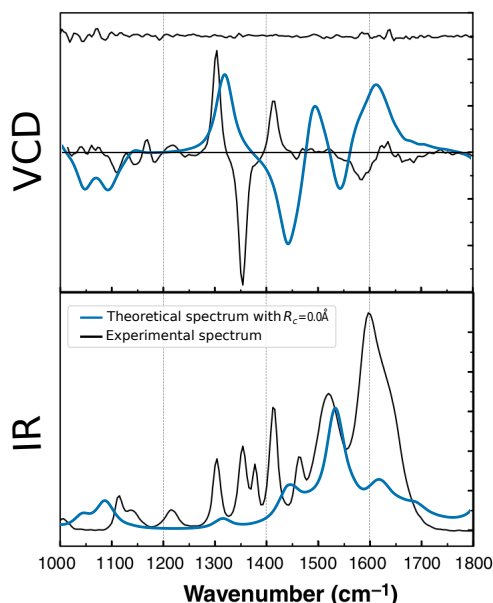


Figure 5.12: IR and VCD spectra of hydrated alanine from experimental [1] and theoretical (FFMD) results at 300 K. The experimental results are background subtracted with a water spectrum and a cutoff value of $R_c = 0.0 \text{ \AA}$ is used for the theoretical spectra.

The IR spectra are quite similar especially above 1400 cm^{-1} . The low resolution of the theoretical spectrum is probably due to the smoothing performed during the Fourier transform of the correlation functions. This smoothing parameter will need to be adjusted in future computations. The main discrepancy is at 1050 cm^{-1} . The intense double peak in the experimental VCD spectrum is also represented in the theoretical results, similarly to the spectra of the alanine crystal. However there is a slight shift for the negative peak.

5.2.2. Statistical errors and convergence rates

The spectra of hydrated alanine shown above are the result from the average of 30 trajectories. The averages of an increasing number of trajectories (N_T) are plotted to examine the convergence and are compared to the full average in figure 5.13. Between successive spectra three new trajectories are added.

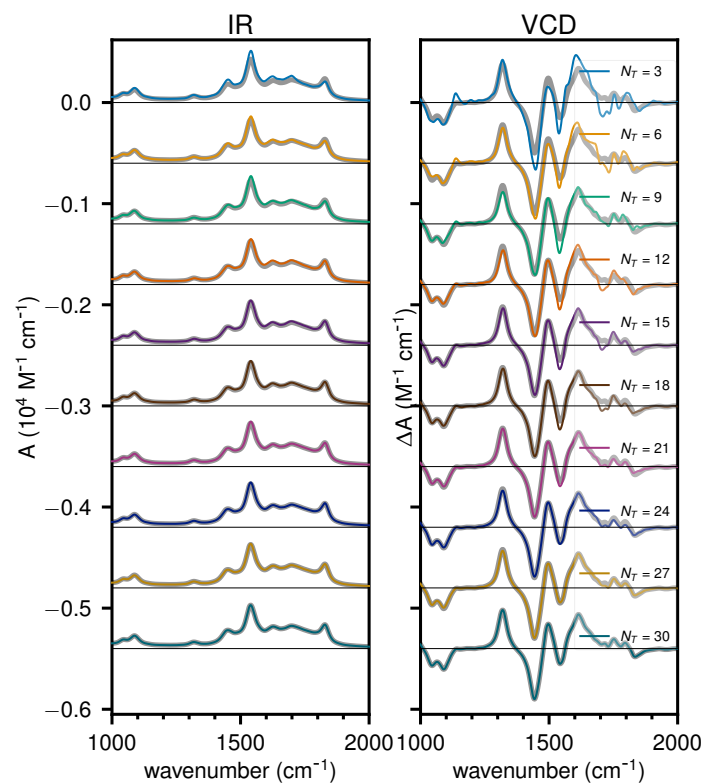


Figure 5.13: IR and VCD spectra of hydrated alanine from the average of 2 to 30 trajectories with cutoff of 3.85 Å, as obtained from FFMD trajectories at 300 K.

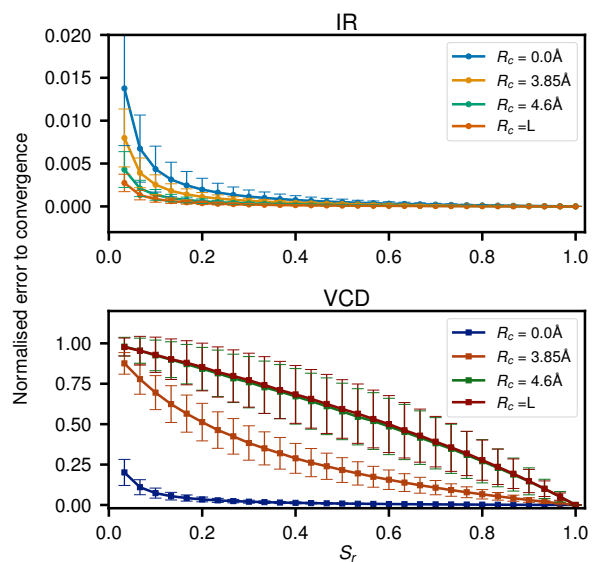


Figure 5.14: Normalised error of the average in the IR and VCD spectra of hydrated alanine from N_T FFMD trajectories of 40 ps at 300 K, compared to the average from 30 trajectories, as a function of the sampling ratio $S_r = N_T/30$.

The normalised errors were determined as a function of the sampling ratio $S_r = N_T/30$ and they are shown in figure 5.14. The error was computed by comparing the spectra from the average of $1 < N_T < 30$ trajectories to the converged spectrum. The normalised errors were determined from the whole frequency range, from 0 to $4\,000\text{ cm}^{-1}$. The high fluctuation of the OH stretching band at 3500 cm^{-1} is thus also taken into account. This explains why the normalised error (figure 5.14) shows a slower convergence than the spectra from the averages of an increasing number of trajectories (figure 5.13) that only show the 1000 to 2000 cm^{-1} region.

The convergence of the VCD spectrum of hydrated alanine is very much more arduous for IR and also for the other phases, especially when water molecules are accounted for. This large error arises from the large fluctuation of the peaks corresponding to the water molecules, which is why the normalised error with a vanishing cutoff is considerably lower than the error with finite cutoffs. This large error from the signal of water molecules confirms the hypothesis that the system has not been sampled enough for the water contributions to average out to zero. This could explain why a residual signal for water is present in the FFMD spectrum. In the future, spectra from a longer sampling time need to be computed.

5.2.3. Assessing dipole moments and spectra from one trajectory

Following the same lines as in the previous chapter, the electric and magnetic dipole moments are determined from the same MD trajectory with DFT and with the force field methods.

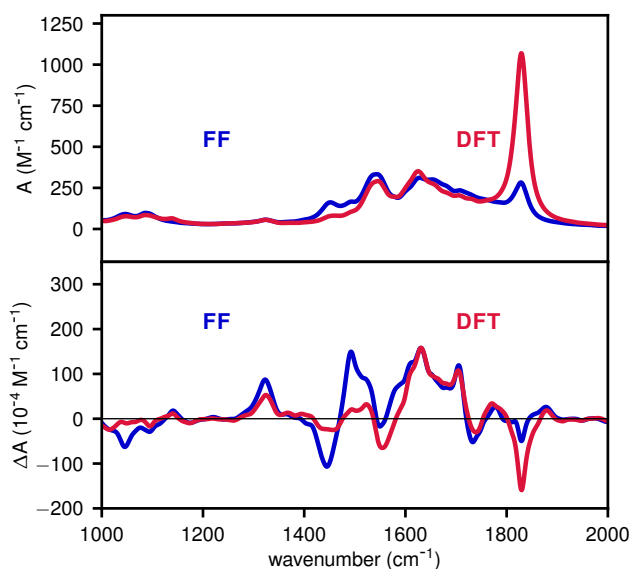


Figure 5.15: Comparison of the IR and VCD spectra of hydrated alanine using the dipole moments calculated with the force field and with DFT from the same FFMD trajectory at 300 K , processed with a cutoff of $R_c = 3.85\text{ \AA}$.

The spectra from both FF and DFT methods are shown in figure 5.15. A cutoff value

of $R_c = 3.85 \text{ \AA}$ was chosen to consider the alanine and its first solvation shell only. The spectra are rather similar and the peaks of both VCD spectra have the same signs. Again, as for zwitterionic alanine in the gas phase, the force field gives rise to more activity in the $1410\text{--}1450 \text{ cm}^{-1}$ range and a lower intensity at 1925 cm^{-1} in the IR spectrum.

For each point along a FFMD trajectory the electric dipole moment $\vec{\mu}$, its time derivative $\dot{\vec{\mu}}$, and the magnetic dipole moment \vec{m} , were determined from the force field and using DFT. The values of the moments from the force field are plotted along the y -axis and the values of the DFT moments are plotted along the x -axis. This yields figures 5.16 to 5.18, where r is the associated Pearson correlation coefficient.

Here three figures comparing the force field and DFT moments are shown. The first figure (5.16) represents the moments of the whole system of alanine and the 211 water molecules. The second figure (5.17) displays only the moments of the alanine, ignoring the moments of the water molecules. The third figure (5.18), depicts the opposite, showing only the moments from the 211 water molecules.

The Pearson correlation coefficients for the dipole moments of the water molecules in figure 5.18 are very high, above 0.96. This possibly comes from water being a system that has been studied deeply and for which the force field parameters are very accurate, in particular the multipolar distribution. The values of r from the dipole moments of only alanine, shown in figure 5.17, are slightly lower. The values of r from the dipole moments of water are reflected in those of alanine and water, in figure 5.16, as there are considerably more molecules of water than of alanine. In all three figures the values of r are very high.

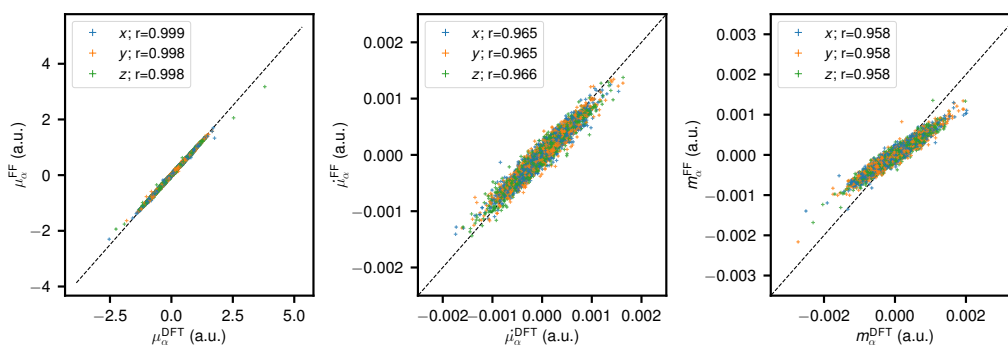


Figure 5.16: Correlation of the electric dipole moments $\vec{\mu}$, the velocity form of the electric dipole moments $\dot{\vec{\mu}}$ and the magnetic dipole moments \vec{m} of alanine and water.

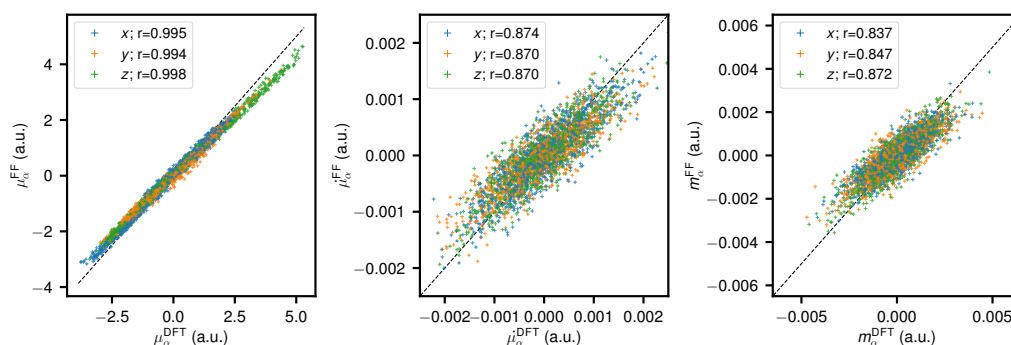


Figure 5.17: Correlation of the electric dipole moments $\vec{\mu}$, the velocity form of the electric dipole moments $\dot{\vec{\mu}}$ and the magnetic dipole moments \vec{m} of the alanine molecule, ignoring the dipole moments from the water molecules.

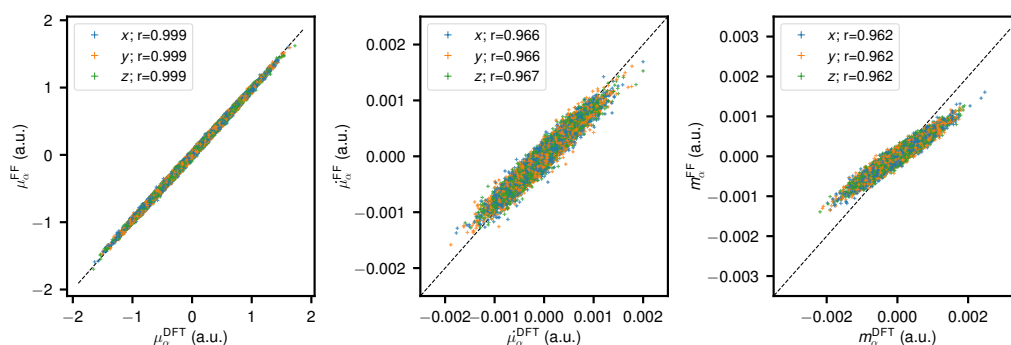


Figure 5.18: Correlation of the electric dipole moments $\vec{\mu}$, the velocity form of the electric dipole moments $\dot{\vec{\mu}}$ and the magnetic dipole moments \vec{m} of the water molecules, ignoring the dipole moments from the alanine molecule.

5.3. Effective mode analyses for crystalline and hydrated alanine

The moments acquired through classical molecular dynamics based on the force field have thus been validated for crystalline and hydrated alanine. Here an effective modes analysis is performed to determine the vibrational modes causing the peaks. Two FFMD trajectories of 40 ps at 300 K were analysed to retrieve the positions and velocities. A first trajectory of the alanine crystal was exploited, then a trajectory of hydrated alanine. While the trajectories were performed on the alanine molecule with its full environment, the power spectra and the effective modes were determined by retrieving the molecular data from a unique alanine molecule. Therefore, the IR spectra and the trajectories themselves fully account for the effect of the environment, similarly to the fragment analysis reported at the end of the previous chapter. These positions and velocities were transformed into internal coordinates to initiate the effective modes analysis. The internal coordinates are detailed in table 5.1.

bond lengths	valence angles	dihedral angles
N1-C2	C3-C2-N1	O4-C3-C2-N1
C2-C3	O4-C3-C2	H5-N1-C2-C3
C3-O4	C2-N1-H5	H8-C2-N1-H7
N1-H5	C2-N1-H6	C9-C2-N1-H7
N1-H6	C2-N1-H7	H10-C9-C2-N1
N1-H7	N1-C2-H8	
C2-H8	C9-C2-N1	
C2-C9	H10-C9-C2	
C9-H10	H11-C9-C2	
C9-H11	H12-C9-C2	
C9-H12	O13-C3-C2	
C3-O13	H5-N1-H6	
	H7-N1-H6	
	H12-C9-H11	
	H10-C9-H11	
	O13-C3-O4	

Table 5.1: List of the 33 internal coordinates used to initialise the effective modes analysis of alanine in the zwitterionic form (12 bond lengths, 16 valence angles, 5 dihedral angles), as relevant for the crystalline and hydrated phases. The labels of the atoms refer to figure 4.1(b).

A power spectrum with an arbitrary intensity and a vectorial depiction of each effective mode were acquired, they are compared to the FFMD IR spectrum in figure 5.19 and 5.20. The same set of internal coordinates is used as for the zwitterionic form of alanine in the gas phase.

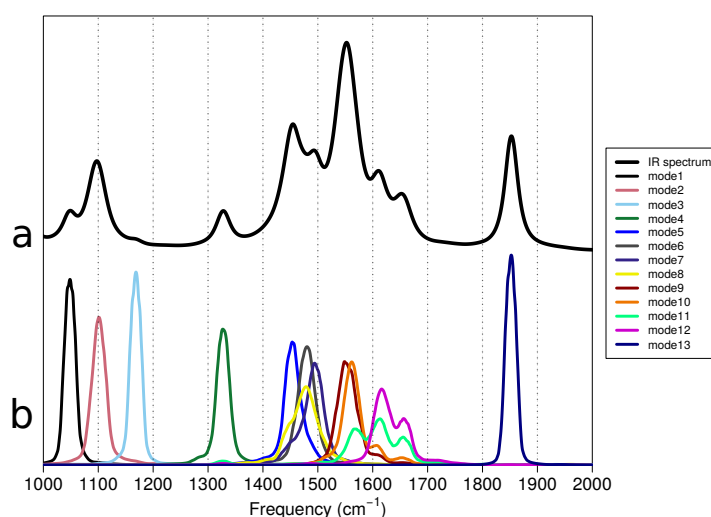


Figure 5.19: (a) IR spectrum as obtained from FFMD trajectories at 300 K of the alanine crystal in the zwitterionic form and (b) its decomposition in effective modes, whose power spectra are shown in colours. The power spectra have arbitrary intensities.

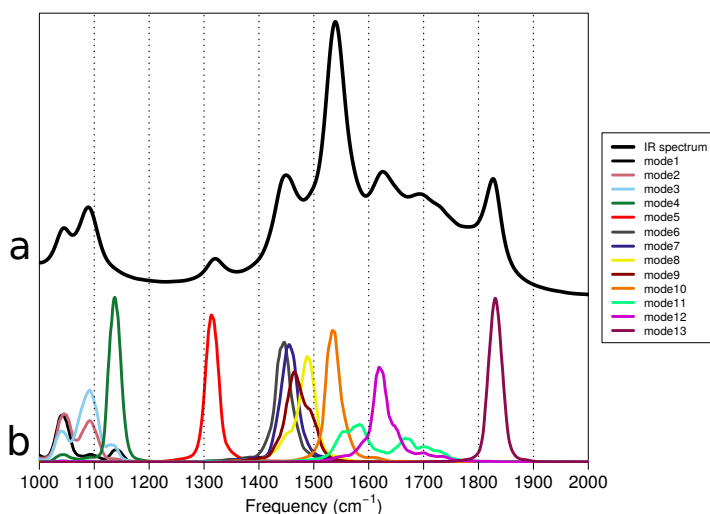


Figure 5.20: (a) IR spectrum as obtained from FFMD trajectories at 300 K of hydrated alanine in the zwitterionic form and (b) its decomposition in effective modes, whose power spectra are shown in colours. The power spectra have arbitrary intensities.

The goal of the analysis reported below is to comment on those similarities and differences based on the effective modes analysis. The main differences between the absorption spectra of hydrated and crystalline alanine are the following:

1. in the region $1000\text{--}1200\text{ cm}^{-1}$, the two main bands are well-resolved in the crystal form, as well as the shoulder at around 1180 cm^{-1} , with the identification of three distinct effective modes bands, whereas, in the hydrated form, the two main bands are predicted to arise from the superposition of three modes, while a fourth mode is responsible for the shoulder, which seems, however, suppressed in water;
2. in the region $1400\text{--}1650\text{ cm}^{-1}$, the lower intensity bands at the sides on the central most intense bands are both split into two contributions in the crystal form, while only one band appears to be present on each side of the intense band when alanine in solvated in water;
3. in the region around 1700 cm^{-1} , a large and not well-resolved band appears in the hydrated form of alanine, which is completely absent in the crystal form.

Additionally, if compared to the zwitterionic form of alanine in the gas phase, the band at around 1300 cm^{-1} is slightly blue-shifted in water and in the crystal, while, at the same time, the band at 1900 cm^{-1} is red-shifted when the molecule is not isolated.

For the decomposition of the hydrated spectrum, there is no effective mode for the peak at $1700\text{--}1800\text{ cm}^{-1}$ as this corresponds to the water asymmetric stretch, not accounted for in the analysis.

In order to analyse the points just discussed, in figure 5.21 some of the effective modes giving rise to the IR absorption bands in the region $1000\text{--}1200\text{ cm}^{-1}$ are reported.

The two modes in the upper panels refer to the hydrated form of zwitterionic alanine (h.), while the modes in the lower panels refer to its crystal form (c.). Note that all modes are very similar to each other, involving some kinds of wagging and twisting of the CH₃ and NH₃ fragments, mainly coupled to the vibrations of the CH bond that is not involved in CH₃. Only mode 3 in the crystal form appears slightly different, as the umbrella mode of NH₃ can clearly be identified.

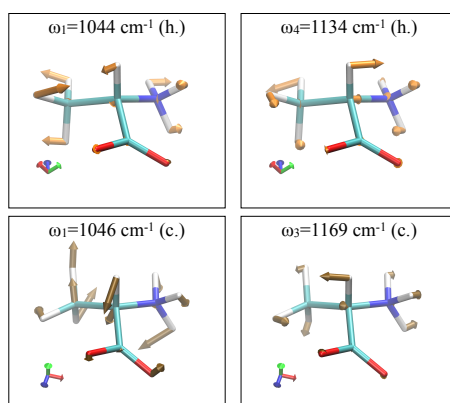


Figure 5.21: Effective modes in the region $1000\text{--}1200 \text{ cm}^{-1}$ for the hydrated (h.) form of zwitterionic alanine (upper panels) and for the crystal (c.) form of zwitterionic alanine (lower panels).

Concerning the appearance of the double peaks in the region $1400\text{--}1650 \text{ cm}^{-1}$ of the IR spectrum of crystal alanine in comparison to the single peaks structure of hydrated alanine, the corresponding effective modes are reported in figure 5.22. In the figure, two modes were selected, one around 1450 cm^{-1} and one slightly above 1600 cm^{-1} .

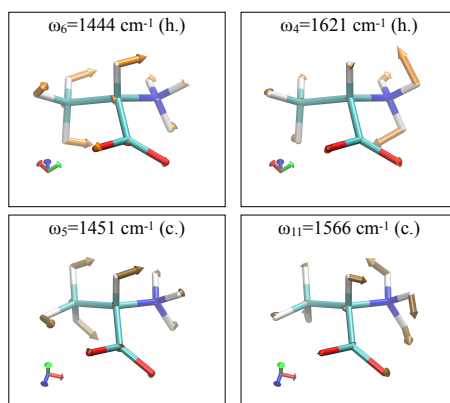


Figure 5.22: Effective modes in the region $1400\text{--}1650 \text{ cm}^{-1}$ for the hydrated (h.) form of zwitterionic alanine (upper panels) and for the crystal (c.) form of zwitterionic alanine (lower panels).

The modes on the left panels and on the right panels of figure 5.22 are very similar: on

the left, both modes involve the umbrella motion of NH_3 coupled to the (out-of-phase) umbrella motion in CH_3 ; on the right, both modes involve the HNH bending motion in NH_3 , coupled to an out-of-plane vibration of the third NH bond in the fragment. Therefore, it appears reasonable to assume that despite the fact the IR spectrum of the crystal form of alanine presents a better resolved structure than its hydrated form, the source of the absorption bands is to be found in modes that are fundamentally very similar. Presumably, the constrained vibrations in the crystal reduce the effect of anharmonicity, thus allowing the absorption bands to be better resolved in frequency.

Finally, in figure 5.23, the effective modes responsible for the appearance of the absorption bands at around 1350 cm^{-1} (left panels) and 1850 cm^{-1} (right panels) are analysed. The effective modes at lower frequency appear to be slightly different. As in water, an umbrella motion involving the fragment CH_3 cannot be found, but rather a combination to HCH bending and CH vibration. As previously, this slight difference can be ascribed to the different rigidities of the environment in which alanine is embedded.

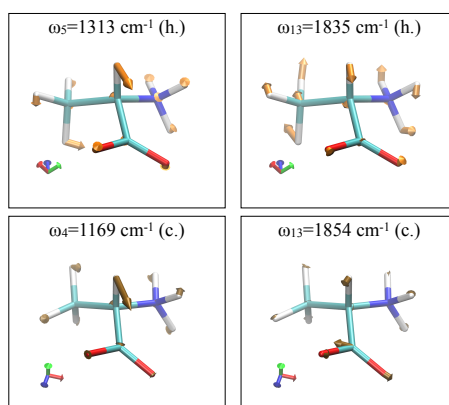


Figure 5.23: Effective modes at around 1350 cm^{-1} (left) and 1850 cm^{-1} (right) for the hydrated (h.) form of zwitterionic alanine (upper panels) and for the crystal (c.) form of zwitterionic alanine (lower panels).

The effective mode at higher frequency allows to very clearly identify the asymmetric stretching motion of COO in the crystal, which becomes instead a wagging motion in water and which is coupled to vibrations in the other fragments of the molecule. It can thus be concluded that in water in the region spectral region $1000\text{--}2000\text{ cm}^{-1}$ the pure antisymmetric stretching motion of COO appears to be suppressed.

5.4. Alanine solvated in dinitrogen

A second solvent for alanine was studied, dinitrogen. The idea of this study is to examine a solvent that does not involve hydrogen bonding with the solute and for which the results are expected to be closer to those in the gas phase, except for temperature effects. At 300 K, dinitrogen is naturally found in the vapor phase, and its influence on the vibrational spectra of alanine would be extremely limited. To impose a dense liquid,

the simulations were performed at 77.35 K. This time a molecule of *neutral* alanine was solvated in a periodic box of 18.27 Å. The density of the box is 0.8171 g/cm³.

The low temperature prevents an extensive exploration of the potential energy surface during the MD simulation, as the alanine molecule stays mostly vibrating around its starting structure. Thus the lowest energy structures of families 2 and 3 of alanine found in the exploration in chapter 4 were used as starting points for FFMD simulations in the gas and solvated phases. Family 1 was excluded as it is not found in the force field simulations.

For each family, 10 different simulations were performed. The first simulation used the structure of its corresponding family and the other nine structures used the last structure of the previous simulation. Each simulation was composed of an initial 5 ps MD simulation in the canonical ensemble to equilibrate the system at 77.35 K. Then a 200 ps MD simulation, for the gas phase and a 40 ps MD simulation for the solvated phase, in the microcanonical ensemble were carried out to retrieve the electric and magnetic dipole moments. Thus for each family a spectrum from the average of 10 spectra was obtained.

5.4.1. Vibrational spectra obtained from force field molecular dynamics

To achieve an averaged spectrum with sufficient sampling at low temperature, thus including both families, the thermal probability $Z_\alpha / \sum_i Z_i$, with Z corresponding to equation (4.12), was determined for each configuration of both families at 77.35 K, using the energies and harmonic frequencies from the force field. The value for each configuration was summed within each family so as to obtain values of the thermal probability for both families. The spectra from the average of 10 trajectories was then multiplied by the corresponding thermal probability to give the spectra at 77.35 K in figure 5.24.

First, the spectrum from the gas phase is compared from the FFMD simulations at 300 K and 77.35 K, then the spectrum from the gas phase at 77.35 K is compared to the spectrum of alanine solvated in N₂ at 77.35 K.

In figure 5.24a, the bands are present at similar locations for both temperatures. As expected, due to the higher temperature, the peaks for the alanine in the gas phase at 300 K, are significantly wider.

In figure 5.24b, at 1250 and 1700 cm⁻¹, there is a dramatic change between the VCD spectrum of alanine in the gas phase and the VCD spectrum of alanine solvated in N₂ at 77.35 K. The IR spectra do not show significant change, only a slight widening of the peaks. This can be explained by the change of configuration of the structure of family 2 when solvated in N₂ compared to the structure shown in chapter 4. The NH₂ group rotates by 180° and the hydrogen of the hydroxyl group turns towards the amino group. This change does not occur when the structure of family 2 is in the gas phase at 77.35 K. The structure of family 3 remains identical to the structure shown in chapter 4 in the gas phase and solvated in N₂ at 77.35 K. This confirms the sensitivity of the VCD spectrum to conformational changes, especially compared to the IR spectra where the change is minimal.

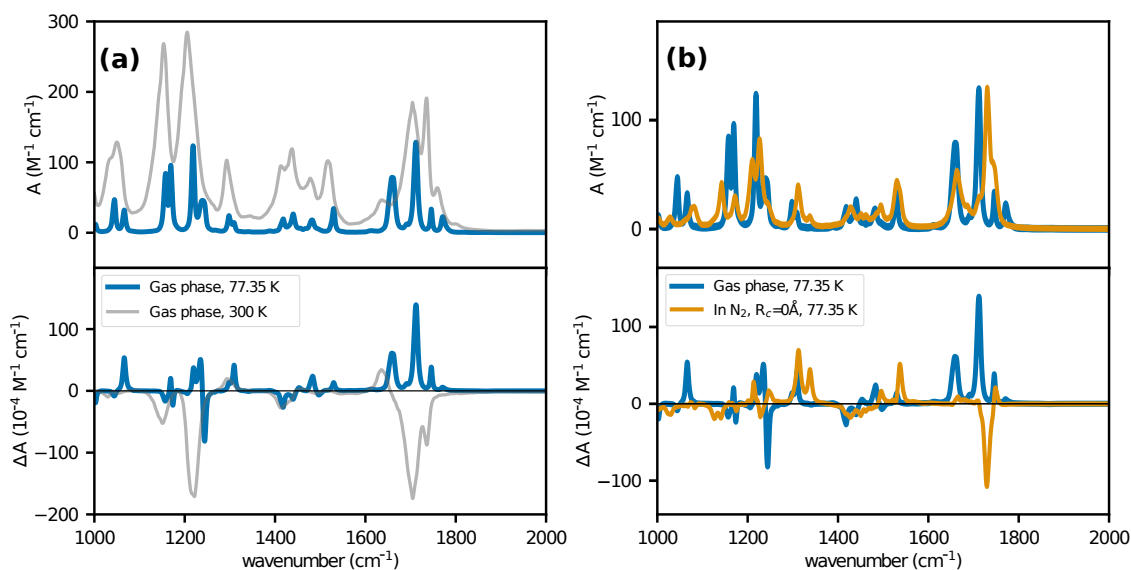


Figure 5.24: (a) IR and VCD spectra of alanine in the neutral form in the gas phase at 77.35 K and 300 K. (b) IR and VCD spectra of alanine in the neutral form in the gas phase and in N_2 at 77.35 K.

5.5. Phenylcyclohexanediol in the crystalline phase and solvated in DMSO

A final system was studied, phenylcyclohexanediol. This molecule has a diol group which contains two hydroxyl groups. These diols can form hydrogen bonds, which is of particular interest. Combining with theoretical results, more information can be extracted from the experimental data. This molecule was examined in the crystalline phase and solvated in DMSO (see figure 5.25). The unit cell of the phenylcyclohexanediol crystal is composed of 8 molecules that form dimers. Each dimer has 2 hydrogen bonds between the 4 hydroxyl groups, represented by the black dotted lines. This unit cell is rather large ($19.767 \times 5.645 \times 19.924 \text{ \AA}$) making FFMD particularly valuable to address such a complex system.

The periodic box, for the crystalline phase, has the same size as the unit cell defined from crystallographic data shown in figure 5.25a.

For the solvated phase, a cubic box of 18.0 \AA was chosen. Phenylcyclohexanediol was solvated in 44 DMSO molecules and the density of the system was 1.0336 g/cm^3 ; it is depicted in 5.25b. Both phases were studied at 150 K, exploiting 20 trajectories of 80 ps to collect the electric and magnetic dipole moments.

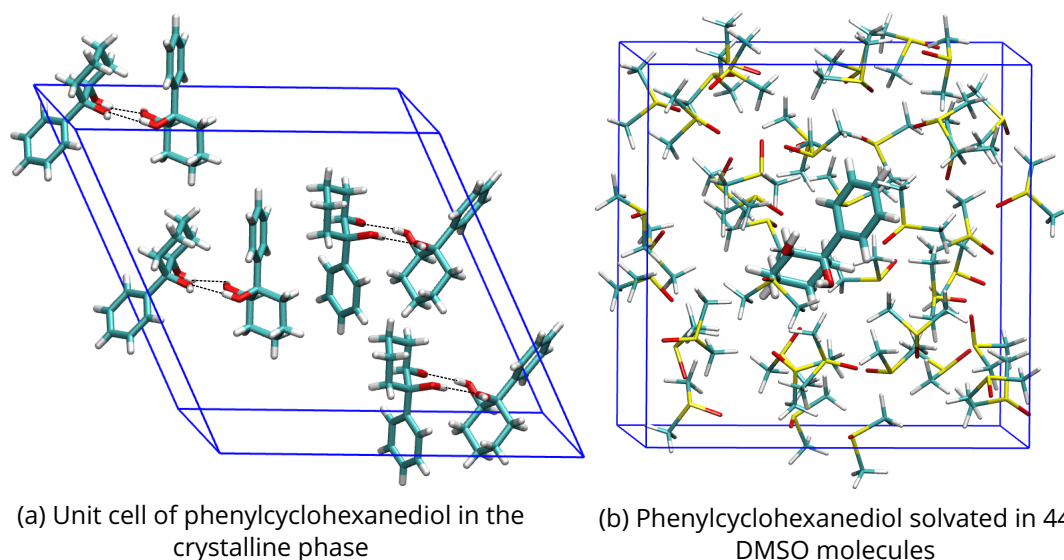


Figure 5.25: Visual depiction of phenylcyclohexanediol in the crystalline phase and solvated in DMSO.

5.5.1. Vibrational spectra obtained from force field molecular dynamics

The main difference between solvated and crystalline phenylcyclohexanediol is the interaction of its hydroxyl groups. When it is solvated in DMSO, its hydroxyl groups interact with DMSO though in the crystalline phase they interact with one another. This major difference influences the IR and VCD spectra of figure 5.26, where the crystal spectra are compared to their solvated phase counterparts.

For solvated phenylcyclohexanediol, as a first step, a vanishing cutoff ($R_0=0$) was chosen, thus leading to ignore the DMSO contribution.

The interest of this system is the OH groups, thus the spectra is shown from 2800 to 4000 cm^{-1} where the bands of the OH stretch are present. The peaks at 3000 cm^{-1} correspond to the CH stretches, which are very similar for both phases in the IR spectra. In the VCD spectra, then are present at the same frequency but with a change of sign which is probably due to the change of environment of the carbon rings.

The peaks in between 3600 and 3800 cm^{-1} correspond to the OH stretching modes. The difference in both the IR and VCD spectra is very noticeable and expected due to the very different environments and configurations of the OH groups. The large peak at 3700 cm^{-1} in the spectrum of phenylcyclohexanediol solvated in DMSO could be composed of multiple peaks. Adjustment of the smoothing factor needs to be tested in the future. In addition, as no solvent is considered in the IR and VCD spectra of phenylcyclohexanediol solvated in DMSO, future work could consider the effect of the DMSO solvent.

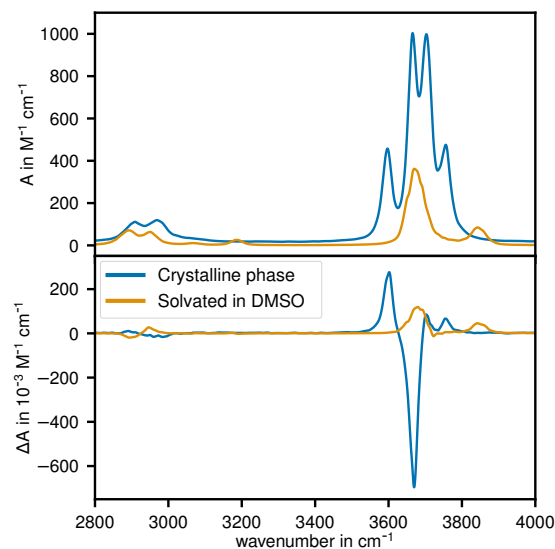
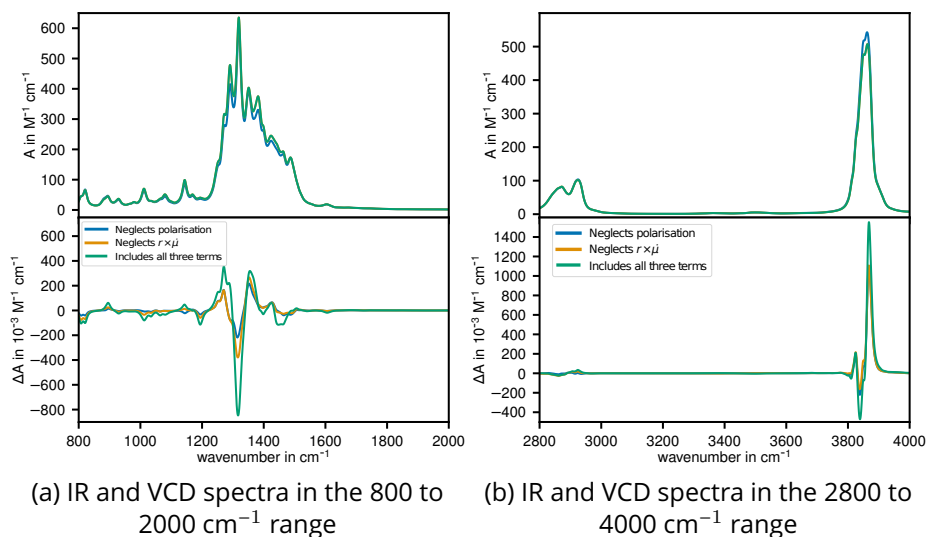


Figure 5.26: IR and VCD spectra averaged over 20 FFMD trajectories of 80 ps at 300 K of crystal phenylcyclohexanediol and solvated in DMSO. For visibility the spectra of the crystalline phase were multiplied by 0.1 for the IR spectrum and 0.02 for the VCD spectrum.

In the following, the spectra of isolated phenylcyclohexanediol are shown. In this system there is no inter-molecular hydrogen bonding as there is only one molecule. Figure 5.27a, shows the influence of the various terms of the magnetic dipole moment in the 800 to 2000 cm^{-1} range and figure 5.27b in the 2800 to 4000 cm^{-1} range.



(a) IR and VCD spectra in the 800 to 2000 cm^{-1} range

(b) IR and VCD spectra in the 2800 to 4000 cm^{-1} range

Figure 5.27: IR and VCD spectra of isolated phenylcyclohexanediol averaged over 20 FFMD trajectories at $T = 300$ K. Both figures emphasize the influence of the polarisation of the molecule and of the time varying induced dipole moments moving around fixed positions term added to the magnetic dipole moment.

In figure 5.27a, the inclusion of all the terms containing the induced dipole moment influences the VCD spectrum the most between 1200 and 1400 cm^{-1} though without any peak inversion, contrary to the alanine spectrum discussed in chapter 4. In figure 5.27b, the influence of the terms containing the induced dipole moment is minimal, we can deduce that the induced dipole moments are small on the OH groups, compensation effects being quite limited as well.

Comparing these spectra to the crystalline and solvated phases, the OH peaks are found to be narrower and shifted to higher frequencies. The peaks corresponding to the CH stretches are slightly shifted to lower frequencies. Both of these shifts are expected due to the change of environment.

5.5.2. Statistical errors and convergence rates

The convergence is examined by comparing the spectrum from the average of 20 trajectories of 80 ps to the average of a lesser number of trajectories. First, the convergence is observed directly through the spectrum, in figure 5.28, then through the normalised error, in figure 5.29.

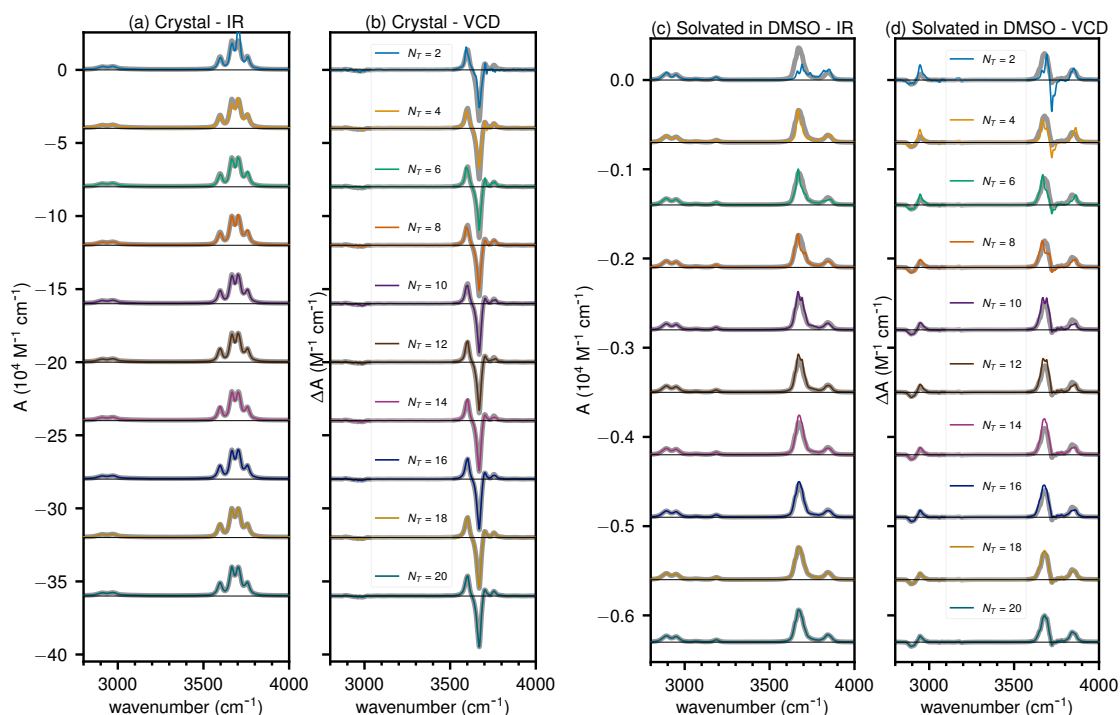


Figure 5.28: IR and VCD spectra averaged of phenylcyclohexanediol in crystalline form and solvated in DMSO. The averages from 2 to 20 FFMD trajectories of 80 ps at 300 K are displayed. The number of averaged trajectories increases with a step of 2. Represented in grey is the average of 20 trajectories.

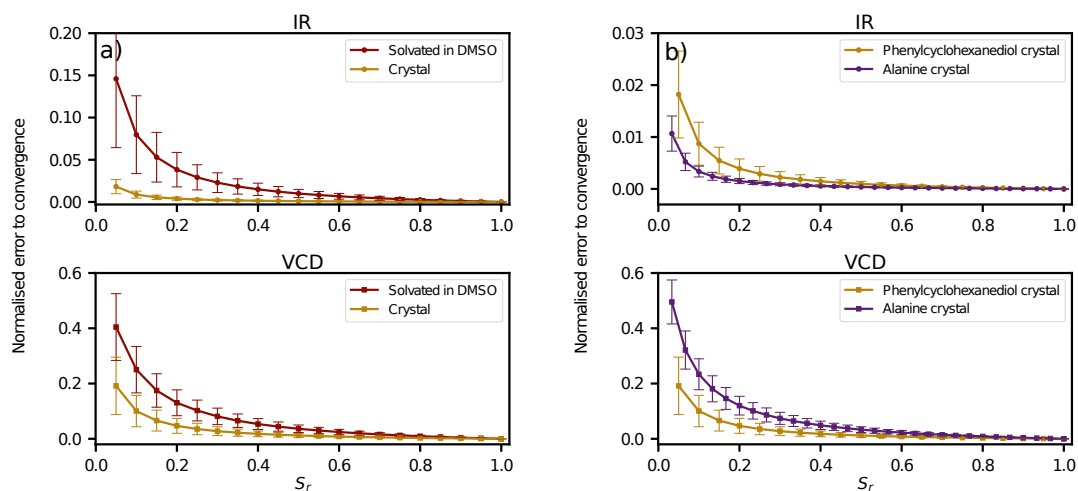


Figure 5.29: a) Normalised error of the average in the IR and VCD spectra of phenylcyclohexanediol in the crystalline phase and solvated in DMSO from N_T FFMD trajectories at 300 K of 80 ps, compared to the average from 20 trajectories, as a function of the sampling ratio $S_r = N_T/20$; b) Normalised error of the average in the IR and VCD spectra of phenylcyclohexanediol in the crystalline phase from N_T FFMD trajectories at 300 K of 80 ps, compared to the average from 20 trajectories, as a function of the sampling ratio $S_r = N_T/20$, and of the alanine crystal from N_T trajectories of 40 ps, compared to the average from 30 trajectories, as a function of the sampling ratio $S_r = N_T/30$.

For the normalised error, in figure 5.29a, the two phases of phenylcyclohexanediol are compared, then, in figure 5.29b, the errors of phenylcyclohexanediol crystal are shown against those of the alanine crystal. The convergence is faster for the crystalline phase than for the solvated phase. This was also the case for crystalline alanine (see figures 5.6 and 5.14) and reflects the improved convergence that results from the presence of multiple molecules with strong interactions within the crystal, making the molecules less fluxional. When comparing the two crystals, again the alanine crystal shows lower IR errors, this time against phenylcyclohexanediol, but higher VCD errors.

The sampling time necessary for the spectra of the phenylcyclohexanediol crystal and of phenylcyclohexanediol solvated in DMSO to converge is 10 trajectories of 80 ps. This time is longer than for the alanine crystal and the hydrated alanine in the case of vanishing cutoff. This is due to the high flexibility of phenylcyclohexanediol resulting in a large number of configurations that need to be all explored.

The dipole moments of the phenylcyclohexanediol systems remain to be compared to DFT dipole moments to confirm their accuracy. An effective modes analysis will also be performed to examine, with more detail, in the future, which band is assigned to each molecular vibration.

5.6. Concluding remarks

Throughout this chapter, five systems in condensed phases were studied: alanine crystal, hydrated alanine, alanine in dinitrogen, phenylcyclohexanediol in DMSO and phenylcyclohexanediol crystal. For these systems, the IR and VCD spectra were determined with classical molecular dynamics based on force fields. For alanine and phenylcyclohexanediol, the statistical errors and convergence rates were determined and they reveal the extended sampling that is needed, though the crystal shows the shortest sampling time required to reach convergence. This further supports the use of classical force fields to simulate VCD spectra for such systems which are barely accessible with methods that explicitly account for electronic structure. The moments were also determined from one trajectory, using DFT and the force field methods. The comparison of these results showed the excellent agreement of the FFMD results with DFT. This extends our validation of polarisable force fields for VCD spectroscopy from the gas phase to a variety of condensed phases that are more relevant in experiments.

The spectra of alanine in the hydrated and crystalline phase were finally decomposed into the contributions of effective modes. This analysis can be used in the future to assess the peak positions and adjust the force field parameters if necessary.

For the phenylcyclohexanediol, the study can be furthered with the effective modes analysis and with the study of the effect of the solvation by DMSO. A comparison of the dipole moments would confirm for a different system the classical molecular dynamics method to achieve VCD spectroscopy.

Bibliography

- [1] L. A. Nafie, *Vibrational optical activity: principles and applications*. The Atrium, Southern Gate, Chichester, West Sussex: John Wiley & Sons, 2011.
- [2] S. Jähnigen, A. Scherrer, R. Vuilleumier, and D. Sebastiani, "Chiral crystal packing induces enhancement of vibrational circular dichroism," *Angewandte Chemie International Edition*, vol. 57, no. 40, pp. 13344–13348, 2018.
- [3] K. Jalkanen, I. Degtyarenko, R. Nieminen, X. Cao, L. Nafie, F. Zhu, and L. Barron, "Role of hydration in determining the structure and vibrational spectra of L-alanine and N-acetyl L-alanine N'-methylamide in aqueous solution: a combined theoretical and experimental approach," *Theoretical Chemistry Accounts*, vol. 119, no. 1, pp. 191–210, 2008.
- [4] E. Tajkhorshid, K. J. Jalkanen, and S. Suhai, "Structure and Vibrational Spectra of the Zwitterion L-Alanine in the Presence of Explicit Water Molecules: A Density Functional Analysis," *The Journal of Physical Chemistry B*, vol. 102, no. 30, pp. 5899–5913, 1998.
- [5] K. Frimand, H. Bohr, K. J. Jalkanen, and S. Suhai, "Structures, vibrational absorption and vibrational circular dichroism spectra of l-alanine in aqueous solution: a density functional theory and rhf study," *Chemical Physics*, vol. 255, no. 2-3, pp. 165–194, 2000.
- [6] J. Sadlej, J. C. Dobrowolski, and J. E. Rode, "VCD spectroscopy as a novel probe for chirality transfer in molecular interactions," *Chemical Society Reviews*, vol. 39, no. 5, pp. 1478–1488, 2010.
- [7] S. Jähnigen, A. Zehnacker, and R. Vuilleumier, "Computation of solid-state vibrational circular dichroism in the periodic gauge," *The Journal of Physical Chemistry Letters*, vol. 12, pp. 7213–7220, 2021.
- [8] CPMD, *Copyright IBM Corp 1990-2019, Copyright MPI für Festkörperforschung Stuttgart 1997-2001.*, 2022. <http://www.cpmd.org/> (accessed April 27th, 2022).

Chapter 6

Conclusions and perspectives

Vibrational circular dichroism and its general theory have been presented. Various methods to obtain VCD spectra were discussed, these methods include anharmonicity and temperature effects with various degrees. The method employed for the implementation of VCD spectroscopy in classical molecular dynamics based on force fields naturally accounts for such effects. Methods to access normal (harmonic) and effective (anharmonic) modes were discussed. From the effective modes, it is possible to assign spectral peaks from molecular dynamics trajectories to specific molecular vibrations.

Trans-1-amino-2-indanol was examined as a case study for VCD on a complex fluxional system. Because of this fluxionality, many structures coexist within tiny energy differences. These structures can be explored to a certain extent with static calculations though more conformers can be accessed through molecular dynamics, which crosses barriers and produces conformations not biased by intuition. In this case study, the effects of the solvent on the spectrum were scrutinised. In this system, the main contribution to the spectrum arises from the local solvent, at a distance less than 1 Å from the solute. This leads to the solvent influencing only indirectly the calculated spectra through structural effects. Ab-initio molecular dynamics simulations were shown to be more dependent on the starting point than simulations using classical molecular dynamics based on force fields. This latter MD can explore for longer periods of time as its computational cost is lower. As this system shows important solvent mobility, its solvation cannot be taken fully into account with small, finite-size clusters. Therefore VCD spectroscopy from FFMD presents itself as a complementary method to the static picture using the cluster-in-a-liquid model and AIMD simulations.

During this thesis, appropriate expressions for the electric and magnetic dipole moments were implemented in classical molecular dynamics based on polarisable force fields to obtain subsequently the IR and VCD spectra from the Fourier transform of the corresponding time correlations functions. Alanine in the gas phase under the zwitterionic and neutral forms was used as a benchmark system. AMOEBA force field was parametrised for this amino acid, which turns out to be the smallest chiral amino acid. It was shown that the inclusion of the polarisation of the atoms is important to capture the VCD signal. Moreover, important sampling times are needed to reach a converged spectrum. These sampling times exceed the capabilities of AIMD. The electric and magnetic dipole moments computed from FFMD were compared successfully to DFT calculations at each time step. This comparison validates the implementation in the gas phase.

With the FFMD spectra validated, the IR spectrum was decomposed into effective modes to assign molecular vibrations to their respective peak. This information is not

directly accessible from the MD simulation, thus an effective mode analysis had to be performed. This analysis is used as an opportunity to decompose the contributions from each functional group to the spectrum.

As the experimental interest lies in condensed phases, the implementation of VCD spectroscopy through classical molecular dynamics was extended to crystalline and solvated phases. Using this implementation, alanine was studied as a crystal, hydrated and solvated in dinitrogen. The crystal and hydrated alanine were simulated at 300 K and alanine in dinitrogen at 77.35 K. The inter-molecular interactions in the alanine crystal and their influence on the VCD spectrum were demonstrated via a shift of the bands corresponding to the NH_3 and CO stretching compared to their counterpart bands in the gas phase. These interactions hinder the convergence of the VCD spectrum of alanine in crystalline form. The IR spectrum of hydrated alanine presents a signal for the bending of the water solvent, as expected. However this is also the case for the VCD spectrum which is more unexpected. This could be due to insufficient sampling time, as shown through the normalised errors. This could also be caused by the water contributions to the spectrum that had not enough time to average out to zero. The force field and DFT moments in the condensed phases were also compared and shown to correlate to a very high degree. This contributed to validate our implementation of VCD spectroscopy with polarisable force fields within the framework of classical MD under diverse chemical environments and thermodynamical conditions. The studies of alanine in dinitrogen and of phenylcyclohexanediol in the crystalline phase and solvated in DMSO confirm the sensitivity of the VCD signal to the conformations of the system and their diversity.

The present work can be extended along various directions. A first direction would be a longer exploration of the hydrated alanine. This would assess the possible convergence of the system at longer sampling times. Moreover the effect, if any, of the periodic box size of the alanine crystal could be examined. For phenylcyclohexanediol, the electric and magnetic dipole moments need to be compared to dipole moments obtained via DFT, and the spectra examined through an effective mode analysis. The effective mode analysis of all the systems can be used to check the values of the frequencies against experimental results and to adjust the force field parameters, if necessary.

Some modes are more susceptible than others to a change of sign in the VCD spectra. As the intensities of the VCD spectra are a function of the cross correlation of the electric and magnetic dipole moments, if the angle $\xi(i)$ between the electric dipole moment and the magnetic dipole moment is smaller than 90° , the intensity of the corresponding peak is positive. If the angle is larger than 90° , the peak is negative [1]. This means that modes corresponding to dipole moments with an angle close to 90° are more susceptible to a change in the sign. These are called non-robust modes and can be changed by small perturbations (experimentally, e.g. solvent effects, and computationally, e.g. level of calculations). If $\xi(i)$ differs from 90° by at least 30° , the mode corresponds to a robust mode and will not change sign with ease [2]. An extension of the implementation could be to determine the angle between the two dipole moments for each molecule, to assess the robustness of the various modes. This will help with the comparison to other theoretical

and experimental results as the analysis will be more precise because only robust modes will then be considered.

A second extension of the code could be the determination of the electronic polarisability, to establish the Raman optical activity spectrum which arises from the difference in intensity of Raman scattered right- and left- circularly polarised light [3, 4].

During the MD simulations, the nuclei were considered as classical particles. However this leads to a slight overestimation of the frequencies, as in the harmonic approximation. Though, solving the quantum problem, thus Schrödinger's equation for nuclei is excessively expensive for the systems treated in this thesis. For system of this size, semi-classical approaches can be used and MD simulations can be corrected to account for vibrational delocalisation. For example, the MD trajectories can be initialized based on the harmonic normal mode coordinates and associated momenta, then given an appropriate quantum mechanical thermal energy, with an amplitude chosen so the excess energy equals the (harmonic) zero-point energy [5]. To approximate quantum-mechanical correlation functions, the ring-polymer molecular dynamics (RPMD) method can also be used. In this method, each of the N classical particles in the system is represented by a harmonic ring-polymer of n monomers and n is pushed toward infinity [6].

The main interest of using force field molecular dynamics is its capacity to treat large, explicitly solvated and fluxional systems. Thus, interesting perspectives lie in the determination of the vibrational signatures of polypeptides, fibrils and proteins. Various advantages of VCD have been shown for these systems. First, VCD facilitates the study of proteins as there is no interference from the water absorption [7]. Second, VCD has an unusual sensitivity to fibril growth and development in solution [8]. Fibrils with supramolecular chirality show highly enhanced VCD intensities and thus can be easily studied with this spectroscopic method [9]. Moreover, metal ions are found in intimate association with nucleic acids in their natural environment and VCD can be used to determine the effect of these ions on DNA structure. This is important as they can stabilize and destabilize biological structures [10].

Bibliography

- [1] V. P. Nicu and E. J. Baerends, "Robust normal modes in vibrational circular dichroism spectra," *Physical Chemistry Chemical Physics*, vol. 11, no. 29, pp. 6107–6118, 2009.
- [2] J. Sadlej, J. C. Dobrowolski, and J. E. Rode, "VCD spectroscopy as a novel probe for chirality transfer in molecular interactions," *Chemical Society Reviews*, vol. 39, no. 5, pp. 1478–1488, 2010.
- [3] L. D. Barron, *Molecular light scattering and optical activity*. The Edinburgh Building, Cambridge CB2 2RU, UK: Cambridge University Press, 2009.
- [4] D. Sidler, M. Meuwly, and P. Hamm, "An efficient water force field calibrated against intermolecular THz and Raman spectra," *The Journal of Chemical Physics*, vol. 148, no. 24, p. 244504, 2018.
- [5] N.-T. Van-Oanh, C. Falvo, F. Calvo, D. Lauvergnat, M. Basire, M.-P. Gaigeot, and P. Parneix, "Improving anharmonic infrared spectra using semiclassically prepared molecular dynamics simulations," *Physical Chemistry Chemical Physics*, vol. 14, no. 7, pp. 2381–2390, 2012.
- [6] X. Huang, S. Habershon, and J. M. Bowman, "Comparison of quantum, classical, and ring-polymer molecular dynamics infra-red spectra of $\text{Cl}^-(\text{H}_2\text{O})$ and $\text{H}^+(\text{H}_2\text{O})_2$," *Chemical Physics Letters*, vol. 450, no. 4-6, pp. 253–257, 2008.
- [7] G. Shanmugam and P. L. Polavarapu, "Vibrational circular dichroism of protein films," *Journal of the American Chemical Society*, vol. 126, no. 33, pp. 10292–10295, 2004.
- [8] S. Ma, X. Cao, M. Mak, A. Sadik, C. Walkner, T. B. Freedman, I. K. Lednev, R. K. Dukor, and L. A. Nafie, "Vibrational circular dichroism shows unusual sensitivity to protein fibril formation and development in solution," *Journal of the American Chemical Society*, vol. 129, no. 41, pp. 12364–12365, 2007.
- [9] M. Pazderková, T. Pazderka, M. Shanmugasundaram, R. K. Dukor, I. K. Lednev, and L. A. Nafie, "Origin of enhanced VCD in amyloid fibril spectra: Effect of deuteration and pH," *Chirality*, vol. 29, no. 9, pp. 469–475, 2017.
- [10] V. Andrushchenko, D. Tsankov, and H. Wieser, "Vibrational circular dichroism spectroscopy and the effects of metal ions on dna structure," *Journal of Molecular Structure*, vol. 661, pp. 541–560, 2003.

Titre: Caractérisation structurale des molécules chirales : développement d'outils de simulation moléculaire de spectres de dichroïsme circulaire vibrationnel

Mots clés : Dichroïsme circulaire vibrationnel, dynamique moléculaire classique, molécules flexibles

Résumé: Le dichroïsme circulaire vibrationnel (VCD) est la faible différence d'absorption des molécules chirales entre la lumière polarisée droite et gauche dans le domaine de l'infrarouge (IR). Cette technique spectroscopique est devenue de plus en plus populaire depuis sa découverte dans les années 1970 et a des applications prometteuses en pharmacologie en raison de sa capacité à déterminer les configurations absolues des molécules chirales. La forme des spectres VCD est très sensible aux changements mineurs de conformation et aux interactions moléculaires, ce qui en fait une sonde de l'isomérisation conformationnelle et de la solvata-tion. L'attribution des spectres VCD expérimentaux nécessite en général une comparaison avec les calculs théoriques, ce qui implique également que le solvant soit précisément décrit. Ce manuscrit décrit comment une modélisation basée sur des champs de force polarisables peut être exploitée pour interpréter les signaux VCD. Cette modélisation est appliquée à des molécules flexibles pour lesquelles les approches conventionnelles de chimie quantique reposant sur des approximations statiques sont limitées. En particulier, il est difficile de prendre en compte les anharmonicités et les effets de température finie, et le solvant est typiquement modélisé de manière simplifiée par un continuum ou par un nombre limité de molécules explicites. Dans le travail présenté ici, le champ de force AMOEBA mis en oeuvre dans le logiciel Tinker a été utilisé pour modéliser les moments dipolaires électriques et magnétiques de diverses molécules pour des applications en phase gazeuse, ainsi que dans des systèmes solvatés ou cristallins. L'échantillonnage et les informations dynamiques ont été obtenus par simulations de dynamique moléculaire classique. Les spectres IR et VCD anharmoniques obtenus à partir des trajectoires ont été interprétés à l'aide d'une analyse des modes effectifs dans laquelle les coordonnées internes sont décomposées linéairement par projection sur des modes effec-

tifs appropriés, les poids correspondants étant obtenus de telle sorte que le pic de chaque mode soit le plus localisé possible. Cette analyse permet de mettre en évidence les contributions des différents groupes fonctionnels de la molécule aux pics individuels du spectre. Comme premier exemple, le cas du trans-1-amino-2-indanol solvaté dans le diméthylsulfoxyde a été examiné en combinant des explorations statiques au niveau de la théorie de la fonctionnelle de la densité, des trajectoires courtes de MD *ab initio* (AIMD) et un large échantillonnage de la surface d'énergie potentielle à l'aide du champ de force. La comparaison avec les mesures existantes confirme l'importance de la flexibilité pour cette molécule, le rôle des molécules de solvant proches et leur contribution aux signaux VCD. La détermination des spectres VCD à partir du champ de force polarisable a ensuite été entreprise et ses performances ont été évaluées par comparaison avec des calculs de structure électronique. Cette comparaison a été effectuée sur la même trajectoire MD, en utilisant l'alanine en phase gazeuse comme système de référence. Les différentes contributions aux moments dipolaires sont discutées. Après avoir validé la méthodologie, la convergence statistique des spectres IR et VCD a été quantifiée via des mesures d'erreurs dédiées. Ces erreurs montrent à leur tour que la convergence des spectres VCD est plus lente que celle des spectres IR et que celle n'est atteinte qu'après un échantillonnage bien plus important que celui réalisable par AIMD. L'approche par champ de force polarisable a ensuite été appliquée à l'alanine en phases condensées. Là encore, la comparaison avec les simulations AIMD confirme que les moments dipolaires électriques et magnétiques peuvent être évalués de manière fiable. Ce travail ouvre la voie à la détermination du spectre VCD de systèmes chimiquement complexes inaccessibles via des méthodes qui tiennent compte explicitement de la structure électronique.

Title: Structural characterisation of chiral molecules: tool development for molecular simulations of vibrational circular dichroism spectra

Keywords: Vibrational circular dichroism, classical molecular dynamics, fluxional molecules

Abstract: Vibrational circular dichroism (VCD) is the weak difference in absorption for chiral molecules between right- and left-polarised light in the infrared (IR) range. It has become an increasingly popular spectroscopic technique since its discovery in the 1970s and has promising applications in pharmacology owing to its ability to determine absolute configurations of chiral molecules. The shape of VCD spectra is highly sensitive to minor changes in conformation and molecular interactions, which makes it a valuable probe of conformational isomerism and solvation. Structural assignment of experimental VCD spectra usually requires successful comparison with theoretical calculations, which also implies that the solvent is described with sufficient chemical accuracy. This manuscript examines how computational modelling based on polarisable force fields can be exploited to assist in the elucidation of VCD signals. This modelling is applied to flexible molecules for which conventional quantum chemical approaches relying on static approximations are necessarily limited. In particular, anharmonicities and finite temperature effects are difficult to account for, and the solvent is typically modelled in a rather rudimentary way through continuum descriptions or very few explicit solvent molecules. In the work presented here, the AMOEBA force field implemented in the Tinker software package was employed to model electric and magnetic dipole moments of various molecules of interest, for applications in the gas phase, as well as the condensed phases of solvated or crystalline systems. Sampling and real-time dynamical information were achieved by means of classical molecular dynamics simulations. Anharmonic IR and VCD spectra obtained from molecular dynamics trajectories were interpreted using an effective mode analysis in which the internal coordinates are linearly decomposed by projection onto suitable effective modes, the corresponding weights being obtained in such a way that the peak of each mode

is as localized as possible. This effective mode analysis sheds light onto the various contributions of the different functional groups of the molecule to individual peaks in the vibrational spectrum. As a first illustrative example, the case of trans-1-amino-2-indanol solvated in dimethyl sulfide was examined by combining static explorations at the level of density-functional theory, short ab initio MD (AIMD) trajectories, and extended sampling of the potential energy surface using the force field. Comparison with existing measurements confirms the importance of flexibility for this molecule, the active role of the nearby solvent molecules, and their contribution to the VCD signals. The determination of VCD spectra from the polarisable force field was then undertaken and its performance was assessed by comparison with electronic structure calculations, namely AIMD. This comparison was performed on the very same MD trajectory, using the alanine amino acid in the gas phase as our benchmark system. The various contributions to the dipole moments arising from the induced dipole moment or its time dependence are discussed. Having validated the methodology, the statistical convergence of the IR and VCD spectra was quantified by defining dedicated error measures. These errors show in turn that convergence of the VCD spectra is generally slower relative to IR spectra and that convergence is achieved only after extended sampling, way more than the amount achievable by AIMD. The force field approach was then extended and applied to alanine in crystalline and solvated phases. Here again, comparison with AIMD simulations confirms that electric and magnetic dipole moments can be reliably evaluated using the polarisable force field. This paves the way towards determining the VCD spectrum of chemically complex and flexible systems that are inaccessible via methods that explicitly account for electronic structure.