Theoretical Tools to Study Solvation in Liquid and Nanoconfined Phases

Vishal Kumar Porwal

UNIVERSITÉ DE LORRAINE

École doctorale C2MP

Theoretical Tools to Study Solvation in Liquid and Nanoconfined Phases

Thèse

présentée et soutenue publiquement le 12 décembre 2022

pour l'obtention du

Doctorat de l'Université de Lorraine

(Docteur en Chimie)

par

Vishal Kumar Porwal

Composition du jury

Président :	Pr. Claude Millot	Université de Lorraine
Rapporteurs :	Dr. Carine Clavaguéra, DR CNRS,	Université de Paris Saclay
	Pr. Abdenacer Idrissi	Université de Lille 1
Examinateurs :	Pr. Rachel Schurhammer,	Université de Strasbourg
	Dr. Mariachiara Pastore, CR HDR CNRS,	Université de Lorraine
Invités :	Dr. Antoine Carof, MCF	Université de Lorraine
	Pr. Cédric Carteret	Université de Lorraine
	Dr. Erwan André, MCF	Université de Lorraine
	Dr. Damien Cornu, MCF	Université de Lorraine
Encadrant :	Dr. Francesca Ingrosso, MCF HDR	Université de Lorraine

Laboratoire de Physique et Chimie Théoriques - UMR CNRS 7019

Mis en page avec la classe thesul.

Acknowledgments

I would like to thank the reviewers and jury members for their participation. I am grateful to my advisor Dr. Francesca Ingrosso, and my collaborators, Dr. Giacomo Prampolini and Dr. Antoine Carof, for their guidance and scientific discussions. I am grateful to the members of the monitoring committee, Dr. Guillaume Stirnemann, and Dr. Mariachiara Postera, for their suggestions and encouragement. I am thankful to Dr. Fabien Pascale and Mr. Christian Causse for the technical set-up and Mrs. Séverine Bonenberger for the help in administrative procedures. The organization of my new life in a new French city would not have been possible without the immense help of the university administration. I am pleased to have an amazing group of friends and coworkers.

The growth of a human infant takes place in several stages, from birth to adulthood. This development is multidimensional and involves the contribution of many aspects and ideologies. In my growing experience of this vast unknown world, I am indebted to my Mother, Father, Brothers, Sisters-in-law, and the newest member of the family, my little Niece.

This acknowledgment would be incomplete without mentioning the inexplicable experience of the Service Universitaire des Activités Physics et Sportives (SUAPS), the Orchester Symphonique Universitaire de Lorraine (OSUL) and the Revel Ensemble à Cordes (REAC).

List of publications

Hydration Effects on the Vibrational Properties of Carboxylates: From Continuum Models to QM/MM Simulations; **Vishal Kumar Porwal**, Antoine Carof, and Francesca Ingrosso (J. Comp. Chem., under review)

Abstract

Confinement on a molecular scale is one of the most fascinating and complex fields in continuous progress. Once rationalized, the consequences of confinement on the molecular and electronic structure of chromophores can be used to fine-tune their optical properties and thus exploit them in the development of photochemical technologies. In material chemistry, untangling the complex nature of nanoconfined phases can help provide essential knowledge to fine-tune the synthesis of new compounds with versatile properties. This project is devoted to the development of ad hoc computational strategies to achieve a molecular interpretation of the impact of the environment on the conformational, vibrational, and optical properties of organic molecules. In collaboration with an experimental group, we analyzed the behavior of organic anions confined in the interlayer of claylike materials. Focusing on the evolution of the carboxylate bands with increasing hydration, we characterized the changes in the binding modes of the anion by using classical molecular dynamics simulations. The second part of the project, carried out with Italian collaborators, is based on an integrated multilevel approach providing a sophisticated force field for 2,2'-bipyridine-3,3'-diol. This molecule undergoes excited state intramolecular proton transfer, and experimental data point to a fine sensitivity of its properties to a nanoconfined environment. Our study of the potential energy surface and the absorption spectrum in water using a sequential classical-quantum mechanical approach brought significant progress in the characterization of the tautomeric equilibria and their effect on the optical properties of the chromophore.

Résumé

Le confinement à l'échelle moléculaire est l'un des domaines les plus fascinants et les plus complexes en constante évolution. Une fois rationalisées, les conséquences du confinement sur la structure moléculaire et électronique des chromophores peuvent être utilisées pour affiner leurs propriétés optiques et ainsi les exploiter dans le développement de nouvelles technologies. En chimie des matériaux, démêler la nature complexe des phases nanoconfinées peut guider dans la synthèse de nouveaux composés aux propriétés polyvalentes. Ce projet est consacré au développement de stratégies computationnelles ad hoc pour parvenir à une interprétation moléculaire de l'impact de l'environnement sur les propriétés conformationnelles, vibrationnelles et optiques de molécules organiques. En collaboration avec un groupe expérimental, nous avons analysé le comportement d'anions organiques confinés dans des matériaux argileux. En nous focalisant sur l'évolution des bandes carboxylates avec l'augmentation de l'hydratation, nous avons caractérisé les changements des modes de liaison de l'anion avec la surface en utilisant des simulations de dynamique moléculaire classique. La deuxième partie du projet, réalisée avec des collaborateurs italiens, est basée sur une approche intégrée multiniveau pour obtenir un champ de force sophistiqué du 2,2'-bipyridine-3,3'-diol. Cette molécule subit un transfert de proton intramoléculaire à l'état excité, et les données expérimentales indiquent une sensibilité de ses propriétés à un environnement nanoconfiné. Notre étude de la surface d'énergie potentielle et du spectre d'absorption dans l'eau par une approche séquentielle classique-quantique a apporté des progrès significatifs sur la caractérisation des équilibres tautomères et sur leur effet sur les propriétés optiques du chromophore.

Résumé détaillé

Ce projet de doctorat a été consacré au développement d'outils théoriques et computationnels pour décrire les effets de solvatation et de confinement dans des systèmes complexes. En particulier, nous avons modélisé les propriétés vibrationnelles de molécules organiques intercalées dans des systèmes argileux (hydroxydes doubles lamellaires - LDH) et les propriétés optiques de chromophores qui sont sensibles aux effets de confinement. Ces deux parties du projet ont bénéficié de collaborations avec d'autres groupes, théoriciens et expérimentateurs. D'importantes modifications du projet initial, qui était censé inclure une mobilité pour travailler dans un laboratoire partenaire à l'étranger, ont dû être apportées en raison des fortes limitations dictées par la pandémie de coronavirus. Cependant, nous pouvons maintenant affirmer que ces modifications nous ont permis de conserver avec succès l'idée originale du projet, c'est-à-dire de développer et de combiner des stratégies de calcul à utiliser dans une grande variété d'applications lorsque le système moléculaire d'intérêt interagit fortement avec son environnement et est soumis à des effets de confinement.

Les minéraux, en particulier les argiles naturelles, sont utilisés comme anti-inflammatoires, antiseptiques, cicatrisants et cosmétiques depuis la préhistoire. Leur utilisation en tant que principes actifs (par exemple, comme protecteurs gastro-intestinaux), et dans la médecine esthétique, est largement connue. La recherche scientifique dans ce domaine n'a cessé d'évoluer, et les avancées les plus récentes ont conduit au développement de nouveaux matériaux biocompatibles inspirés de la nature et basés sur des approches synthétiques. Parmi ces matériaux, les hydroxydes doubles lamellaires présentent une toxicité faible ou nulle, une bonne biocompatibilité et la possibilité de conduire à un relargage contrôlé. Les

LDH sont constitués de couches inorganiques ayant une surface chargée positivement et s'empilant les unes sur les autres, et incorporant entre deux couches des molécules d'eau et les ions négatifs, pour permettre un environnement neutre. Des anions de natures très différentes peuvent être intercalés pour obtenir des systèmes hybrides, comme des molécules inorganiques, organiques ainsi que biologiques, et des macromolécules (protéines, ADN), conduisant à des matériaux hybrides bioinorganiques. Les molécules intercalées dans les LDH ayant une charge négative portée par des groupes carboxylate, comme les acides carboxyliques déprotonés, ont été systématiquement synthétisées depuis les années 1970, et il a été suggéré que l'environnement LDH pouvaient assurer la stabilisation, le transport et la libération de médicaments, comme par exemple l'ibuprofène, contenant de tels groupes.

Parmi les anions intercalés contenant des groupes carboxylates, les acides aminés ont attiré une attention croissante pour la production de matériaux hybrides. Les argiles anioniques naturelles pourraient avoir joué un rôle dans la concentration des unités d'acides aminés, fournissant un environnement propice à la polymérisation menant aux polypeptides, la première étape vers l'origine de la vie dans des conditions abiotiques. L'aspartate, la forme déprotonée de l'acide aspartique, possède une chaîne latérale déprotonée à pH neutre et basique et il était présent dans les premières protéines synthétisées naturellement. Il a été suggéré que dans le mécanisme de l'interaction anion-surface, relayée par des molécules d'eau intercalées, le groupe amino des acides aminés pourrait jouer un rôle dans l'ancrage de la surface. Si l'on compare la structure du succinate et de l'aspartate, la seule différence structurelle est l'absence/présence d'un groupe amino. L'étude des dif-

férences dans les modes de liaison de ces deux molécules contenant des unités carboxylate pourrait donc apporter des informations à exploiter pour comprendre un tel mécanisme. Dans un premier temps, nous avons décidé de nous concentrer sur la description des propriétés structurelles et vibrationnelles des carboxylates modèles dans l'eau, où ils peuvent agir comme donneurs de liaisons hydrogène et, dans le cas des acides aminés, comme accepteurs de liaisons hydrogène (en fonction de l'acidité des atomes d'hydrogène de l'acide aminé).

Les molécules contenant des groupes carboxylates sont utilisées dans un large éventail d'applications. La structure ainsi que les propriétés électroniques de la partie carboxylate en font une unité idéale pour que la molécule soit exploitée en tant que ligand. En chimie de coordination, les carboxylates, en particulier les multicarboxylates, sont utilisés pour construire des polymères de coordination et de nouveaux matériaux. Ils peuvent être utilisés pour synthétiser des MOF (metal organic frameworks), des matériaux photocatalytiques, et le mécanisme de liaison carboxylate-métal dans les systèmes biologiques affecte l'affinité et la sélectivité des métaux et, par conséquent, la fonction des métalloprotéines.

Malgré l'importance de ces systèmes, les travaux computationnels dans la littérature, portant sur la structure, la capacité de formation de liaisons hydrogène et leur influence sur les propriétés vibratoires de carboxylates modèles, manquaient encore. Nous avons donc développé une approche bottom-up (des anions solvatés dans un solvant continu aux simulations hybrides QM/MM) pour explorer ces propriétés (Chapitre 2). Ceci a permis d'obtenir des résultats dans liquide, à comparer ensuite avec nos résultats de dynamique

moléculaire dans le système confiné. Nous avons étudié l'espace conformationnel et les propriétés vibrationnelles de trois anions contenant des groupes carboxylates : le succinate, l'aspartate zwitterionique et l'aspartate ayant un groupe amino déprotoné. Plus spécifiquement, nous nous sommes concentrés sur la bande d'élongation C-O du carboxylate, qui est sensible à l'environnement et, par conséquent, peut être utilisée comme une sonde de la structure locale et de la solvatation, et sur une exploration de l'espace conformationnel des molécules interagissant avec l'eau. Nous sommes partis des anions dans un solvant continuum, puis nous sommes passés à une approche de microsolvatation (simplifiée), dans laquelle nous avons inclus une molécule d'eau à côté de chaque groupe polaire et immergé l'agrégat ainsi obtenu dans un solvant continuum représenté comme un diélectric. En outre, nous avons effectué des simulations MD hybrides QM/MM des anions dans l'eau explicite, incluant ainsi les effets de température finie.

Dans les calculs de structure électronique, deux méthodes DFT et une méthode semi-empirique ont été employées. Les résultats sur les géométries prédites, tant pour les anions que pour les clusters, étaient remarquablement similaires. Cela pourrait être surprenant, compte tenu des artefacts découlant de l'application d'hamiltoniens semi-empiriques pour décrire les interactions non covalentes. Cependant, nous pouvons émettre l'hypothèse que les fortes interactions électrostatiques entre les anions chargés négativement (possédant des groupes polaires) et l'eau peuvent être plus facilement modélisées que l'interaction entre les systèmes neutres puisque les limites des méthodes semi-empiriques sont plus apparentes lorsque de faibles interactions non covalentes sont présentes.

En comparant les résultats obtenus pour la bande d'élongation C-O du carboxylate

sur la base de l'analyse en modes normaux, nous avons constaté que l'accord global avec les mesures expérimentales est satisfaisant et en accord avec la littérature. Une conclusion intéressante que nous pouvons tirer est que l'on observe des différences plus importantes lors de la comparaison des résultats obtenus avec la même méthode de chimie quantique, mais un modèle de solvatation continuum différent, que lors de la comparaison de deux méthodes de chimie quantique différentes. Cette constatation doit nous mettre en garde contre la réalisation de calculs de telles propriétés sans une analyse comparative minutieuse.

Dans le Chapitre 3, nous avons étudié des LDH intercalés avec des molécules contenant des carboxylates (succinate, aspartate, glutamate). Les travaux expérimentaux de nos collaborateurs, basés sur les techniques de diffraction des rayons X, ont fourni des informations sur les paramètres structurels de l'hôte inorganique (distance entre les cations dans les couches ou espacement entre les couches) et les spectroscopies vibrationnelles ont apporté des informations sur l'évolution du spectre vibrationnel des invités intercalés lors de l'hydratation. Nous avons proposé une approche basée sur des simulations de dynamique moléculaire des systèmes synthétisés et caractérisés par le groupe expérimental, incluant une description atomistique à la fois de la surface inorganique et de la région interlamellaire avec des quantités croissantes d'eau, en choisissant différents états d'hydratation le long de l'isotherme d'adsorption expérimental. Les expériences ont été réalisées sous humidité contrôlée, donnant ainsi accès au nombre précis de molécules d'eau par anion dans chaque système. Bien que l'application de simulations moléculaires sur les carboxylates et les acides aminés intercalant la LDH ait fourni des informations re-

marquables dans des études publiées dans la littérature, nous n'avons pas connaissance de travaux théoriques antérieurs examinant explicitement comment ces différentes structures moléculaires peuvent induire des différences dans le mécanisme d'adsorption à l'intérieur du matériau et pour différents taux d'hydratation.

Des différences significatives ont été observées dans la structure de la région interlamellaire en fonction des espèces intercalées considérées, ce qui nous a motivés à pousser plus loin l'analyse pour rationaliser l'effet de la charge et de la structure de l'anion sur l'organisation locale. La réorientation de l'anion dans la région interlamellaire joue un rôle très important dans le mécanisme d'adsorption. Or, en augmentant la quantité d'eau, chaque anion réagit différemment à l'augmentation de l'espace interlamellaire.

Le succinate se réoriente assez brusquement (ce qui est cohérent avec la forte augmentation de l'espace interlamellaire et avec la tendance de l'isotherme d'adsorption) et passe d'une situation dans laquelle il se lie à la surface dans un arrangement bidenté pontant, à une situation dans laquelle les modes de liaison sont principalement dans un arrangement bidenté par chélation. En présence de taux d'hydratation plus importants, les anions peuvent être trouvés soit dans une orientation verticale par rapport à la surface, soit horizontalement par rapport la surface. L'évolution de différentes populations d'anions, interagissant différemment avec la surface, est cohérente avec l'évolution de la bande d'élongation du carboxylate. L'aspartate et le glutamate subissent des transitions plus douces, et dans le cas du glutamate, une réorientation complète n'est pas observée puisque l'anion est plus volumineux et, en même temps, présente une plus grande mobilité conformationnelle. Dans les systèmes intercalant des acides aminés, le groupe amino peut

participer à la liaison avec la surface, bien que cela se produise principalement dans les systèmes anhydres ou à faible hydratation. Dans ces conditions, les acides aminés voisins peuvent également interagir le long d'un chemin de liaison hydrogène (COO...HNH) qui a été déclaré propice à la réactivité chimique conduisant à la formation de liaisons peptidiques, et donc à la polymérisation, pour former des polypeptides dans des conditions abiotiques. Des extensions intéressantes du présent travail porteront sur les propriétés des chromophores confinés dans les LDH et, éventuellement, la réactivité chimique avec des applications possibles dans le domaine de la photocatalyse.

Dans la deuxième partie de ce projet de doctorat, nous nous sommes intéressés à l'opportunité de développer un champ de force sophistiqué pour décrire un chromophore organique, dont le spectre d'absorption s'est avéré sensible aux effets de confinement. Cette molécule, le [2,2'-bipyridyl]-3,3'-diol (BPOH2), a été initialement proposée par Abou-Zied en 2006 comme modèle de paire de bases dans l'ADN. L'interaction et la reconnaissance entre deux paires de bases dans la structure de l'ADN sont médiées par la formation de liaisons hydrogène intermoléculaires. La réaction de (double) transfert de proton le long des directions de ces liaisons peut conduire à des tautomères énol-imine qui, bien qu'ils ne soient pas stables, peuvent survivre suffisamment longtemps pour conduire à un mauvais appariement des bases et apporter des altérations du code génétique.

Les travaux d'Abou-Zied ont montré la sensibilité particulière de la sonde à l'environnement confiné dans une microcavité de cyclodextrine. L'idée de ces travaux était d'évaluer le degré d'hydrophobicité de cavités de différentes tailles (de la cyclodextrine α à la cyclodextrine γ) et de différentes natures (différents substituants aux unités pyranose). Lorsque

la cavité est trop petite pour que le BOPH2 soit encapsulé (cyclodextrine α), la liaison se fait en formant un complexe 1:2 avec deux unités de cyclodextrine. C'est la situation qui conduit aux effets d'hydrophobicité les plus importants. D'autre part, lorsque la taille de la cavité est suffisamment grande (cyclodextrine γ), le chromophore peut entrer dans la cavité et être entouré de quelques molécules d'eau. Cela semble stabiliser une forme di-zwitterion (la même que celle qui est censée être stabilisée dans l'eau liquide), donnant lieu à des bandes d'énergie plus basses dans le spectre d'absorption du BOPH2.

Bien que quelques études théoriques dans la littérature se soient intéressées à cette molécule, pratiquement toutes se sont concentrées sur le mécanisme de transfert de protons dans l'état excité. D'autre part, l'évolution du spectre d'absorption en présence de solvants de différentes natures n'avait pas été modélisée. En particulier, l'existence de formes di-zwitterioniques (dicéto) n'est qu'une hypothèse faite pour expliquer le comportement spectroscopique dans l'eau et cela n'a pas encore été démontré. De plus, jusqu'à présent, aucune analyse computationnelle n'a abordé le comportement spectroscopique de la molécule sous confinement.

Dans ce travail de thèse, nous avons construit un champ de force sophistiqué du BPOH2 (Chapitre 4). Ce projet a profité d'une collaboration récemment établie avec le Dr Giacomo Prampolini, ICCOM-CNRS, Pise. Le Dr. Prampolini a co-développé une procédure rigoureuse et semi-automatisée, implémentée dans un code nommé Joyce, permettant de construire des champs de force à partir de calculs quantiques, qui peuvent être utilisés pour décrire une molécule dans son état fondamental ainsi que dans un état électronique excité. À long terme, notre objectif final sera de donner une interprétation

moléculaire des différences mesurées dans les spectres d'absorption et de fluorescence lorsque la molécule est confinée dans la macrocavité cyclodextrine, et de fournir une interprétation du rôle du solvant dans le mécanisme de transfert de proton à l'état excité. La modélisation de BPOH2 entouré d'une cavité moléculaire conduirait à exclure les modèles de continuum pour décrire l'environnement. Un traitement quantique complet du système, d'autre part, conduirait à des schémas de calcul coûteux, même avec des approches hybrides QM/MM, surtout si des propriétés moléculaires comme le spectre UV/VIS sont visées.

Un compromis raisonnable peut être atteint si des méthodologies hybrides séquentielles classiques/QM sont appliquées. Elles consistent à mener des simulations de dynamique moléculaire, à partir desquelles des configurations du système d'intérêt sont extraites et utilisées pour effectuer des calculs de chimie quantique. La propriété d'intérêt est finalement dérivée comme une convolution des résultats obtenus au niveau QM sur les différents instantanés. Toutefois, si l'évolution de la structure du système au cours de la simulation MD (en tenant compte des effets de température) n'est pas raisonnablement prédite, certains artefacts peuvent apparaître dans la description de la propriété recherchée. Dans notre travail, l'application du protocole de Joyce nous a donné accès à une surface d'énergie potentielle pour BPOH2 au niveau de précision QM et à la possibilité de traiter la dynamique moléculaire en présence de l'état excité. L'étude de l'état fondamental que nous avons réalisée avec le Dr. Prampolini nous a conduit à de nombreux résultats intéressants..

Au cours de la paramétrisation de la molécule de l'état fondamental, nous avons

pris conscience de l'importance de deux autres tautomères qui sont stabilisés dans l'eau : une forme intermédiaire céto-énol et une forme dicéto. En particulier, ils semblent être responsables de bandes de basse énergie dans le spectre d'absorption qui sont particulièrement sensibles à l'environnement et pourraient fournir des informations sur le degré d'hydratation (et donc d'hydrophilicité) de l'espace confiné dans les cavités de la cyclodextrine. En effet, nous avons montré par des calculs QM que ces deux formes sont stables dans l'eau et que le maxima de leurs spectres d'absorption sont localisés aux longueurs d'onde correspondant à ces bandes, l'accord étant très satisfaisant. Cette découverte, bien qu'extrêmement prometteuse, nous a obligés à fournir un champ de force différent pour chaque tautomère, nécessitant des efforts pour la description du système à l'état fondamental plus importants que prévu. Nous sommes actuellement en train de finaliser cette étape. D'autres perspectives naturelles découlant de ce travail sont l'obtention du spectre d'absorption pour les trois tautomères en utilisant l'approche séquentielle dans l'eau et dans d'autres solvants. Nous passerons ensuite à la complexation du chromophore dans la cavité de la cyclodextrine, un sujet pour lequel le laboratoire a une longue expérience. A plus long terme, nous passerons à la modélisation de la réactivité chimique, en particulier à la description du transfert intramoléculaire de protons dans l'état excité, une question particulièrement difficile qui suscite beaucoup d'attention dans la communauté scientifique.

Contents

Acknowledgments i

List of publications iii

Abstract vi

Résumé vii

Résumé détaillé ix

Introduction 1

References 11

Chapter 1
Theoretical Framework

1.1 Introduction 18

1.2 Electronic structure calculations 19

1.2.1 The Adiabatic and Born-Oppenheimer Approximations 20

1.2.2 Limitations for multielectron system 22

1.2.3 Sempiempirical quantum chemistry methods 23

1.2.4 Density Functional Theory 25

1.2.5 Continuum solvation models 28

1.3 Molecular dynamics simulations 30

1.3.1 Integration of Newton's equations of motion . . . 32
1.3.2 Periodic boundary conditions and thermodynamic ensembles . . . 34
1.4 Interaction energy of the system in terms of a force field . . . 35
1.4.1 Bonding Interactions . . . 37
1.4.2 Non-Bonding Interactions . . . 38
1.4.3 Cutoff schemes for non-bonded interactions . . . 39
1.5 Force Field parametrization with the Joyce protocol . . . 40
1.6 Analysis of the simulated trajectory . . . 46
1.6.1 Radial distribution functions . . . 47
1.6.2 Calculation of power spectra . . . 47
1.7 Summary . . . 50

References **53**

Chapter 2
Hydration Effects on the Vibrational Properties of Carboxylates

2.1 Introduction . . . 58
2.2 Computational details . . . 60
2.3 Results and discussion . . . 62
2.3.1 Results obtained using electronic structure calculations . . . 63
2.3.2 Results obtained using QM/MM molecular dynamics . . . 79
2.4 Conclusions . . . 86

References **89**

Chapter 3
MD Simulations of Nanoconfined Carboxylates in LDH

3.1 Introduction . . . 96
3.2 Computational studies of LDH . . . 99

3.3 Computational Setup . . . 101
3.3.1 Simulation protocol . . . 103
3.4 Results and discussion . . . 103
3.4.1 Radial distribution function . . . 104
3.4.2 Orientation, conformations, binding modes of the anions in the interlayer . . . 112
3.4.3 Number density along the z-axis . . . 118
3.4.4 Vibrational properties of intercalated anions . . . 121
3.5 Conclusions . . . 127

References **131**

Chapter 4
Modeling the Optical Properties of 2,2'-Bipyridine-3,3'-Diol in Water

4.1 Introduction . . . 136
4.2 Integrated multilevel QM/MM computational approach . . . 138
4.2.1 QM calculations . . . 139
4.2.2 Intramolecular FF parametrization with Joyce . . . 139
4.2.3 MD simulations . . . 140
4.2.4 Snapshot extraction and calculation of the UV-VIS spectra . . . 141
4.3 Results and discussion . . . 142
4.3.1 QM calculations . . . 142
4.3.2 FF parametrization . . . 145
4.3.3 MD simulations . . . 148
4.3.4 Absorption spectra . . . 151
4.4 Preliminary conclusions . . . 154

References **155**

Conclusions and Future Perspectives **159**

References **165**

Appendix A First appendix **169**

Appendix B Second appendix **171**

B.1 Experimental measurements for LDH intercalating organic anions 172

B.2 Force fields used in MD simulations 173

B.2.1 Clay force field 173

B.2.2 Generalized amber force field 176

B.2.3 Force fields for water 177

B.3 Supporting information on the MD simulations of LDH intercalating organic anions 179

References **185**

Appendix C Third appendix **187**

C.1 Conformers of BPOH2 187

C.2 FF parameters of diol 188

C.3 FF parameters of diketo 191

C.4 MD simulations, calculation of absorption spectra 194

List of Figures **195**

Introduction

Minerals, in particular natural clays, have been used as anti-inflammatories, antiseptics, cicatrizers, and cosmetics since prehistory.[1] Their use as active principles (such as, for instance as gastrointestinal protectors) and in therapeutic activity and aesthetic medicine is widely known. The scientific research in this field has continuously evolved, and the most recent advances have led to the development of new biocompatible materials inspired by nature and based on synthetic approaches. The knowledge of the molecular structure and of the mechanisms related to the action of clays has pushed even further their potential use in the pharmaceutical industry, in particular for the delivery of drugs and bio-active molecules.[2] Among these materials, layered double hydroxides (LDH) possess low or null toxicity, good bio-compatibility, and the possibility of leading to a controlled release. LDH comprises inorganic layers having a positively charged surface and packing together, including water and negative ions in the interlamellar region to allow a neutral environment. Anions of very different nature can be intercalated to obtain hybrid systems: inorganic, organic as well as biological molecules and macromolecules (proteins, DNA),[3–5] leading to bioinorganic hybrid materials.[6–11] LDH intercalating molecules with a negative charge born by carboxylate groups, such as deprotonated car-

boxylic acids, have been systematically synthesized since the 1970s,[12], and it has been suggested that they can provide stabilization, transport, and release of drugs, such as for instance ibuprofen, containing such groups.[13]

Among the possible intercalates containing carboxylate groups, amino acids have attracted increasing attention for the production of hybrid materials.[14–16] Natural anionic clays might have played a role in concentrating amino acid units, providing a propitious environment for the polymerization leading to polypeptides, the first step towards the origin of life in abiotic conditions.[17] Aspartate, the deprotonated form of aspartic acid, has a deprotonated side chain at neutral and basic pH and it was present in the first naturally synthesized proteins.[18] It has been suggested that in the mechanism of the anion-surface interaction, mediated by intercalated water molecules, the amino group in amino acids may play a role in anchoring the surface.[19] If we compare the structure of succinate and aspartate, the only structural difference is the absence/presence of an amino group (see Figure 1). Studying the differences in the binding modes of these two molecules containing carboxylate units would therefore bring some information to be exploited in understanding such a mechanism.

It has been shown that the conformations adopted by intercalated biomolecules can either be retained or modified, inducing unfolding of the secondary, tertiary, or quaternary structure,[4] and that intercalated DNA reveals to be stabilized compared to bulk water.[20] Given the importance of the structure of biomolecules to deliver their biological function, understanding the way in which nanoconfinement within LDHs affects the configuration, the dynamics, and the reactivity of intercalated species is key to a rational

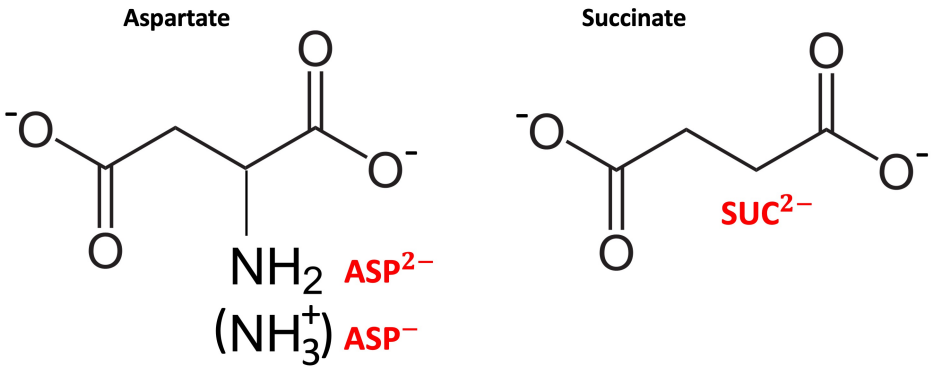


Figure 1: Structure of the aspartate anions (which will be named ASP^- and ASP^{2-}, depending on the protonation state) compared with the structure of the succinate anion (SUC^{2-}).

design for hybrid materials.

The structure of LDH intercalated with some carboxylic acids,[21] amino acids[22] and DNA[20] has been studied by means of molecular dynamics (MD) simulations. Indeed, significant differences have been observed in the structure of the interlamellar region based on the species considered, which motivates us to push further the analysis to rationalize the effect of the anion charge and structure on the local organization. For instance, different amino acids can lead to different orientations with respect to the inorganic layer and the organization water is strongly dependent on the balance between the interactions with the hydrophilic and hydrophobic groups. The results in the literature show the complexity of the local environment in LDH, a complexity that we shall approach by means of the interplay between theory and experiments.

In the original version of this Ph.D. project, we were meant to carry out an investigation of the effect of nanoconfinement on the structure and the dynamics of polypeptides of increasing size, the goal of which would have been be to provide a molecular understand-

ing of the factors that may lead either to stabilization (and thus protection) or to relevant conformational changes preventing the molecule from delivering its function. This project was established in collaboration with Prof. C. Carteret, Dr. E. André (Université de Lorraine, LCPME), and Prof. A. Bastida (Universidad de Murcia, Espagne) and was funded by the PICS program of the CNRS. Systems of increasing complexity would be selected, synthesized, and characterized by the experimental members of the project (LCPME), and in a parallel way, MD simulations would be carried out on them to explicitly study the conformational space (theoretical work to be performed in collaboration with Prof. Bastida). The latter collaboration would have helped to shed light on the mechanism of vibrational energy transfer involving the peptides and would have been based on previous work done together,[23–25] In addition, more recent results by the Bastida's group showed that conformational changes of dipeptides in water could be induced simply by equilibrium fluctuations of the internal energy,[26] an interesting hypothesis to be tested in confined systems. The original project had to be modified since the Covid health crisis led to a strong reduction of mobility and the stays that were planned in Bastida's group in order to acquire the basics of the instantaneous normal modes approach that he developed, as well as the training with the computer codes, could not take place.

In the first part of this Ph.D. project, we decided to focus on the description of the structural and vibrational properties of intercalated carboxylates. Experimental work based on X-ray diffraction techniques provided information on the structural parameters of the inorganic host (distance between cations within the layers or interlayer spacing) and vibrational spectroscopies brought information about the evolution of the vibrational

spectrum of the intercalated guests upon hydration.[27] We proposed an approach based on molecular dynamics simulations of the systems synthesized and characterized by the experimental group, including an atomistic description of both the inorganic surface and of the interlamellar region with increasing amounts of water, by choosing different hydration states along the experimental adsorption isotherm. Experiments were carried out under controlled humidity, thus giving access to the precise number of water molecules per anion in each system. Although remarkable insights were provided by applying molecular simulations on LDH intercalating carboxylates and amino acids,[19, 21, 28–30] we are not aware of previous theoretical work explicitly examining how these different molecular structures can induce differences in the adsorption mechanism inside the material and for different hydration states. When analyzing the results obtained from molecular dynamics, we were confronted with a complex interplay between different interactions, which was particularly difficult to disentangle. In particular, the way in which the carboxylate units interact with the surface, with water, and with each other required important efforts, given the sensitivity of this polar group to the environment, particularly with respect to the formation of the hydrogen bond.

We decided to first focus on model carboxylates in water, in which they can act as hydrogen bond donors and, in the case of amino acids, as hydrogen bond acceptors, depending on the acidity of the amino hydrogen atoms. This opened the door to an independent computational study carried out within this project, which delivered interesting results. Molecules containing carboxylate groups are used in a wide range of applications. The structure as well as the electronic properties of the carboxylate moiety make it an ideal

unit for the molecule to be exploited as a ligand. In coordination chemistry, carboxylates, in particular multi-carboxylates, behave as bridging ligands and are used to build coordination polymers and new materials.[31] They can be used to synthesize metal-organic frameworks,[32] photocatalytic materials,[33], and inorganic open frame materials used for sorption and catalytic processes.[34] Given their strong interaction with lanthanides, they are used to stabilize lanthanide ions in water solution and reduce their toxicity, thus allowing the development of luminescent probes for cell imaging.[35, 36]

In organic synthesis using atom and step economy principles, carboxylates can assist in the catalytic activity of Ru complexes.[37] Polypyridine-Ru complexes functionalized with carboxylate groups have been employed for the design of dye-synthesized solar cells since those groups allow anchoring with the TiO_2 electrode surface and in electrocatalysis.[38, 39]

The mechanism of carboxylate-metal in biological systems binding affects the affinity and selectivity of metals and, thus, the function of metalloproteins.[40, 41] The affinity of carboxylates to metal cations are also exploited in the technology based on cellulose to produce nanofibrils, facilitating the process leading to biocompatible hydrogels,[42] and in the mechanism leading to hydrogels based on alginates and pectin.[43, 44] In the field of biocompatible materials, the carboxylate units of amino acids, peptides, and other biomolecules are involved as ligands in the formation of metal-biomolecule frameworks.[45]

The ability of carboxylates to act as anchoring groups is particularly important in the interaction with mineral surfaces. The possibility of mineral dissolution in water strongly depends on the interaction at the mineral surface/water interface. In principle,

it is possible to control the solid growth or dissolution by triggering the adsorption of ions that are initially dispersed in the surrounding solution. In particular, polycarboxylates are used to control particle aggregation for calcite in water.[46] In the field of metallurgy, the separation of minerals from ore bodies is often based on flotation through the use of surfactants containing fatty acids, which interact via coordination with calcium ions on the surface.[47] Modifications of the wettability of reservoir rock, leading to an improved oil recovery process, can be induced through the desorption of organic carboxylates from the rock surface by interaction with surfactants.[48–50]

Apart from two studies based on a microsolvation approach,[51, 52] computational work addressing the structure, the hydrogen bond formation capability, and its influence on the vibrational properties of model carboxylates was still missing. We thus developed a bottom-up approach (from anions solvated in a continuum solvent to hybrid QM/MM simulations) to explore those properties. This provided results in the bulk, to be compared with our molecular dynamics results in the confined system.

In the second part of this Ph.D. project, we were interested in the opportunity of developing a sophisticated force field to describe an organic chromophore, the absorption spectrum of which has been shown to be sensitive to confinement effects.[53] This molecule, [2,2'-bipyridyl]-3,3'-diol (BPOH2, see Figure 2), was originally proposed by Abou-Zied in 2006 as a model of a base pair. The interaction and recognition between two base pairs in the DNA structure are mediated by the formation of intermolecular hydrogen bonds. Double-proton transfer reaction along the directions of those bonds can lead to enol-imine tautomers that, though not stable, may survive long enough to lead to base mispairing,

a precursor of alterations of the genetic code.

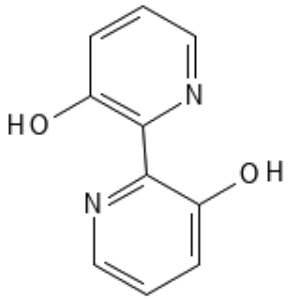


Figure 2: The structure of [2,2'-bipyridyl]-3,3'-diol

BPOH2 is well known for giving rise to excited state intramolecular proton transfer (ESIDPT) in solution, and molecular dynamics simulations of a DNA strand containing it showed that it does not perturb the local structure of the helix.[54] The complex photochemical properties of this molecule made it an ideal probe to study ESIDPT, since the early 1980's, with increasing attention on the solvent effects (see for instance [55–59] and references therein) In most recent years, the literature on this subject has grown rapidly. Some studies focused on how a confined environment can be used to tune its optical properties.[54, 60–63]

In particular, the work by Abou-Zied showed the peculiar sensitivity of the probe to the confined environment within a cyclodextrin microcavity, The idea of that work was to assess the degree of hydrophobicity of cavities of different sizes (from α to γ cyclodextrin) and nature (different substituents at the pyranose units). When the cavity is too small for BOPH2 to be encapsulated (α cyclodextrin), binding occurs by forming a 1:2 complex with two cyclodextrin units. This is the situation leading to the larger hydrophobicity effects. On the other hand, when the cavity size is large enough (γ cyclodextrin), the chromophore can enter the cavity and be surrounded by a few water molecules. This

appears to stabilize a di-zwitterion form (the same that is claimed to be stabilized in liquid water), giving rise to lower energy bands in the absorption spectrum of BOPH2.

Although some theoretical studies in the literature have focused on this interesting molecule, practically all of them have focused on the mechanism of proton transfer in the excited state. On the other hand, the evolution of the absorption spectrum in the presence of solvents of different natures has not been modeled in depth. In particular, the existence of di-zwitterionic (diketo) forms is just a hypothesis made to explain the spectroscopic behavior in water and this has not been demonstrated yet. In addition, so far no computational analysis has addressed the spectroscopic behavior of the molecule under confinement, with the exception of a small docking study in Reference [62].

This part of the Ph.D. project took advantage of a recently established collaboration with Dr. Giacomo Prampolini, ICCOM-CNRS, Pisa. Dr. Prampolni co-developed a rigorously and semi-automated procedure, implemented into a code named Joyce, allowing to build quantum mechanically driven force fields that can be used to describe a molecule in its ground as well as in its electronically excited state.[64, 65] Modeling BPOH2 surrounded by a molecular cavity would lead to excluding continuum models to describe the environment. A full quantum mechanical treatment of the system, on the other hand, would lead to costly computational schemes, even with hybrid QM/MM approaches, especially if converged molecular properties, such as the UV/VIS spectrum, need to be obtained. A reasonable compromise can be reached if sequential classical/QM hybrid methodologies are applied. They consist in running classical molecular dynamics simulations, from which configurations of the system of interest are extracted and used to

perform quantum chemistry calculations. The property of interest is finally derived as a convolution of the results obtained at the QM level on the different snapshots. However, if the evolution of the structure of the system during the MD simulation (taking into account temperature effects) is not reasonably predicted, some artifacts may result in the description of the desired property.[66]

In our work, the application of the Joyce protocol gave us access to a potential energy surface for BPOH2 at the QM accuracy level and to the possibility of treating molecular dynamics in the presence of the excited state. The study of the ground state that we performed with Dr. Prampolini led us to many interesting results for the ground state properties. The study on the excited state will therefore represent our perspective work.

This manuscript is organized as follows. In Chapter 1, we present the theoretical background of our work. In Chapter 2, we report the study of carboxylates in bulk water using a bottom-up approach. The molecular dynamics simulations of LDH intercalating aspartate, glutamate, and succinate are described in Chapter 3. Our study of the ground state properties of BPOH2 using Joyce and a sequential classical-QM approach is reported in Chapter 4. Finally, some general conclusions and future perspectives will be presented.

References

(1) M. I. Carretero, *Appl. Clay Sci.*, 2002, **21**, 155–163.

(2) J.-H. Choy, S.-J. Choi, J.-M. Oh and T. Park, *Appl. Clay Sci.*, 2007, **36**, 122–132.

(3) L. Desigaux, M. B. Belkacem, P. Richard, J. Cellier, P. Léone, L. Cario, F. Leroux, C. Taviot-Guého and B. Pitard, *Nano Lett.*, 2006, **6**, 199–204.

(4) Z. An, S. Lu, J. He and Y. Wang, *Langmuir*, 2009, **25**, 10704–10710.

(5) Q. Wang and D. O'Hare, *Chem. Rev.*, 2012, **112**, 4124–4155.

(6) C. Forano, T. Hibino, F. Leroux and C. Taviot-Guého, *Dev. Clay Sci.*, 2006, **1**, 1021–1095.

(7) D. G. Evans and R. C. Slade, *Layered double hydroxides*, 2006, 1–87.

(8) L. Mohapatra and K. Parida, *J. Mater. Chem. A*, 2016, **4**, 10744–10766.

(9) C. Del Hoyo, *Appl. Clay Sci.*, 2007, **36**, 103–121.

(10) G. Mishra, B. Dash and S. Pandey, *Appl. Clay Sci.*, 2018, **153**, 172–186.

(11) K.-H. Goh, T.-T. Lim and Z. Dong, *Water Res.*, 2008, **42**, 1343–1368.

(12) S. Carlino, *Solid State Ion*, 1997, **98**, 73–84.

(13) J. Zhang, F. Zhang, L. Ren, D. G. Evans and X. Duan, *Mater. Chem. Phys.*, 2004, **85**, 207–214.

(14) N. T. Whilton, P. J. Vickers and S. Mann, *J. Mater. Chemistry*, 1997, **7**, 1623–1629.

(15) H. Nakayama, N. Wada and M. Tsuhako, *Int. J. Pharm.*, 2004, **269**, 469–478.

(16) S. Aisawa, S. Takahashi, W. Ogasawara, Y. Umetsu and E. Narita, *J. Solid State Chem.*, 2001, **162**, 52–62.

(17) J. D. Bernal, *Proc. Phys. Soc. B*, 1949, **62**, 597.

(18) A. Meister, *Biochemistry of the Amino Acids.* Academic Press, 1965.

(19) V. Erastova, M. T. Degiacomi, D. G Fraser and H. C. Greenwell, *Nat. Commun*, 2017, **8**, 1–9.

(20) M.-A. Thyveetil, P. V. Coveney, H. C. Greenwell and J. L. Suter, *J. Am. Chem. Soc.*, 2008, **130**, 4742–4756.

(21) P. P. Kumar, A. G. Kalinichev and R. J. Kirkpatrick, *J. Phys. Chem. C*, 2007, **111**, 13517–13523.

(22) S. P. Newman, T. Di Cristina, P. V. Coveney and W. Jones, *Langmuir*, 2002, **18**, 2933–2939.

(23) F. Ingrosso, G. Monard, M. Hamdi Farag, A. Bastida and M. F. Ruiz-López, *J. Chem. Theory Comput.*, 2011, **7**, 1840–1849.

(24) M. H. Farag, A. Bastida, M. F. Ruiz-López, G. Monard and F. Ingrosso, *J. Phys. Chem. B*, 2014, **118**, 6186–6197.

(25) M. H. Farag, M. F. Ruiz-Lopez, A. Bastida, G. Monard and F. Ingrosso, *J. Phys. Chem. B*, 2015, **119**, 9056–9067.

(26) M. A. Soler, J. Zúñiga, A. Requena and A. Bastida, *Phys. Chem. Chem. Phys.*, 2017, **19**, 3459–3463.

(27) B. Grégoire, E. André, C. Ruby and C. Carteret, *Curr. Inorg. Chem.*, 2015, **5**, 169–183.

(28) A. G. Kalinichev, P. Padma Kumar and R. James Kirkpatrick, *Philos. Mag.*, 2010, **90**, 2475–2488.

(29) B. Grégoire, V. Erastova, D. L. Geatches, S. J. Clark, H. C. Greenwell and D. G. Fraser, *Geochim. Cosmochim. Acta*, 2016, **176**, 239–258.

(30) A. Tsukanov and S. Psakhie, *Sci. Rep.*, 2016, **6**, 1–8.

(31) B.-H. Ye, M.-L. Tong and X.-M. Chen, *Coord. Chem. Rev.*, 2005, **249**, 545–565.

(32) S. L. James, *Chem. Soc. Rev.*, 2003, **32**, 276–288.

(33) X.-X. Xie, Y.-C. Yang, B.-H. Dou, Z.-F. Li and G. Li, *Coord. Chem. Rev.*, 2020, **403**, 213100.

(34) C. Rao, S. Natarajan and R. Vaidhyanathan, *Angew. Chem. Int. Ed.*, 2004, **43**, 1466–1496.

(35) F. S. Richardson, *Chem. Rev.*, 1982, **82**, 541–552.

(36) A. Thibon and V. C. Pierre, *Anal. Bioanal. Chem.*, 2009, **394**, 107–120.

(37) L. Ackermann, *Acc. Chem. Res.*, 2014, **47**, 281–295.

(38) M. Grätzel, *J. Photochem. Photobiol. C Rev.*, 2003, **4**, 145–153.

(39) B. Das, A. Rahaman, A. Shatskiy, O. Verho, M. D. Karkas and B. Åkermark, *Acc. Chem. Res.*, 2021, **54**, 3326–3337.

(40) T. Dudev and C. Lim, *J. Phys. Chem. B*, 2004, **108**, 4546–4557.

(41) T. Dudev and C. Lim, *Acc. Chem. Res.*, 2007, **40**, 85–93.

(42) H. Dong, J. F. Snyder, K. S. Williams and J. W. Andzelm, *Biomacromolecules*, 2013, **14**, 3338–3345.

(43) A. M. Du Poset, A. Zitolo, F. Cousin, A. Assifaoui and A. Lerbret, *Phys Chem Chem Phys*, 2020, **22**, 2963–2977.

(44) A. Assifaoui, A. Lerbret, H. T. Uyen, F. Neiers, O. Chambin, C. Loupiac and F. Cousin, *Soft Matter*, 2015, **11**, 551–560.

(45) I. Imaz, M. Rubio-Martinez, J. An, I. Sole-Font, N. L. Rosi and D. Maspoch, *Chem. Comm.*, 2011, **47**, 7287–7302.

(46) C. Geffroy, A. Foissy, J. Persello and B. Cabane, *J. Colloid Interface Sci.*, 1999, **211**, 45–53.

(47) Y. Lu and J. D. Miller, *J. Colloid Interface Sci.*, 2002, **256**, 41–52.

(48) D. C. Standnes and T. Austad, *J. Pet. Sci. Eng.*, 2000, **28**, 123–143.

(49) R. Zhang, N. Qin, L. Peng, K. Tang and Z. Ye, *Appl. Surf. Sci.*, 2012, **258**, 7943–7949.

(50) J. Kubelka, S. Bai and M. Piri, *J. Phys. Chem. B*, 2021, **125**, 1293–1305.

(51) C. C. Sutton, G. V. Franks and G. da Silva, *Spectrochim Acta A Mol Biomol Spectrosc*, 2015, **134**, 535–542.

(52) B. Hernández, F. Pflüger and M. Ghomi, *J. Comp. Chem.*, 2020, **41**, 1402–1410.

(53) O. K. Abou-Zied, *J. Phys. Chem. B*, 2010, **114**, 1069–1076.

(54) O. K. Abou-Zied, *J. Photochem. Photobiol A*, 2006, **182**, 192–201.

(55) H. Bulska, *Chem. Phys. Lett.*, 1983, **98**, 398–402.

(56) J. Sepioł, A. Grabowska, H. Bulska, A. Mordziński, F. P. Salgado and R. Rettschnick, *Chem. Phys. Lett.*, 1989, **163**, 443–448.

(57) V. Barone and C. Adamo, *Chem. Phys. Lett.*, 1995, **241**, 1–6.

(58) H. Zhang, P. Van der Meulen and M. Glasbeek, *Chem. Phys. Lett.*, 1996, **253**, 97–102.

(59) D. Marks, H. Zhang, M. Glasbeek, P. Borowicz and A. Grabowska, *Chem. Phys. Lett.*, 1997, **275**, 370–376.

(60) K. Rurack, K. Hoffmann, W. Al-Soufi and U. Resch-Genger, *J. Phys. Chem. B*, 2002, **106**, 9744–9752.

(61) S. Mandal, S. Ghosh, H. H. K. Aggala, C. Banerjee, V. G. Rao and N. Sarkar, *Langmuir*, 2013, **29**, 133–143.

(62) K. Gavvala, A. Sengupta, R. K. Koninti and P. Hazra, *Phys. Chem. Chem. Phys.*, 2014, **16**, 933–939.

(63) N. Kundu, P. Banerjee, R. Dutta, S. Kundu, R. K. Saini, M. Halder and N. Sarkar, *Langmuir*, 2016, **32**, 13284–13295.

(64) I. Cacelli and G. Prampolini, *J. Chem. Theory Comput.*, 2007, **3**, 1803–1817.

(65) V. Barone, I. Cacelli, N. D. Mitri, D. Licari, S. Monti and G. Prampolini, *Phys. Chem. Chem. Phys.*, 2013, **15**, 3736–3751.

(66) O. Andreussi, I. G. Prandi, M. Campetella, G. Prampolini and B. Mennucci, *J. Chem. Theory Comput.*, 2017, **13**, 4636–4648.

Chapter 1

Theoretical Framework

Contents

1.1 Introduction . . . 18
1.2 Electronic structure calculations . . . 19
1.2.1 The Adiabatic and Born-Oppenheimer Approximations . . . 20
1.2.2 Limitations for multielectron system . . . 22
1.2.3 Sempiempirical quantum chemistry methods . . . 23
1.2.4 Density Functional Theory . . . 25
1.2.5 Continuum solvation models . . . 28
1.3 Molecular dynamics simulations . . . 30
1.3.1 Integration of Newton's equations of motion . . . 32
1.3.2 Periodic boundary conditions and thermodynamic ensembles . . . 34
1.4 Interaction energy of the system in terms of a force field . . . 35
1.4.1 Bonding Interactions . . . 37
1.4.2 Non-Bonding Interactions . . . 38
1.4.3 Cutoff schemes for non-bonded interactions . . . 39
1.5 Force Field parametrization with the Joyce protocol . . . 40
1.6 Analysis of the simulated trajectory . . . 46
1.6.1 Radial distribution functions . . . 47
1.6.2 Calculation of power spectra . . . 47
1.7 Summary . . . 50

1.1 Introduction

In this work, we focused on a molecular understanding of the environment and confinement effects on the structure, dynamics, and optical properties of two systems: organic anions within hybrid layered double hydroxides (LDH) and the bipyridine diol BPOH2 in water. In the first case, the system of interest did not undergo modifications involving bond making or breaking. In order to get some insights into available experimental measurements, we decided to address the analysis of the interlamellar space and the interpretation of the vibrational properties of the intercalated molecule. On the one hand, we needed to run reasonably long simulations to sample this complex environment and on the other hand, a description of the quantum nature of the electron behavior was not necessary and we could therefore rely on the classical dynamics with the help of an available accurate force field.

Nonetheless, in order to get deeper insights into the binding modes of the carboxylate groups, which are key for the interaction of the intercalated molecules with the inorganic layer, we applied methodologies based on quantum chemistry methods to evaluate the frequency of relevant normal modes and how they are influenced by an environment containing hydrogen bond donors.

The case of BPOH2 was quite different since electronic excited states were relevant to our study, and some intramolecular reactivity (tautomerization, proton transfer) was involved as the phenomena of interest. A fine description of electronic properties was mandatory, requiring electronic structure methods based on quantum chemistry. To sam-

ple the environment (solvent, in which the organic molecule was embedded, a fine force field development was applied through the Joyce protocol[1], capable of building parameters to reconstruct the potential energy surface at the quantum mechanical (QM) level.

The scope of this Chapter is to present a summary of the methods that were necessary to perform our work: electronic structure calculations based on density functional theory, classical molecular dynamics (MD) simulations, and force field development. In the former two cases, standard methods were applied as well as commercial codes (Gaussian16[2], Amber16[3]). Therefore, a synthetic view is presented here, in order to define the notation and discuss the reasons for our choice of methods. As for the latter, we adapted the Joyce[1] strategy to the development of the force field of the different tautomers of interest. Since this is a non-standard method, a detailed description is provided.

1.2 Electronic structure calculations

Atoms and molecules at equilibrium can be described by the time independent Schrödinger equation, written as:[4]

$$H\Psi(\mathbf{R},\mathbf{r}) = E\Psi(\mathbf{R},\mathbf{r}) \tag{1.1}$$

Here, H represents the Hamiltonian of the system and E stands for the ground state electronic energy of the system. $\Psi(\mathbf{R},\mathbf{r})$ is the wave function of the system, which is a function of the nuclear coordinates $\mathbf{R}$ and of the electron coordinates $\mathbf{r}$. The Hamiltonian

operator is defined as:

$$H = \underbrace{T_n + T_e}_{T} + \underbrace{V_{ne} + V_{ee} + V_{nn}}_{V}, \tag{1.2}$$

where the T_n and T_e are the kinetic energy operators of the nuclei and electrons, respectively. V_{ne}, V_{ee}, and V_{nn} are the potential energy operators of nuclei-electron, electron-electron, and nuclei-nuclei pairs. The analytical resolution of Equation 1.1 is impossible beyond two particles. It is thus necessary to make a certain number of approximations to be able to solve it in an approximate way.[5]

1.2.1 The Adiabatic and Born-Oppenheimer Approximations

The electron mass is much smaller than the mass of the nuclei (e.g., the ratio between the mass of a proton and the mass of an electron is about 1800), implying that nuclei have much slower motions than electrons. For a molecule at equilibrium, one can assume that the electrons move in a field created by the immobile nuclei. This allows us to split the Hamiltonian into nuclear H_n and electronic H_e parts:

$$H = H_e + H_n, \tag{1.3}$$

and to write total wave function as a product of electronic and nuclear wave functions:

$$\Psi(\mathbf{R}, \mathbf{r}) = \Psi_e(\mathbf{R}, \mathbf{r}) \cdot \Psi_n(\mathbf{R}). \tag{1.4}$$

H_e and H_n are cast in terms of their corresponding kinetic and potential energy operators:

$$H_e = T_e + V_{ne} + V_{ee}, \tag{1.5}$$

$$H_n = T_n + V_{nn}. \tag{1.6}$$

This approximation thus leads to a separable time-independent Schrödinger equation. In the Born-Oppenheimer approximation, the motion of the nuclei is considered negligible hence T_n is neglected, and the repulsion term between the nuclei can be considered as constant. The relevant surviving terms in Equation 1.2 give the so called electronic Hamiltonian:[5]

$$H_{elec} = \underbrace{T_e + V_{ne} + V_{ee}}_{H_e}. \tag{1.7}$$

The nuclear and the electronic parts are solved separately. In particular, in the Born-Oppenheimer approximation, the nuclei move on a potential energy surface constructed by solving the electronic Schrödinger equation for each set of fixed nuclear coordinates ($\mathbf{R}$ are fixed). In the coming sections, we will discuss the solution for the electronic part, which is written as follows

$$H_e \Psi_e(\mathbf{R}, \mathbf{r}) = E_e \Psi_e(\mathbf{R}, \mathbf{r}) \tag{1.8}$$

The adiabatic approximation considers that the wave function ($\Psi_e(\mathbf{R}, \mathbf{r})$) does not change its eigenstate during the process.In other words, the system does not hop from one potential energy surface to another. The Born-Oppenheimer and the adiabatic approximations are good until two or more energy states come close to each other in terms of their energies.[4]

1.2.2 Limitations for multielectron system

Schrödinger's equation does not have an exact solution for atoms (and molecules) with more than one electron.[6] To find approximated solutions for multielectron systems, the most popular strategy is to develop the total wave function, ($\Psi_e(\mathbf{R}, \mathbf{r})$), using anti-symmetric multielectronic functions with respect to the exchange of two electrons built from a product of one electron functions (ϕ_i) antisymmetrized with respect to electron exchange in which the functions ϕ_i are built as a Linear Combination of Atomic Orbitals (LCAO), based on one electron functions:[6]

$$\Psi_e(\mathbf{R}, \mathbf{r}) \approx \sum_{i=1}^{N} \phi_i(\mathbf{R}, \mathbf{r}) \approx \sum_{\nu=1}^{K} c_\nu \phi_\nu(\mathbf{R}, \mathbf{r}). \tag{1.9}$$

In a molecular system, $\Psi_e(\mathbf{R}, \mathbf{r})$ is the multielectronic wave function, and $\phi_\nu(\mathbf{R}, \mathbf{r})$ are one electron functions, centered on the atoms, representing an atomic basis set.

The development of atomic basis sets is a large and continuously evolving field in quantum chemistry but describing the different types of basis sets that have been built is beyond the scope of this Chapter. Finally, in a multielectron system, a reliable description of the wave function needs to take into account the spin quantum number.[6] However, an explicit treatment of the spin was not necessary in this work, since the molecules of interest are closed shell systems.

To solve the electronic part of the Schrödinger equation (Equation 1.8), different methods have been developed, each corresponding to a different physical model. One ab initio electronic structure strategy based on a wave function approach is the Hartree-Fock

(HF) method, based on a variational framework and on a self-consistent field approach. According to this methodology, each electron is assumed to interact with an average potential exerted by the other electrons (V_{ee}), so that the instantaneous position of an electron is not affected by the presence of a neighboring electron. In reality, the motions of electrons are correlated and they tend to 'avoid' each other more than Hartree-Fock theory would suggest, giving rise to lower energy. For this reason, post Hartree-Fock approaches are developed, such as perturbation methods, as the Møller-Plesset theory, coupled cluster theory, and multireference methods, the latter being particularly useful to describe excited states.[5, 6] On the other hand, alternative, not necessarily ab initio, methodologies can be envisaged. In this work, we performed quantum chemistry calculations using a Density Functional Theory level with hybrid functionals as well as semiempirical methods based on a wave function approach. These two methodologies will be briefly presented in the following.

1.2.3 Sempiempirical quantum chemistry methods

Starting with the work of Hückel, then of Pariser, Parr, and Pople, mainly addressing π systems, the quantum chemistry method using parametrized Hamiltonian quickly evolved towards different possible approximations. Such approximations treated differently the overlap between basis functions in the expression of the self-consistent field terms in a Hartree-Fock like approach: zero-differential overlap (ZDO), complete neglect of differential overlap (CNDO), neglect of diatomic differential overlap (NDDO) are the main proposed approximations. In the 1980s, Stewart proposed an improved version of

the MNDO code by Dewar and collaborators, originally meant to accurately reproduce the properties of organic molecules in their ground state (the training set of the first version of MNDO was formed by 138 small molecules). This was called the PM3 method, which has continued to be improved until the more recent PM6 and PM7 versions, eventually eliminating some of the artifacts of PM3 for describing crystal structures and for the description of hydrogen bonds. For a complete presentation of these methods, see [7] and references therein.

In these methods, Slater Type Orbitals (STO) are used as atomic basis sets, and some assumptions are made on the calculations of the integrals necessary to iteratively solve the Hartree-Fock equations (numerical solution of the Schrödinger equation) with the STO basis set, through the definition of the so called Fock matrix. Compared with the HF ab initio method, semiempirical methods such as the PMX (PM3-PM7) family simplify the expression of the matrix elements containing two-electron terms and some one-electron terms to make the iterative process faster. The corresponding terms are replaced by ad hoc chosen analytical functions depending on parameters, which are optimized to describe the experimental properties (ionization potentials, geometries, dipole moments, heats of formation) of a large number of molecules forming a training set. Compared with the MNDO method, the PMX methods are based on a more sophisticated optimization scheme.

1.2.4 Density Functional Theory

The central idea behind Density Functional Theory (DFT) is the relationship between the total electronic energy and the overall electronic density. In 1964, Hohenberg and Kohn showed the one-to-one correspondence between the ground state electron density distribution ($\rho(\mathbf{r})$) of a many electron system and the (ground state) electron energy E_0.[8] According to the Hohenberg and Kohn theorem, E_0 is a functional of ρ:

$$E_0 = E_e[\rho] \tag{1.10}$$

If V_{ne} is the term describing the interaction between electrons and nuclei, one can express the ground state electron energy in terms of a functional F of the density:

$$E_e[\rho] = V_{ne}[\rho] + F[\rho], \tag{1.11}$$

where the potential energy term is expressed in terms of Coulombic interaction:

$$V_{ne}[\rho] = \int \rho(\mathbf{r}) v_e(\mathbf{r}) d\mathbf{r} \tag{1.12}$$

Here, $v_e(\mathbf{r})(= -\frac{Z_A}{|\mathbf{r}-\mathbf{R}|})$ is called the external potential energy function. Despite the fact that V_{ne} is known, the functional $F[\rho]$ is unknown. Using the variational principle, Hohenberg and Kohn proved the dependency of $F[\rho]$ on a universal exchange-correlation (EC) energy functional $E_{XC}[\rho]$, electron-electron Coulombic energy term ($V_{ee}[\rho]$), and the kinetic energy $T_e[\rho]$ but they did not provide an explicit definition of all these terms.

The formulation of DFT was further improved by Kohn and Sham[9]. Their landmark paper suggested a practical way to solve the Hohenberg-Kohn theorem for a set of interacting electrons. They explicitly formulated the $F[\rho]$ in Equation 1.11 as the sum of a kinetic energy term ($T_e[\rho]$), an electron-electron Coulombic energy term ($V_{ee}[\rho]$), and an exchange-correlation energy term (E_{XC}):

$$F[\rho] = T_e[\rho] + V_{ee}[\rho] + E_{XC}[\rho]. \tag{1.13}$$

The ground state electronic energy can finally be written as:

$$E_e[\rho] = V_{ne}[\rho] + T_e[\rho] + V_{ee}[\rho] + E_{XC}[\rho]. \tag{1.14}$$

V_{ee}, the Coulombic interaction between electrons, can be expressed in terms of density $\rho(\mathbf{r})$ as:

$$V_{ee}[\rho] = \frac{1}{2}\int\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \tag{1.15}$$

As a consequence, after incorporating all the terms the final form of the energy of the N-electron system would be:

$$\begin{aligned} E_e[\rho] = &\sum_{i=1}^{N}\int \phi_i(\mathbf{r})\left(-\frac{\nabla^2}{2}\right)\phi_i^*(\mathbf{r})d\mathbf{r} + \frac{1}{2}\int\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{XC}[\rho] \\ &- \sum_{i=1}^{N}\int \frac{Z_A}{|\mathbf{r}-\mathbf{R}|}\rho(\mathbf{r})d\mathbf{r} \end{aligned} \tag{1.16}$$

Kohn and Sham expressed the density ($\rho(\mathbf{r})$) of the system as the sum of the square

moduli of a set of one-electron orthonormal orbitals:

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2 \tag{1.17}$$

By applying variational condition and using the above definition of the electron density, the formulation of DFT was further ameliorated by adding its orbital formulation[9] and presented as the well known one-electron Kohn-Sham equations:

$$\left\{ -\frac{1}{2}\nabla^2 - \left(\sum_{A=1}^{M} \frac{Z_A}{\mathbf{r}_{1A}} \right) + \int \frac{\rho(\mathbf{r}_2)}{\mathbf{r}_{12}} d\mathbf{r}_2 + V_{XC}[\mathbf{r}_1] \right\} \phi_i(\mathbf{r}_1) = \epsilon_i \phi_i(\mathbf{r}_1), \tag{1.18}$$

where Z_A is the atomic number of nucleus A, M is the total number of nuclei, ϵ_i is the orbital energy of the i^{th} orbital represented by the basis function $\phi_i(\mathbf{r}_1)$. $V_{XC}[\mathbf{r}_1]$ is known as the exchange-correlation functional, and it is related to the exchange-correlation energy as follows:

$$V_{XC}[\mathbf{r}] = \left(\frac{\delta E_{XC}[\rho]}{\delta \rho(\mathbf{r})} \right) \tag{1.19}$$

The other two terms being in principle exact, the only approximation in DFT applied to molecular systems would be represented by the search for a suitable functional to describe the exchange-correlation term. Increasing levels of improvement in the search for the best functional are usually referred to as Jacob's ladder in DFT.[6] In the Local Density Approximation (LDA), one follows the assumption that the electron density is the same as that provided by a homogeneous gas of electrons at each point within the molecule. The Generalized Gradient Approximation (GGA) is based on providing an exchange-correlation functional as well as its gradient (gradient corrected functionals:

B88, LYP, P86, PBE, etc.). It is possible to combine different approaches by mixing the exchange term from one functional and the correlation term of another (e.g. B88-LYP). In the adiabatic approximation, the exchange-correlation term can be expressed as a linear combination of an EC DFT term and the Hartree-Fock exchange term.[7] Different levels of theory are used to optimize such coefficients, giving rise to different hybrid functionals (B3LYP, B3PW91, M0X, etc.). There is no clear rule to decide which functional to use in a quantum chemistry calculation. Experience can help in either performing a benchmark study or in seeking justification for a given choice through similar work already published in the literature.

1.2.5 Continuum solvation models

The possibility to include the effect of the environment on the properties of molecules modeled through quantum chemistry calculations has allowed significant steps forward in the description of chemical systems. As a matter of fact, continuum solvation models are among the most efficient ones to include condensed phase effects, since they lead to reasonable computational times compared with those method requiring that a large number of explicit solvent molecules is included in the calculation, as in ab initio or DFT-based molecular dynamics, or in hybrid quantum mechanics/molecular mechanics schemes. The latter methodologies also require special techniques to sample the conformational space, given the very large number of degrees of freedom. On the other hand, continuum methods focus on the description of the chemical system of interest (thus a very limited portion of the overall system) replacing the structure of the solvent by a continuum bath having

an effect on the wave function of the solvated molecule(s).

In this work, we performed quantum chemistry calculations using two different solvation methods and compared the results obtained in the two cases, in Chapter 2 (application to hydrated carboxylates). In addition, the quantum chemistry calculations performed in Chapter 4 to develop the intramolecular force field of a bipyridine diol also included solvent effect through a continuum model. We, therefore, provide a short description of these different solvation schemes.

Polarizable Continuum Model, PCM[10] The basic idea of this solvation model is to replace the solvent surrounding a molecular solute with a dielectric medium characterized by its dielectric constant. The solute is embedded in a cavity, built from interlocking spheres centered on the atoms, in a way that it reproduces reasonably well the molecular shape. The interaction between solute and solvent is treated as a quantum mechanical operator that represents the perturbation included by the solvent on the gas phase Hamiltonian of the solute. Although some developments exist to treat nonelectrostatic interactions, in the basic versions of PCM only the electrostatic term is taken into account.

The charge distribution of the solute polarizes the dielectric medium in return. The solution of the electrostatic problem needs a self-consistent iterative procedure. Finally, the interaction between the solute charge distribution and the electric polarization field of the solvent gives access to the computation of the electrostatic contribution to the solvation free energy.

The electrostatic problem is solved by defining apparent surface charges on the cavity, resulting from the discretization of the surface charge induced by the presence of the

solute as a response to the dielectric medium. The **CPCM**[11, 12] version of the polarizable continuum model, used in Chapter 4, is based on the COSMO (Conductor Like Screening Model) method,[13] leading to a simplification of the boundary conditions of the electrostatic problem. Such approximation is compensated by the use of adjustable parameters in the expression of the solvation energies, tuned on experimental values. In the **IEFPCM** (Integral Equation Formalism PCM),[14, 15] used in Chapter 2, the electrostatic potential is developed in terms of the proper Green function. The solution of the electrostatic problem is then given based on the theory of integral equations.

Solvation Model Density, SMD[16] As in PCM, the solute is immersed in a dielectric medium that polarizes the solute and is polarized by it, and the electrostatic problem is solvent through the definition of a surface charge on the solute cavity. However, instead of using a discretization of the surface charge, the SMD model uses a polarized continuous quantum mechanical charge density of the solute. In addition, the geometrical parameters used to define the cavity have been optimized to reproduce experimental standard free energies of solvation of molecules belonging to a wide training set.

1.3 Molecular dynamics simulations

Computational methods based on a quantum mechanical Hamiltonian are too costly to be applied when a large system and/or a large number of atoms need to be taken into account. In addition, a much heavier formalism has to be included for time-dependent phenomena, a subject that is developed in the field of quantum dynamics. In situations

in which the Born Oppenheimer approximation holds, classical statistical mechanics can be used to retrieve macroscopic properties of a system of molecules from ensemble averages along molecular dynamics (MD) trajectories that evolve in time, based on Newton's equations of motion:

$$\frac{d^2\mathbf{R}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i} \tag{1.20}$$

Equation 1.20 describes the motion of a particle of mass m_i along a displacement vector $\mathbf{R}_i$ with force $\mathbf{F}_i$ in the direction of vector $\mathbf{R}_i$.[6] If we use an atomistic representation of the system, the particles are the atoms composing it. The forces $\mathbf{F}_i$ exerted on an atom i with a position vector $\mathbf{R}_i$ at time t can be determined by the derivative of the total potential energy function $E_{tot}(\mathbf{R})$.

$$\mathbf{F}_i = -\frac{\partial E_{tot}(\mathbf{R})}{\partial \mathbf{R}_i} \tag{1.21}$$

therefore, the velocity $\mathbf{v}(\mathbf{R}_i)$ and positions $\mathbf{R}_i$ of a particle i at time t can be written in terms of the total potential energy $E_{tot}(\mathbf{R})$ as follows:

$$\frac{d\mathbf{v}(\mathbf{R}_i)}{dt} = -\frac{1}{m}\frac{\partial E_{tot}(\mathbf{R})}{\partial \mathbf{R}_i} \tag{1.22}$$

$$\frac{d^2\mathbf{R}_i}{dt^2} = -\frac{1}{m}\frac{\partial E_{tot}(\mathbf{R})}{\partial \mathbf{R}_i} = \frac{F_i}{m} \tag{1.23}$$

The goal here would be to determine the coordinates $\mathbf{R}_i(t+\delta t)$ of a particle at the instant $t+\delta t$ knowing its acceleration $\mathbf{a}_i(t) = \frac{d^2\mathbf{R}_i}{dt^2}$ at the previous instant t that we calculated knowing the interaction forces $\mathbf{F}_i(t)$.

1.3.1 Integration of Newton's equations of motion

To study a system of atoms using MD, one needs to calculate the force acting on each atom of the system ($\mathbf{F}_i(t)$ is the force acting on atom i) from the potential $V(\mathbf{R})$ at time t using Equation 1.21. Once the forces are calculated, Newton's equations of motion are integrated to obtain velocities ($\mathbf{v}_i$) and positions ($\mathbf{R}_i$) at time $t + \delta t$. Two different types of integration methods can be used, referred to as Finite Difference Methods and Predictor-Corrector Integration.

Finite difference methods are used to generate MD trajectories with an analytical expression of continuum model potentials. The total force on each particle at time t is calculated as the vector sum of its interactions with other particles. When the integration of the equations of motion is done using finite difference methods, one assumes that the positions and dynamic properties (velocities, accelerations, etc.) can be approximated as Taylor series expansion:

$$\mathbf{R}_i(t + \delta t) = \mathbf{R}_i(t) + \delta t \mathbf{v}_i(t) + \frac{1}{2}\delta t^2 \mathbf{a}_i(t) + ... \tag{1.24}$$

where $\mathbf{v}_i$ and $\mathbf{a}_i$ is the velocity and acceleration, respectively.

The Verlet integration algorithm uses the positions $\mathbf{R}_i(t)$, $\mathbf{R}_i(t-\delta t)$, and accelerations $\mathbf{a}_i(t)$ to calculate new positions $\mathbf{R}_i(t + \delta t)$.[17] The calculations of $\mathbf{R}_i(t + \delta t)$ can be done

as following:

$$\mathbf{R}_i(t+\delta t) = \mathbf{R}_i(t) + \delta t \mathbf{v}_i(t) + \frac{1}{2}\delta t^2 \mathbf{a}_i(t) + ... \tag{1.25}$$

$$\mathbf{R}_i(t-\delta t) = \mathbf{R}_i(t) - \delta t \mathbf{v}_i(t) + \frac{1}{2}\delta t^2 \mathbf{a}_i(t) - ... \tag{1.26}$$

After adding these two equations and neglecting the higher order terms will give

$$\mathbf{R}_i(t+\delta t) = 2\mathbf{R}_i(t) - \mathbf{R}_i(t-\delta t) + \delta t^2 \mathbf{a}_i(t) \tag{1.27}$$

The velocities can be calculated from here in many ways. One of the methods includes taking the difference of equations 1.25 and 1.26 and divide it by $2\delta t$:

$$\mathbf{v}_i(t) = \frac{[\mathbf{R}_i(t+\delta t) - \mathbf{R}_i(t-\delta t)]}{2\delta t} \tag{1.28}$$

The implementation of the Verlet algorithm is straightforward and it is not memory demanding.

The predictor-corrector methods[18] come from a larger family of integration algorithms. In MD simulation codes, this method is implemented in three basic steps:

- New positions, velocities, accelerations, and higher-order terms are predicted according to the Taylor expansion in Equation 1.24
- The forces are evaluated at the new positions to give $\mathbf{a}_i(t+\delta t)$
- These accelerations are then compared with the accelerations that are predicted

from the Taylor series expansion, $\mathbf{a}_i^c(t + \delta t)$

The difference between the predicted (Taylor expansion) and calculated accelerations is then used to correct the positions, velocities, etc., in the correction step.

1.3.2 Periodic boundary conditions and thermodynamic ensembles

The limitations dictated by the computational resources available nowadays are such that, even in the case of the largest MD simulation performed so far,[19] comprising 1.6 billion atoms, the number of particles that can be explicitly modeled with this approach stays much smaller than Avogadro's number. To deal with this limitation, a representative box of the model system is considered, which is replicated throughout the space to form an infinite lattice similar to a three-dimensional crystal. The replication of such a box will create periodic copies of the particles, giving rise to a framework called Periodic Boundary Conditions (PBC).[20]

If a particle within the box moves during the simulation, its images move within every periodic box with the same magnitude as if mimicking the original motion. Hence, considering a particle that leaves the central box, it will be replaced by its image. PBCs are implemented using the minimum image convention: the minimum distance between two particles is estimated, and the particles on the edges of the box are allowed to interact with a mimicking particle from the neighboring cell.[20] In order to run MD simulations under thermodynamic conditions that are directly comparable with realistic situations, they

can be run in thermodynamic ensembles such as the microcanonical (constant number of particles N, volume V, total energy E), canonical (constant N, V, temperature T), grand canonical (constant chemical potential μ, V, T), and Isothermal-isobaric (constant N, P, T) ensembles. In the case of constant temperature, common approaches include (1) modifying the velocities (e.g. weak-coupling), (2) introducing fictitious particles in an extended system (e.g. Nosé-Hoover), or (3) introducing friction (e.g. Langevin dynamics). As in the case of temperature control, constant pressure can be modeled by a barostat that modifies the volume of the system using (1) scaling the box dimensions (e.g. weak-coupling), (2) introducing fictitious particles (e.g. Parinello-Rahman), or (3) introducing a piston. Some thermostats and barostats are better suited for systems far from equilibrium, whereas others are better for production simulations.[21]

1.4 Interaction energy of the system in terms of a force field

In this section, we will present the analytical expression used to describe the different terms of potential energy. We have seen that the first step to define Newton's equations of motion for MD simulations is the knowledge of the forces acting on each atom by performing a derivative of the potential energy (Equation 1.21). It is therefore necessary to model each atom-atom interaction in terms of functional forms deriving from a physical model. The easiest way to think of it is to separate intramolecular forces (to be described in terms of intramolecular motions) with respect to intermolecular forces (most of the

time dominated by Coulombic and van der Waals interactions), in a classical framework. As a matter of fact, even though the functional forms for such terms can be the same, the parameters describing them strongly depend on the specific system considered. The set of parameters developed to describe one particular system is referred to as the force field (FF). Appropriately developed force fields are validated against the structure and the interaction energies for a set of model compounds.

In the field of molecular dynamics, the expression of total potential energy E_{tot} of a system is decomposed into the so called bonding and non-bonding components, representing intramolecular and intermolecular interactions, respectively:

$$E_{tot} = E_{bonded} + E_{non-bonded} \tag{1.29}$$

E_{bonded} and $E_{non-bonded}$ can be further decomposed into several components as following:

$$E_{bonded} = E_{str} + E_{bend} + E_{dih} \tag{1.30}$$

$$E_{non-bonded} = E_{elec} + E_{vw} \tag{1.31}$$

E_{str}, E_{bend}, and E_{dih} represent the potential energies of stretching, bending, and torsion motions, respectively. E_{elec} and E_{vw} represent the potential energy for electrostatic and van der Waals interactions. Such partitions, and their functional forms, will be described in the following sections. We stress here that only a non-polarizable FF will be described in this work, though important developments in MD also provide polarizable FF, taking into account polarization terms in the intermolecular electrostatic interactions. Such FF

lead to an increased simulation time but they become necessary for those systems in which polarizability plays a major role. In our case, especially for the work on layered double hydroxides, Coulombic interactions played the most important role as well as van der Waals interactions at short range, and only a non-polarizable FF was employed.

1.4.1 Bonding Interactions

Recalling the partition of the bonding terms defined in Equation 1.30, we first focus on the bond stretching term. This term is generally represented by a harmonic function:

$$E_{str} = \frac{1}{2} k_{str} (l - l_0)^2 \tag{1.32}$$

where k_{str} is the force constant, l is the instantaneous bond length, and l_0 is the equilibrium bond length.

The bending of an angle between three atoms is also assumed as a harmonic motion and its potential is written as:

$$E_{bend} = \frac{1}{2} k_{bend} (\theta - \theta_0)^2 \tag{1.33}$$

where k_{bend} is the force constant, θ is the instantaneous bond angle, and θ_0 is the equilibrium bond angle.

The torsion potential of dihedral angles is further split into two classes because of the nature of the torsional motion. If the dihedral angle is stiff (Improper dihedral) then

it will be treated with a harmonic function:

$$E_{sdihed} = \frac{1}{2} k_{sdihed} (\phi - \phi_0)^2 \tag{1.34}$$

where k_{sdihed} is the force constant, ϕ is the instantaneous dihedral angle, and ϕ_0 is the equilibrium dihedral angle. For the flexible (proper) dihedrals, periodic functions are used to represent the energies:

$$E_{fdihed} = \frac{1}{2} \sum_n k_n [1 + \cos(n\phi + \phi_0)] \tag{1.35}$$

where ϕ is the instantaneous dihedral angle between the atoms, ϕ_0 is the phase shift angle and k_n is the multiplicative constant with n being a non-negative integer constant that indicates the periodicity (often called multiplicity).[20]

1.4.2 Non-Bonding Interactions

When focusing on Coulombic and van der Waals (intermolecular) interaction energy, the non-bonding interactions comprise the following terms:

$$E_{elec} = \frac{1}{4\pi\epsilon_0} \sum_{i \neq j} \frac{q_i q_j}{r_{ij}} \tag{1.36}$$

is a Coulombic term, where ϵ_0 is the permittivity of free space (vacuum), q_i and q_j are atomic charges of atoms i and j, and r_{ij} is the distance between the atoms i and j, and

$$E_{vw} = \sum_{i \neq j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \tag{1.37}$$

is a Lennard-Jones potential form, where r_{ij} gives the distance between the pair of atoms i and j. ϵ_{ij} is the van der Waals well depth for atoms i and j, and σ_{ij} is the distance at which $E_{vw} = 0$.

1.4.3 Cutoff schemes for non-bonded interactions

Cutoff schemes are used to reduce the cost of calculating non-bonded interactions. In the case of van der Waals interactions, which decrease rapidly to zero as the distance between two molecules increases, it is usually computed using the terms corresponding to Equation 1.37 within one-half of the simulation box. On the other hand, larger cutoffs are necessary to treat long-range interactions such as those described by the Coulombic term. Some of the most used cutoff schemes are applying a spherical cutoff, the Ewald and the particle-mesh Ewald (PME) methods, and fast multipole schemes.[20]

The spherical cutoff schemes are composed of three basic categories of cutoff techniques: truncation, switch, and shift formulations. **Truncation** abruptly defines the values of the energy to be zero after a certain limit b and it does not alter the values of the energy and forces for the distances $r < b$ (similarly to what is done to treat the van der Waals term). **Switching** switches the interaction term to a nonzero value at distances

$a < r < b$ but leaves the values unchanged at distances $r < a$. **Shift** alters the function more gradually for all distances $r < b$.[7]

In the approaches based on the Ewald sum method, the electrostatic energy is split into short-range electrostatic (E_{elec}^{SR}) and long-range electrostatic (E_{elec}^{LR}) energy terms:

$$E_{elec} = E_{elec}^{SR} + E_{elec}^{LR} \tag{1.38}$$

The short range electrostatic remains the same as in Equation 1.36, the long range part is a function of Fourier transform of E_{elec}^{SR} and the charge density ρ.[20]

1.5 Force Field parametrization with the Joyce protocol

In this work, some FF already available in the literature were used to model layered double hydroxides, water, and organic molecules. In the case of the bipyridine diol, a complete study would necessitate simulations of solvation and of confinement with the molecule in the ground as well as in the excited state. In addition, more than one tautomers are present in water, leading to the necessity of ad hoc FF capable of finely describing the potential energy surface of the molecule. To this end, we have taken advantage of a collaboration established in the laboratory with Dr. G. Prampolini, who co-developed a few years ago the Joyce protocol.[1] This protocol allows us to obtain FF parameters from first principles and to obtain a perfect match between the QM and the

MM potential energy surfaces.

If we consider a system that contains a solute immersed in a solvent, then Equations 1.29-1.31 can be decomposed further to take into account the interactions of solute with its environment. From now on, we shall thus focus on a solute (*solu*) plus solvent (*solv*) system:

$$E_{tot} = E_{inter}^{solv-solv} + E_{inter}^{solu-solv} + E_{intra}^{solu} + E_{intra}^{solv}, \tag{1.39}$$

where $E_{inter}^{solv-solv}$ and $E_{inter}^{solu-solv}$ are the interaction energies of intermolecular solvent-solvent and solute-solvent interactions, respectively. E_{intra}^{solu} and E_{intra}^{solv} are the intramolecular interaction energies of solute and solvent, respectively. For the sake of simplicity, the superscript labels of Equation 1.39 will be dropped, leading to a simplified version of Equation 1.39:

$$E_{tot}^{simple} = E_{inter} + E_{intra}, \tag{1.40}$$

E_{inter} is computed as a sum of Lennard-Jones and Coulombic pair interaction energies between i, j components of the system:

$$E_{inter} = E_{LJ} + E_{Coul}. \tag{1.41}$$

In this notation, the Coulombic and the Lennard-Jones term are written as

$$E_{Coul} = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N_{pairs}} \sum_{j=1}^{N_{pairs}} \frac{q_i q_j}{\mathbf{r}_{ij}}, \tag{1.42}$$

$$E_{LJ} = \sum_{i=1}^{N_{pairs}} \sum_{j=1}^{N_{pairs}} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{\mathbf{r}_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{\mathbf{r}_{ij}} \right)^{6} \right] \tag{1.43}$$

where i and j are atoms belonging to a pair of different molecules. The total intramolecular term E_{intra} is the sum of the internal energies of all the M molecules composing the system:

$$E_{intra} = \sum_{K=1}^{M} V_K^{intra}. \tag{1.44}$$

The intramolecular potential V_K^{intra} can be written as a function of a set of generalized nuclear internal coordinates describing the target molecule K. The intramolecular force field is expressed as a sum of different terms as

$$V^{intra} = E_{stretch} + E_{bend} + E_{Stors} + E_{Ftors} + E_{nb}^{intra} + E_{Coupl}, \tag{1.45}$$

where the first three terms $E_{stretch}$, E_{bend}, and E_{Stors} correspond to the so called "hard" internal coordinates, bond stretching, angular bending, and the so called "stiff" dihedral angles, respectively. We recall that these potentials are modeled as harmonic potentials:

$$E_{stretch} = \frac{1}{2} \sum_{\mu}^{N_{bonds}} k_\mu^s (l_\mu - l_\mu^0)^2, \tag{1.46}$$

where N_{bonds} is the total number of bonds, k_μ^s is the force constant of the bond μ, and l_μ^0 is the equilibrium bond length;

$$E_{bend} = \frac{1}{2} \sum_{\mu}^{N_{angles}} k_\mu^b (\theta_\mu - \theta_\mu^0)^2, \tag{1.47}$$

where N_{angles} is the total number of angles, k^b_μ is the force constant of the angle μ, and θ^0_μ is the equilibrium bond angle;

$$E_{Stors} = \frac{1}{2}\sum_{\mu}^{N_{Sdihed}} k^t_\mu(\phi_\mu - \phi^0_\mu)^2 \tag{1.48}$$

where N_{Sdihed} is the total number of stiff dihedral angles, k^t_μ is the force constant of the angle μ, and ϕ^0_μ is the equilibrium dihedral angle.

As was already discussed, model functions used for flexible (also named "soft:) dihedrals are expressed as the combination of periodic functions as

$$E_{Ftors} = \sum_{\mu}^{N_{Fdihed}} \sum_{j=1}^{N_{cos_\mu}} k^d_{j\mu}[1 + cos(n^\mu_j \delta_\mu - \gamma^\mu_j)], \tag{1.49}$$

where N_{cos_μ} is the number of cosine functions employed to describe the potential of the δ_μ dihedral.

Finally, if needed the intramolecular non-bonded contribution is computed using Lennard Jones terms:

$$E^{intra}_{nb} = \underbrace{\sum_{i=1}^{N_{nb}} \sum_{j=i+1}^{N_{nb}} 4\epsilon^{intra}_{ij}\left[\left(\frac{\sigma^{intra}_{ij}}{\mathbf{r}_{ij}}\right)^{12} - \left(\frac{\sigma^{intra}_{ij}}{\mathbf{r}_{ij}}\right)^{6}\right]}_{Intramolecular\ LJ} + \underbrace{\sum_{i=1}^{N_{nb}} \sum_{j=i+1}^{N_{nb}} \frac{q_i q_j}{\mathbf{r}_{ij}}}_{Intramolecular\ Coulombic} \tag{1.50}$$

where the i and j are the indexes of atoms of the same molecule, therefore $j \neq i$ to avoid self interaction. The ϵ^{intra}_{ij} and σ^{intra}_{ij} may not be the same employed in the Equation1.43. This term is particularly relevant for large, flexible molecules exploring a wide conformational space, leading to different parts of the molecule coming close to contact during the

evolution of molecular motions.

The Joyce protocol provides a semi-automated way to parameterize the intramolecular part of the force field, V^{intra}, defined in Equation 1.45. This is done by using the energy and its first and second derivatives (gradient and Hessian) obtained through quantum chemistry electronic structure calculations. In order to reduce the number of degrees of freedom in the process that leads to such a fine optimization of the FF parameters, V_g^{intra} can also be expressed in terms of a linear combination of functions f_a, which are defined of a set of redundant internal coordinates (RICs):

$$V^{intra}(q) = V(q) = \sum_{a=1}^{N_{func}} p_a f_a(q) \tag{1.51}$$

where symbol q contains all the RICs and p_a is the normalization coefficient for each RIC. The best force field parameters, which can be used to represent the internal molecular motion, are obtained by minimizing the following merit function I, which is written as a sum over the considered molecular geometries I_g.

$$I = \sum_{g=0}^{N_g} I_g \tag{1.52}$$

where I_g is written as follows:

$$\begin{aligned} I_g = W_g[(E_g - E_0) - V_g]^2 + \sum_{K=1}^{3N-6} \frac{W'_{Kg}}{3N-6}[E'_{Kg} - V'_{Kg}]^2 \\ + \sum_{K\leq L}^{3N-6} \frac{2W^{"}_{KLg}}{(3N-6)(3N-5)}[E^{"}_{KLg} - V^{"}_{KLg}]^2 \end{aligned} \tag{1.53}$$

Here, I_g is a merit function for the instantaneous geometry g and it depends on the following terms: the intramolecular potential energy of the geometry V_g, its gradient V_g' and $V_g^{"}$, the total energy of the geometry E_g, the energy of the reference geometry E_0 and the energy gradient (Hessian E_g' and $E_g^{"}$). All those quantities are computed by QM calculations. The indices K and L run over the normal coordinates and include all the normal modes except the rotational and translational ones. W, W', and W" are the weights of energy and energy gradient terms at each geometry.

An accurate description of large amplitude torsional movements needs to take into account the fraction of the torsional energy coupling with the changes of the other internal coordinates. This information, which is contained in a PES determined through a QM calculation, is lost if the intramolecular FF is developed exclusively by using independent internal coordinates (e.g. normal modes) since the motion of one internal coordinate is independent of the motions of the others. A straightforward solution for this problem would require the inclusion of a relevant number of coupling functions as shown in Equation 1.45.

A more simple and direct solution is to ignore the coupling with those internal coordinates that are not directly involved in the internal rotation and perform a QM relaxed scan, each point of the relaxed scan being assigned to a virtual geometry corresponding to a rigid rotation of the scanned dihedral. This route, named Frozen Internal Rotation Approximation (FIRA), has the effect of attributing the torsional energy to the torsional term only (see Equation 1.49), whereas, in the QM calculations, it is distributed on several coupled internal coordinates. It might be worth mentioning that if a few specific internal coordinates undergo a significant variation upon internal rotation, the coupling between

them and the scanned dihedral can be taken into account either explicitly, through the proper coupling function, or implicitly, with the use of the intramolecular LJ functions.

To resume the Joyce parametrization protocol, we present a flow chart in Figure 1.1.

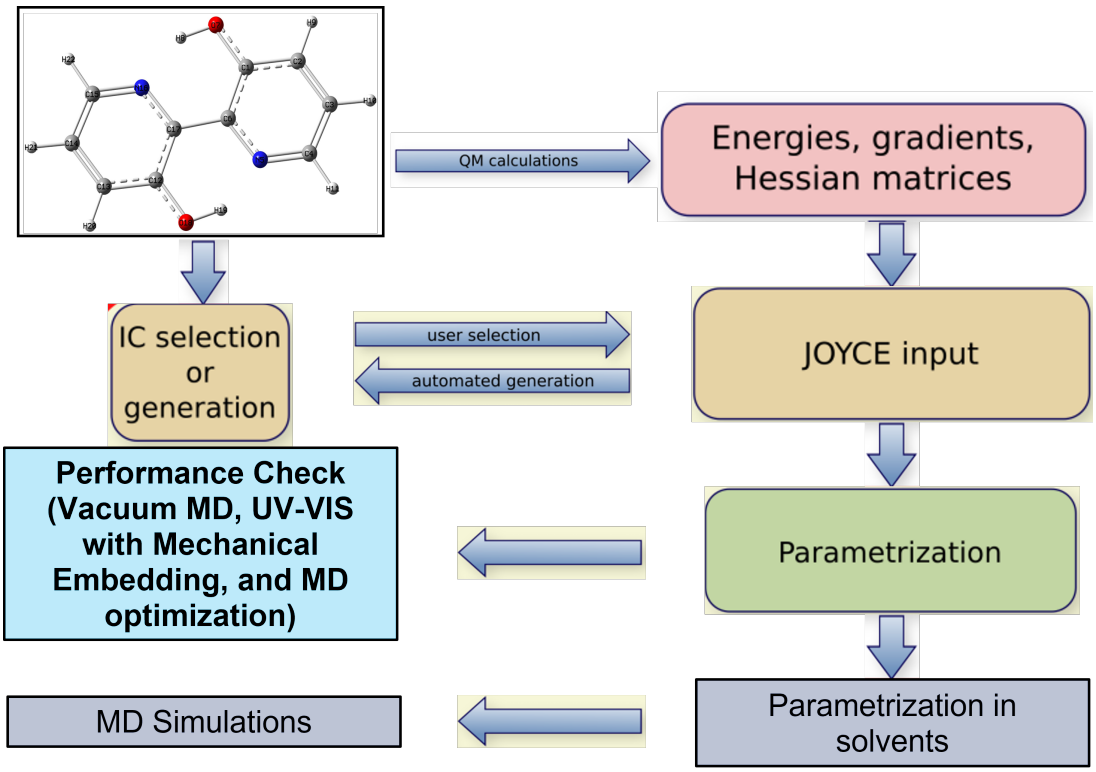


Figure 1.1: Flow chart of the Joyce protocol.[1]

1.6 Analysis of the simulated trajectory

For all the systems considered in this work, we have carried out a thorough study of the local structure described in terms of atom-atom pair correlation functions, characterizing the main intermolecular interactions (hydrogen bonds, hydrophobic/hydrophilic interactions). In addition, in Chapters 2 and 3, we carried out an analysis of the vi-

brational properties of molecules surrounded by different environments by estimating the infrared absorption spectrum. The theoretical background used to compute these two properties will therefore be resumed in the following Sections.

1.6.1 Radial distribution functions

In a homogeneous system, radial distribution functions describe the ratio between the average local particle density at a distance r from one target particle and the bulk density.[6] The probability of finding one particle at a distance r from the reference particle is given by the expression:

$$P(r) = 4\pi r^2 g(r) dr \tag{1.54}$$

where $g(r)$ is the RDF.

The integration of the first peak of the RDFs up to the distance corresponding to the position of the first minimum $\mathbf{r}_{min}$ provides a definition of the coordination number, $n_{\mathbf{r}_{min}}$ as:

$$n_{\mathbf{r}_{min}} = 4\pi\rho \int_0^{\mathbf{r}_{min}} g(\mathbf{r})\mathbf{r}^2 d\mathbf{r} \tag{1.55}$$

1.6.2 Calculation of power spectra

In this work, two methods have been employed to estimate the infrared absorption spectrum of the molecules of interest, one based on the recorded velocities along MD trajectories and the other based on the computation of the time evolution of the molecular dipole moment.

The velocity density of states (VDOS) is computed from the atomic velocities time correlation function (TCF).[22] Its power spectrum $P_v(\omega)$ is defined in terms of a Fourier transform:

$$P_v(\omega) = Q_v(\omega) \sum_i^N \int_{-\infty}^{+\infty} \langle v_i(t) \cdot v_i(0) \rangle e^{i\omega t} dt, \tag{1.56}$$

where $Q_v(\omega)$ is a quantum correction, $v_i(t)$ is the velocity of atom i, N is the total number of atoms, ω is the frequency and $\langle v_i(t) \cdot v_i(0) \rangle$ is the velocity autocorrelation function. This spectrum contains information on the vibrational motions of the molecule.

The infrared absorption spectrum of a molecule can also be evaluated based on MD trajectories by using a formulation stemming from Kubo's relationship[23]. According to this development, the infrared intensity can be expressed in terms of the Fourier transform of the molecular dipole time correlation function as:

$$I(\omega) = \frac{1}{2\pi} Q(\omega) \int_{-\infty}^{+\infty} dt \; exp(-i\omega t) < \boldsymbol{\mu}(0) \cdot \boldsymbol{\mu}(t) > \tag{1.57}$$

in which $Q(\omega)$ is a prefactor taking into account the quantum nature of the time correlation function.[24, 25] An explicit treatment of this factor was not included in our calculation since this goes beyond the purposes of our analysis.

We made use of the Wiener-Khinchine (WK) method, with a Fast Fourier Transform (FFT) algorithm, to compute the power spectra (VDOS, dipole TCF). For the dipole moment time correlation function, the maximum entropy method (MEM) was used as well.[26] In the following, we shall briefly discuss the WK approach.

Let us use $x_T(t)$ to represent a time dependent properties of the system (like dipole

moment and velocity) in such a way that

$$\begin{aligned} x_T(t) &= x(t) \qquad\qquad for\ -T \leq t \leq T \\ &= 0 \qquad\qquad otherwise \end{aligned} \tag{1.58}$$

The Fourier transform of the x_T be $A(\omega)$ is then expressed as:

$$A(\omega) = \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt \tag{1.59}$$

and, by Fourier inversion, we obtain:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} A(\omega) e^{i\omega t} d\omega \tag{1.60}$$

Using Parseval's theorem, one can write:

$$\int_{-\infty}^{\infty} x_T^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |A(\omega)|^2 d\omega \tag{1.61}$$

Now we can define the spectral density as $S(\omega)$ of $x(t)$ as:

$$S(\omega) = \lim_{T \Rightarrow \infty} \frac{1}{2T} |A(\omega)|^2, \tag{1.62}$$

so that

$$\overline{x^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) d\omega. \tag{1.63}$$

If we consider a correlation function $C(\tau)$ for a continuous stationary random process $x(t)$ defined as:

$$C(\tau) = \lim_{T \Rightarrow \infty} \frac{1}{2T} \int_{-T}^{T} x(t+\tau)x(t)dt, \tag{1.64}$$

the Wiener-Khintchine theorem states that the correlation function and the spectral density are simply Fourier transform of each other:[27]

$$C(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{i\omega\tau} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) \cos\omega\tau \ d\omega \tag{1.65}$$

and

$$S(\omega) = \int_{-\infty}^{\infty} C(\tau) e^{-i\omega\tau} \ d\tau = 2 \int_{0}^{\infty} C(\tau) cos(\omega\tau) \ d\tau \tag{1.66}$$

We can thus transform Equation 1.56 as follows:

$$P_v(\omega) = \frac{1}{2\pi NT} \sum_{i}^{N} \left| \int_{+\infty}^{-\infty} v_i(t) e^{i\omega t} dt \right|^2 \tag{1.67}$$

1.7 Summary

In this Chapter, we have provided a summary of the theoretical methods that were used in the Ph.D. project. The scope of this Chapter was not to give a detailed description of each method, since such methods were applied without further and original development. Our purpose was to establish a general background that will be used in the following Chapter to present our work, provide an interpretation of our results and

discuss the different methodological choices that were made during the different stages of the project.

References

(1) I. Cacelli and G. Prampolini, *J. Chem. Theory Comput.*, 2007, **3**, 1803–1817.

(2) 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*, Gaussian Inc. Wallingford CT, 2016.

(3) D. Case, R. Betz, D. Cerutti, T. Cheatham, III, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko,

T. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, D. Roe, A. Roitberg, C. Sagui, C. Simmerling, W. Botello-Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao and P. Kollman, *AMBER 2016*, University of California, San Francisco, 2016.

(4) F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons, Chichester, UK ; Hoboken, NJ, Third edition, 2017.

(5) A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications, Mineola, N.Y, 1996.

(6) A. R. Leach, *Molecular Modelling: Principles and Applications*, Prentice Hall, Harlow, England ; New York, 2nd ed, 2001.

(7) C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, Wiley, Chichester, West Sussex, England ; Hoboken, NJ, 2nd ed, 2004.

(8) R. G. Parr and W. Yang, *J. Am. Chem. Soc.*, 1984, **106**, 4049–4050.

(9) W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.

(10) J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3094.

(11) V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.

(12) M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comp. Chem.*, 2003, **24**, 669–681.

(13) A. Klamt and G. Schurmann, *J. Chem. Soc., Perkin Trans. II*, 1993, 799–805.

(14) E. Cancès, B. Mennucci and J. Tomasi, *J. Chem. Phys.*, 1997, **107**, 3032–3041.

(15) B. Mennucci, E. Cances and J. Tomasi, *J. Phys. Chem. B*, 1997, **101**, 10506–10517.

(16) A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378–6396.

(17) L. Verlet, *Phys. Rev.*, 1967, **159**, 98–103.

(18) C. W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, N.J, 1971.

(19) J. Jung, C. Kobayashi, K. Kasahara, C. Tan, A. Kuroda, K. Minami, S. Ishiduki, T. Nishiki, H. Inoue, Y. Ishikawa et al., *J. Comp. Chem.*, 2021, **42**, 231–241.

(20) T. Schlick, *Molecular Modeling and Simulation: An Interdisciplinary Guide: An Interdisciplinary Guide*, Springer New York, New York, NY, 2010, vol. 21.

(21) S. Genheden, A. Reymer, P. Saenz-Méndez and L. A. Eriksson, in *Computational Tools for Chemical Biology*, 2017, pp. 1–38.

(22) J. M. Dickey and A. Paskin, *Phys. Rev.*, 1969, **188**, 1407–1418.

(23) D. A. McQuarrie, *Statistical Mechanics*, University Science Books, Sausalito, CA, 2000.

(24) S. Egorov and J. Skinner, *Chem. Phys. Lett.*, 1998, **293**, 469–476.

(25) R. Ramirez, T. López-Ciudad, P. Kumar P and D. Marx, *J. Chem. Phys.*, 2004, **121**, 3973–3983.

(26) *Numerical recipes in C++: the art of scientific computing*, ed. W. H. Press, Cambridge University Press, Cambridge, UK ; New York, 2nd ed, 2002.

(27) D. A. McQuarrie, *Statistical Mechanics*, University Science Books, Sausalito, Calif, 2000.

Chapter 2

Hydration Effects on the Vibrational Properties of Carboxylates: Succinate and Aspartate

Contents

2.1 Introduction 58
2.2 Computational details 60
2.3 Results and discussion 62
2.3.1 Results obtained using electronic structure calculations 63
2.3.2 Results obtained using QM/MM molecular dynamics 79
2.4 Conclusions 86

2.1 Introduction

The carboxylate stretch band may be exploited to investigate the local structure surrounding the intercalates, as it is very sensitive to inter(intra)molecular interactions and provides a probe of the local environment.[1–3] Computational work has been carried out to elucidate hydration [4], coordination[5] and both hydration and coordination[6] effects on this infrared (IR) band for carboxylates. However, to our knowledge, a systematic study on how structural differences (e.g., succinate vs. aspartate), the protonation state of the amino group, and the formation of intermolecular hydrogen bonds affect the carboxylate stretch band is still lacking, especially with respect to comparison between different computational methods. As a matter of fact, a fine description of the band position and shape requires sophisticated theoretical approaches (see, for instance, recent work by Vieira Pinto et al. [7] and references therein). On the other hand, a reasonable compromise can be achieved through normal mode analysis from quantum chemistry calculations. However, a validation of such an approach, possibly including a limited number of explicit water molecules, has to be provided on each considered system.[4]

As the first step in our study of molecules containing carboxylate groups, we focused on succinate (SUC^{2-}), zwitterionic aspartate (ASP^{-}), the most stable form in a neutral water solution, and aspartate with a deprotonated amino group (ASP^{2-}), the most stable form in basic solution and in the interlamellar space of LDH.[8] For their structures, we recall Figure 1. We focused our attention on applying different models to describe hydration effects and on comparing the results obtained with different quantum chemistry methods. In a bottom-up approach, firstly we performed electronic structure calculations

using a continuum representation of the water solvent. We then included a number of water molecules corresponding to the number of polar groups (two for succinate, three for aspartate),[4, 9] in a continuum and finally, once the PM6[10] semiempirical level was validated through comparison with density functional theory (DFT) calculations, we ran hybrid quantum mechanics-molecular mechanics (QM/MM) molecular dynamics (MD). Such an approach allowed us to investigate the local water network and the conformational space of the solute (by umbrella sampling[11] along the C-C-C-C dihedral) in the presence of an explicit solvent. An evaluation of the IR spectrum in the carboxylate stretch region was also provided by computing the power spectrum of the (QM) molecular dipole. We would like to point out that the choice of including only a very limited number of water molecules in the quantum calculations instead of developing a more systematic microsolvation approach (for bioorganic molecules, see, for instance, Refs. [12–19]), was supported by the work of Sutton et al. on carboxylates[4] and by Hernández et al. on aspartate,[9] which provide an excellent framework for validation.

This chapter is organized as follows. After a description of the computational procedures in Section 2.2, we describe the results obtained based on electronic structure calculations (Section 2.3.1) and those based on QM/MM molecular dynamics (Section 2.3.2). Conclusions are drawn in Section 2.4.

2.2 Computational details

We performed electronic structure calculations using Gaussian16[20], and Gaussview6 [21] for some analysis and for creating images of the molecular structures. For each anion (SUC^{2-}, ASP^{-} and ASP^{2-}, as defined in Figure 1), and for each of the quantum chemistry methods that will be discussed in the following, we first performed a relaxed scan along the C-C-C-C dihedral. We then performed a full optimization on the minima found along the constrained potential energy surface (PES) and characterized them through an analysis of the frequencies of the normal modes (at 298 K and 1 bar). The stationary points would be considered minima only if no imaginary frequency was found. The normal mode frequencies and intensities from the quantum calculations were processed by means of a convolution of Gaussian functions having a full width at half maximum of 20 cm^{-1}, to obtain infrared bands. The assignment of the bands was confirmed by an analysis of the normal coordinates and by a visual inspection of the normal modes using Gaussview6.

The choice of the methods was based on recent work that has already been published in the literature: density functional theory using the M05-2X[22] functional and the cc-pVDZ basis set,[23] used by Sutton et al. for carboxylates among which succinate,[4]; DFT with the B3LYP[24] functional and the 6-311++G(d,p) basis set,[25, 26] used by Hernández et al. for aspartate.[9] Results obtained with these methods have been compared with PM6[10] semiempirical calculations. Two different continuum models of solvation were used, namely, the solvation model density (SMD)[27] and the integral equation formalism version[28, 29] of the polarizable continuum model (PCM, see Chapter 1).

For each minimum along the PES of the three molecules, we added one water molecule sitting next to the polar groups (i.e. 2 water molecules around succinate - one per carboxylate group- and 3 water molecules around aspartate - one per carboxylate groups and one for the amino group). The structures thus obtained were optimized and normal modes analysis based on the harmonic approximation was performed to confirm the obtention of minima using the same choice of methods already described for the anions, in the presence of water as a continuum medium. This procedure follows the ideas in Refs. [4, 9], in which such a microsolvation approach was applied to analyze the formation of a solute-solvent hydrogen bond. In aspartate, such a band is sensitive to couplings with the amino group,[30], and it is, therefore, important to include the effect of the interaction of explicit water molecules with this group.

Having found a picture that is qualitatively the same for the different levels of theory, we proceeded to use PM6 based QM/MM molecular dynamics with the AMBER16[31] suite of programs. These simulations were performed by Dr. F. Ingrosso and the theoretical framework of QM/MM will not be discussed here, though the computational setup will be provided below. After minimization using the steepest descent algorithm, we performed a 10 ps equilibration in the NPT ensemble using Andersen's thermostat[32] (velocity rescaling every 1 ps) and Berendsen's barostat[33] in a cubic box containing one solute molecule and about 4000 water molecules, using periodic boundary conditions and a modified Ewald scheme for treating long-range electrostatics in QM/MM based dynamics.[34] The force field used for water is TIP3P,[35] widely employed to describe water within biological systems. The time step used in the simulations was 1 fs. After equili-

bration, each trajectory was propagated in NPT (1 bar, 300 K) for 500 ps and used for data analysis by employing post-processing tools built in the AMBER16 package as well as in-house codes. In order to have a qualitative description of the configurational space of the three anions in solutions, umbrella sampling[11] simulations were performed along the C-C-C-C dihedral with the QM/MM scheme. The potential of mean force (PMF) was then built by applying the weighted histogram analysis method (WHAM).[36] Umbrella potentials were placed along the sampling coordinate varying from -180° to 180° (20 windows). The restraining potential was set to 35 kcal/mol/rad^{-1} and the simulations were run in the NPT ensemble, using the same conditions as described above for the equilibrium trajectory.

2.3 Results and discussion

For each of the anions considered, quantum chemistry calculations using two continuum models of solvation will be presented, followed by the calculations on clusters including explicit water molecules (Section 2.3.1). In Section 2.3.2, we shall describe the solvation properties that we investigated for the anions in bulk water by using a QM/MM scheme, providing a qualitative comparison with the picture stemming from static calculations.

2.3.1 Results obtained using electronic structure calculations

2.3.1.1 Succinate: SUC^{2-}

We shall start the presentation of the results by providing a comparison with similar studies in the literature. In line with the work in Reference [4], we found three conformers of succinate, the structure of which is reported in the top panel of Figure 2.1 along with some useful definitions. The conformers will be named SUC_1^{2-}, SUC_2^{2-}, SUC_3^{2-}, corresponding to the S1, S2, S3 conformers obtained by Sutton et al. The results obtained by those authors were confirmed by performing a relaxed scan along the χ dihedral and reoptimizing the minima found along the constrained surface. A comparison of the minima sampled by this methodology using M05-2X/cc-pVDZ and SMD with the published structures.[4] SUC_1^{2-} corresponds to a gauche conformation of the two carboxylates with respect to the bond formed by the two methylene carbons, whereas SUC_2^{2-} and SUC_3^{2-} correspond to an anti conformation and different because of the relative orientation of the planes containing the COO^- groups.

The values of the C_1-C_2-C_3-C_4 dihedral as well as the relative energies with respect to SUC_1^{2-} are shown in Table 2.1. Three quantum chemistry levels were used as well as two different continuum models of solvation. The overall agreement among the different methods used in describing the most relevant features of the three conformers is satisfying. In the case of SUC_1^{2-}, the value of the dihedral is slightly different in the case of PM6 calculations, and this method describes a SUC_3^{2-} conformer that is practically isoenergetic to SUC_2^{2-}. In addition, PM6+SMD predicts higher energy for SUC_2^{2-} and SUC_3^{2-} than

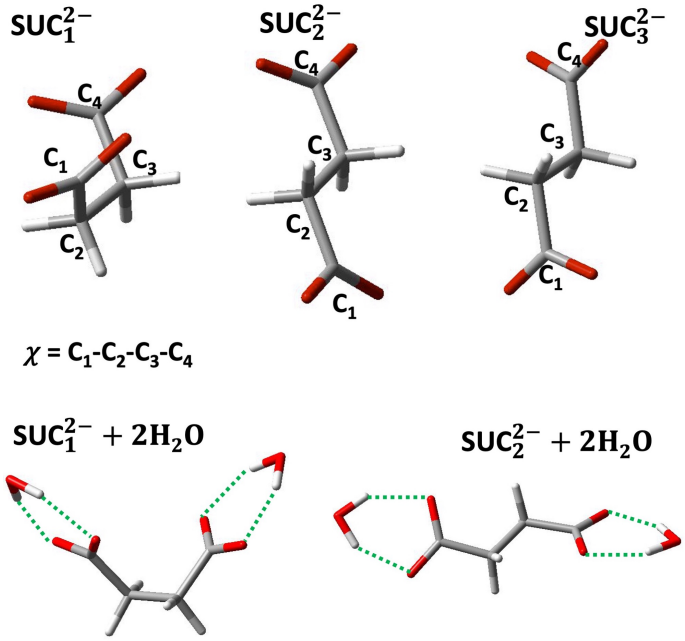


Figure 2.1: Top panel: structure of the three conformers of succinate and definition of the dihedral used in the analysis. SUC_1^{2-}, SUC_2^{2-}, SUC_3^{2-} correspond to the S1, S2, S3 conformers in Reference [4], respectively. Bottom panel: structure of the optimized clusters formed by SUC_1^{2-} and SUC_2^{2-} and two water molecules next to the charged groups. Green dotted lines are shown as a guide for the eye along the succinate-water hydrogen bond directions.

PM6+PCM. However, considering how close in energy such conformers are overall, the observed differences between two different quantum energy methods are of the same order of magnitude as the differences observed when the same QM method is used with two different solvation models. Notably, using M05-2X/cc-pVDZ+PCM predicts a lower energy SUC_3^{2-} conformers whereas the opposite is found when using M05-2X/cc-pVDZ+SMD.

Table 2.1: Summary of structural properties and relative energies for the SUC_1^{2-}, SUC_2^{2-}, SUC_3^{2-} conformers: the χ dihedral is defined in Figure 2.1. The relative energy is computed with respect to the energy of the SUC_1^{2-} conformer.

SUC_1^{2-}		
Method	χ	**ΔE (kJ/mol)**
B3LYP/6-311++G(d,p), PCM	-69°	0.0
B3LYP/6-311++G(d,p), SMD	-66°	0.0
M05-2X/cc-pVDZ, PCM	-69°	0.0
M05-2X/cc-pVDZ, SMD	-63°	0.0
PM6, PCM	-59°	0.0
PM6, SMD	-42°	0.0
SUC_2^{2-}		
Method	χ	**ΔE (kJ/mol)**
B3LYP/6-311++G(d,p), PCM	179°	-0.74
B3LYP/6-311++G(d,p), SMD	178°	+1.09
M05-2X/cc-pVDZ, PCM	177°	+1.06
M05-2X/cc-pVDZ, SMD	176°	+3.62
PM6, PCM	178°	+2.40
PM6, SMD	174°	+11.20
SUC_3^{2-}		
Method	χ	**ΔE (kJ/mol)**
B3LYP/6-311++G(d,p), PCM	-179°	-0.77
B3LYP/6-311++G(d,p), SMD	-176°	+1.30
M05-2X/cc-pVDZ, PCM	-162°	-0.51
M05-2X/cc-pVDZ, SMD	-173°	+4.19
PM6, PCM	-178°	+2.40
PM6, SMD	-173°	+11.20

We turn now to the discussion of the carboxylate stretch band in the infrared spectrum of succinate. A digitized Fourier Transform Infrared (FTIR) spectrum from Reference [37] is shown in black in Figure 2.2, together with the results that we obtained from normal mode analysis based on quantum chemistry calculations. No scaling factor was used in the present analysis for the reported frequencies. Intensities are rescaled to make them comparable. When looking at the different spectra, one basic observation is that the computed spectra from conformers SUC_2^{2-} and SUC_3^{2-} display negligible differences. This is to be expected since the two structures are extremely similar. Two bands are observed in the computed spectra with one exception, namely the SMD calculation using PM6, for which the two bands merge.

The analysis of normal modes allows us to attribute the two bands to symmetric (S, lower frequency) and antisymmetric (AS, higher frequency) stretches. In the comparison with the experiment, the B3LYP/6-311++G(d,p)+PCM spectra in water seem to provide the best agreement, both for the position of the two peaks, the frequency gap between them and the relative height of the peaks, especially for SUC_1^{2-}. However, when using B3LYP/6-311++G(d,p)+SMD the gap between the peaks is too small, since the AS peak is red-shifted, and this is particularly true for SUC_1^{2-}. Finally, the results obtained with PM6+PCM are quite reasonable, though slightly blue-shifted.

The next step in our methodology was to include in our quantum chemistry calculations explicit water molecules interacting directly with the charged groups. We then optimized one cluster containing two water molecules, one by each carboxylate, for the three conformers considered. In all cases, starting from a cluster with SUC_3^{2-} led to a

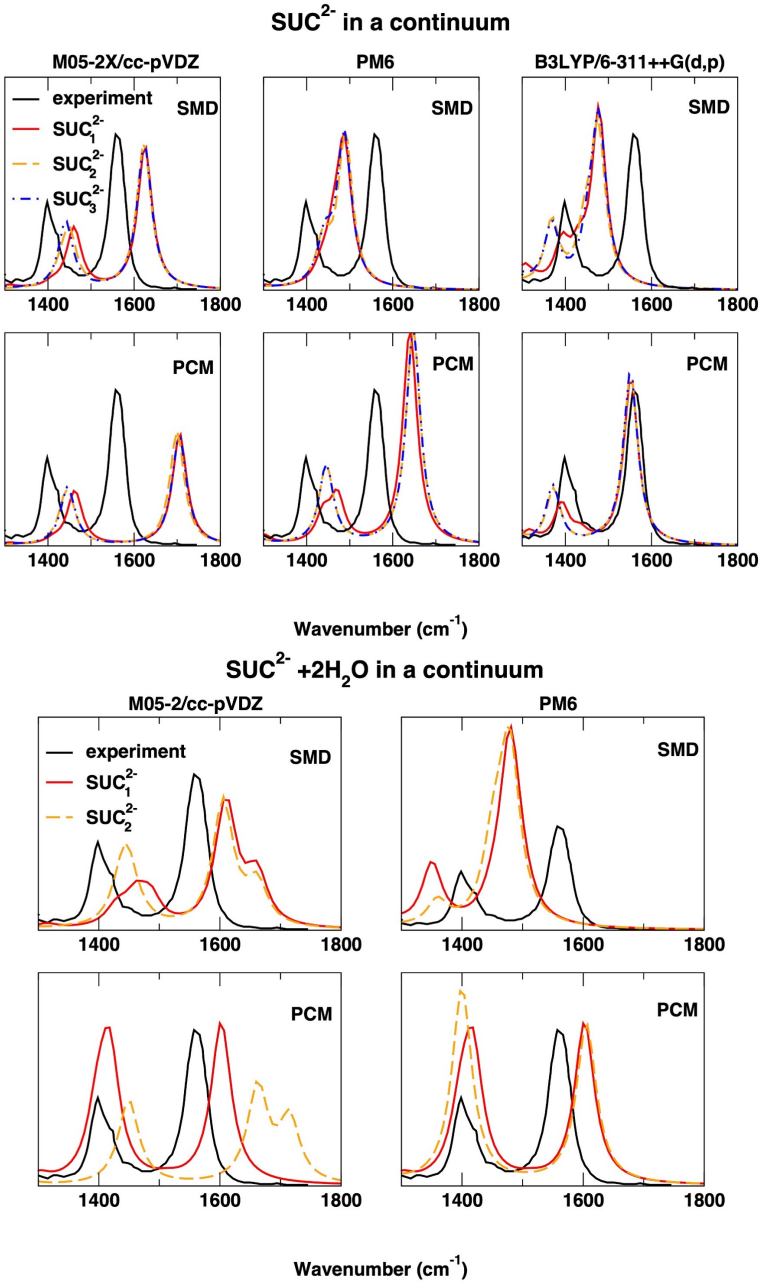


Figure 2.2: Infrared spectra obtained using different implicit solvation models (top panel) and incorporating two interacting water molecules (bottom panel). A comparison is provided with an experimental FTIR spectrum of succinate in water solution[37]. Intensities are expressed in arbitrary units.

change in conformation to SUC_2^{2-}, which is confirmed by the results in Reference [4]. The structures of the two optimized clusters are reported in the bottom panel of Figure 2.1. For the following discussion on such clusters, we shall keep the M05-2X/cc-pVDZ level used by Sutton et al. to compare with the semiempirical PM6 results. The DFT method predicts the SUC_2^{2-} cluster to be slightly more stable than SUC_1^{2-} (by -1.10 kJ/mol with SMD and by -2.93 kJ/mol with PCM). The same trend is obtained with PM6+SMD (SUC_2^{2-} more stable than SUC_1^{2-} by -5.66 kJ/mol), while the SUC_1^{2-} cluster is predicted to be more stable by +0.68 kJ/mol using PM6+PCM. As in the case of the isolated molecules in a continuum, such energy differences are extremely small.

In both SUC_1^{2-} and SUC_2^{2-} minima, the water molecules coordinate to the carboxylate to form two hydrogen bonds each, giving rise to structures that are visually very symmetric. Using the DFT+SMD, neither the value of the O-C-O angle (125-126°) or that of the C-O distances (1.26-1.27 Å) is perturbed by the formation of the succinate-water clusters. It is worth observing that, for each carboxylate, the two hydrogen bonds do not form at the same distance: we measure 1.71 A and 2.27 Å using DFT+SMD, whereas 1.70 A and 2.37 Å are measured for the results based on DFT+PCM. On the other hand, independently of the conformer, these hydrogen bond distances are quite similar when using PM6+PCM (1.79 and 1.80 Å), while they are different with PM6+SMD (1.79 and 1.91 Å). The O-C-O angle is about 120° and not much affected by the interaction with explicit water, and the two hydrogen bond distances are comparable (1.79-1.80 Å), with both PM6+PCM and PM6+SMD.

Considering now the infrared spectra for the microsolvation model, the S and AS

peaks obtained for the two conformers are very similar in all cases, apart from DFT+PCM where the two cluster spectra differ (bottom panel of Figure 2.2). The result obtained with PM6+SMD is improved compared with the case in which no explicit water molecule was taken into account. However, the analysis of normal modes shows that this is due to the water bend mode strongly coupling to the S stretch. On the other hand, in the case of the DFT+PCM results, the water bend absorbs between 1700 cm^{-1} and 1750 cm^{-1}, thus strongly affecting the band shape for the AS stretch. Overall, the agreement with the experimental band is fairly good but this might be misleading, since the contribution of the water bend is present in all cases in the coupling with the symmetric stretch, whereas this should not be directly comparable with the experiment, in which only the succinate absorption is obtained as a difference spectrum with respect to a blank.

2.3.1.2 Aspartate: ASP^-

Considering the anions from the aspartic acid amino acid, in the case of ASP^-, recent work by Hernández et al.[9] provided a deep analysis of the potential energy surface (PES) of the anion in a continuum solvent (water). By using the same DFT functional (B3LYP) and the same basis set (6-311++G(d,p)), we reproduced the same results using our approach, based on a relaxed scan along the C-C_α-C_β-C_γ dihedral and reoptimizing the minima thus found. The structure of the two conformers along with the atom numbering and the definition of some relevant geometrical coordinates are displayed in the top panel of Figure 2.3.

These two minima correspond roughly to an anti (ASP_1^-) and a gauche (ASP_2^-)

conformation, considering the orientation of the carboxylate groups with respect to the C_α-$C_\beta$bond. The two conformers correspond to g-g+ and g+g- in Reference [9], respectively. To describe the molecular structures, two more dihedrals were used in that work, named χ_1 and χ_2. The values of the dihedral characterizing the conformers as well as the relative energies with respect to ASP_1^- are reported in Table 2.2. A comparison is provided between the PCM and SMD solvation models and between the DFT level and the semiempirical PM6 results. Computations using M05-2X/cc-pVDZ led to an unphysical proton transfer from the charged amino group to one of the carboxylate oxygen atoms. In both conformers, the stabilization of the optimized structures in water is linked with the formation of an intramolecular hydrogen bond between the acidic hydrogen atoms and the negatively charged carboxylate oxygen.

Table 2.2: Summary of structural properties and relative energies for the ASP_1^- and ASP_2^- conformers: the χ_i dihedrals are defined in Figure 2.3 and the relative energy is computed with respect to the energy of ASP_1^-.

ASP_2^-				
Method	χ	χ_1	χ_2	**ΔE (kJ/mol)**
B3LYP/6-311++G(d,p), PCM	-72°	49°	-34°,149°	+0.48
B3LYP/6-311++G(d,p), SMD	-65°	58°	-43°,140°	+2.26
PM6, PCM	-64°	56°	-43°,137°	-2.57
PM6, SMD	-51°	71°	-44°,136°	-7.05
ASP_1^-				
Method	χ	χ_1	χ_2	
B3LYP/6-311++G(d,p), PCM	-170°	-50°	-143°,38°	0.00
B3LYP/6-311++G(d,p), SMD	-179°	-57°	-146°,36°	0.00
PM6, PCM	-177°	-58°	-142°,39°	0.00
PM6, SMD	169°	-69°	-141°,40°	0.00

DFT calculations predict ASP_2^- to be more stable than ASP_1^- while the opposite

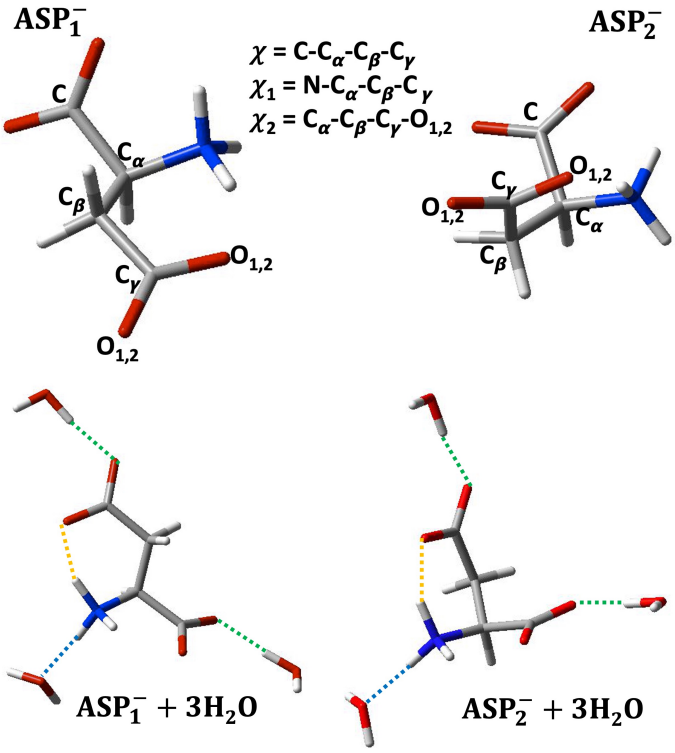


Figure 2.3: Top panel: structure of the two conformers of deprotonated aspartate and definition of the relevant dihedrals used in the analysis. Bottom panel: structure of the optimized clusters formed by ASP_1^- and ASP_2^- and three water molecules next to the charged groups. Relevant H-bond directions are displayed as a guide for the eye. Green dotted lines: intermolecular H-bonds, water is the donor. Blue dotted lines: intermolecular H-bond, water is the acceptor. Yellow dotted line: intramolecular H-bond.

trend is predicted by PM6. As in the previous case, the differences in energies are very small. The structures predicted by the two methods are very similar with both solvation models, the only exception being the structure of ASP_2^- with SMD, for which the χ dihedral is different (about 10°).

In the top panel of Figure 2.4 we present a comparison of the computed carboxylate stretch bands using normal mode analysis. A recorded spectrum of this species not being available in the literature, we shall use the data provided in the work by Pearson and Slifkin[30]. They assign a band at about 1400 cm^{-1} to S stretch and 1590 cm^{-1} to asymmetric stretch. The S deformation of NH_3^+ occurs at about 1500 cm^{-1}. A recorded spectrum not being presented in that work, we only reported the band positions in the plot. As a general remark, all computed spectra predict reasonably well the S band but not the N-H deformation. When using DFT+PCM and PM6+PCM, the AS band tends to be blue-shifted. The analysis of the normal modes allowed us to observe that the NH_3^+ deformation couples strongly to the AS stretch and that, in addition, the lower energy tail of the band is associated with a contribution from the C-H deformation.

To discuss the results obtained for the calculations, including explicit water molecules (structures displayed in the bottom panel of Figure 2.3), we shall focus on the results obtained with PCM. For the ASP_2^- conformer and for B3YLP/6-311++G(d,p), there is no significant effect on the O-C-O angles (125-128°) upon formation of solute-water hydrogen bonds compared with the result obtained on the solute in the continuum. In the absence of explicit water molecules, the two C_γ-O bonds are different: the two distances are equal to 1.24 and 1.26 Å. The longer distance is associated with the O atom involved in an

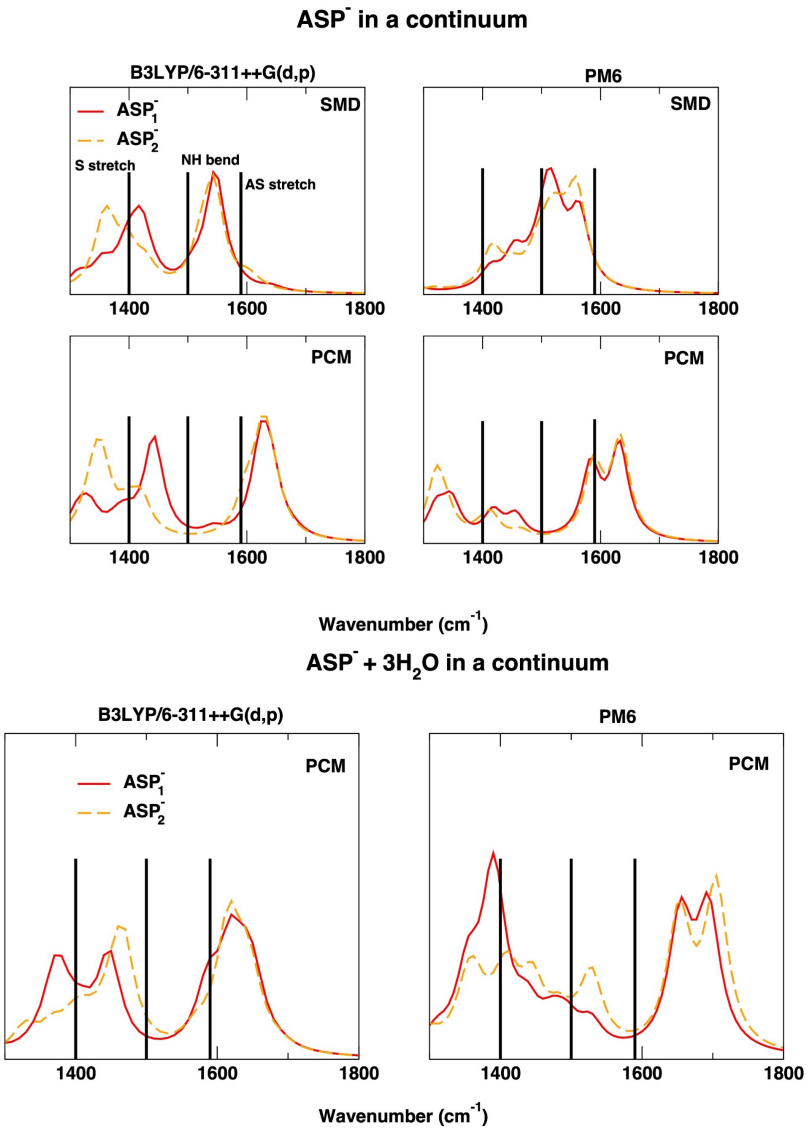


Figure 2.4: Infrared spectra for the two minima ASP_1^- and ASP_2^-. Top panel: results for the anions in a continuum solvent. Bottom panel: results for the anions surrounded by three water molecules next to the polar groups in PCM. Experimental frequencies estimated in Reference[30] for different modes are displayed as black lines. Intensities are expressed in arbitrary units.

intramolecular H-bond with one of the amino hydrogen atoms (bond displayed in yellow in Figure, N...H distance: 1.61 Å).

Once the complex is formed, the N...H distance increases slightly (to 1.80 Å). Each of the carboxylate groups forms one H-bond with one water, a situation that is different with respect to what was described for succinate. The C-O distances are all equivalent and very little affected by the formation of H-bonds (1.25 Å). The O..H distances in the presence of explicit water are equal to 1.73 Å. For the ASP_1^- conformer, the general trend is very similar to what was discussed for ASP_2^-. The intramolecular hydrogen bond involving one amino H atom is formed with one oxygen atom bonded to the C carbon (instead of C_γ) and moves from 1.60 to 1.70 Å after interaction with explicit water molecules. The results obtained with PM6 are very similar and will not be discussed in detail. The only relevant difference appears for ASP_1^-, where both oxygen atoms of C_γ form H-bonds with one water molecule (H-bond distances: 1.90 and 1.88 Å). In addition, for both conformers, the H_2O...HN distance is shorter (1.80 Å) than the one computed with DFT.

To conclude our discussion on the ASP^- conformers surrounded by explicit water molecules, we present computed absorption spectra from normal mode analysis in the bottom panel of Figure 2.4. In the DFT spectra, water bend is mostly coupling with AS stretch whereas NH_3^+ deformation is coupling with the S stretch band. C-H deformation at C_α plays a role below 1400 cm^{-1}, providing peak splitting on the S band for ASP_1^-. As for the succinate spectra, water bend couples with AS stretches. In the PM6 result, the AS band is split and the normal mode analysis shows that this is due to the two carboxylate groups oscillating at different frequencies. The NH_3^+ deformation couples

with S stretches, giving rise to complex features on this band. Coupling with the water bend is strong around 1370 cm^{-1}.

2.3.1.3 Aspartate: ASP^{2-}

Moving from ASP^- to ASP^{2-}, we find some similarities as well as some interesting differences. Two minima are found along the PES, named ASP_1^{2-} and ASP_2^{2-}, the structure of which is shown in the top panel of Figure 2.5. Following the same steps as for the previous anion, we provide results for DFT and PM6 calculations in SMD and in PCM water. Some of the relevant parameters to describe ASP_1^{2-} and ASP_2^{2-} are collected in Table 2.3. ASP_1^{2-} is predicted to be the most stable conformer with all methods but one exception, namely DFT+SMD. This might be related to a slightly different geometry predicted by this method for ASP_2^{2-}, displaying a somewhat different orientation of the carboxylates compared to the other methods. The same effect is observed with PM6+SMD and it also leads to a less stable ASP_1^{2-} compared to what is predicted by PM6+PCM. As for the other two anions, the energy differences are small. As a general remark, the carboxylate moieties are more free to reorient compared to ASP_1^- and ASP_2^- because the formation of an intramolecular H-bond is not as favorable as in those cases (the H atoms belonging to the amino site being less acidic).

In the absorption spectra, the contribution from NH deformation is much smaller than in the ASP^- case. The symmetric stretch is at about 1400 cm^{-1} and the intensity is very low. On the other hand, a strong contribution from C-H deformation is playing a role to give rise to a large and intense AS stretch band. The C-H bend and CH_2 wagging

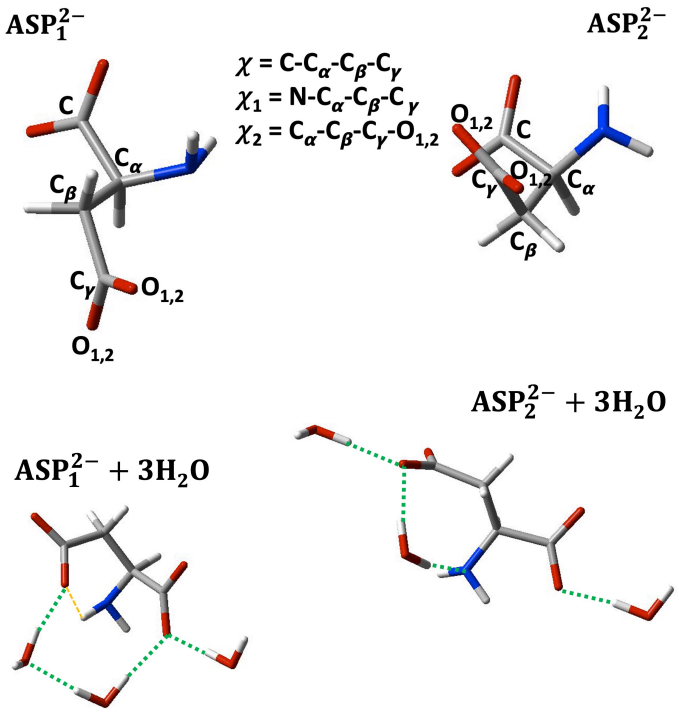


Figure 2.5: Top panel: structure of the two conformers of aspartate having a deprotonated amino group (ASP_1^{2-}, ASP_2^{2-}) and definition of the relevant dihedrals used in the analysis. Bottom panel: structure of the optimized clusters formed by ASP_1^{2-} and ASP_2^{2-} and three water molecules sitting next to the charged groups. Green dotted lines: intermolecular H-bonds, water is the donor. Yellow dotted line: intramolecular H-bond.

Table 2.3: Summary of structural properties and relative energies for the ASP_1^{2-} and ASP_2^{2-} conformers: the χ_i dihedrals are defined in Figure 2.5 and the relative energy is computed with respect to the energy of the ASP_2^{2-} conformer.

ASP_1^{2-}				
Method	χ	χ_1	χ_2	**ΔE (kJ/mol)**
B3LYP/6-311++G(d,p), PCM	168°	-66°	-48°, 134°	+11.4
B3LYP/6-311++G(d,p), SMD	168°	-65°	-50°, 132°	-9.20
PM6, PCM	163°	-73°	-63°, 118°	+5.48
PM6, SMD	-179°	55°	-61°, 120°	+2.92
ASP_2^{2-}				
Method	χ	χ_1	χ_2	**ΔE (kJ/mol)**
B3LYP/6-311++G(d,p), PCM	-74°	53°	-126°, 55°	0.00
B3LYP/6-311++G(d,p), SMD	-65°	63°	-150°, 32°	0.00
PM6, PCM	-60°	65°	-63°, 118°	0.00
PM6, SMD	-62°	60°	-96°, 83°	0.00

are predicted at about 1350 cm^{-1} and no contribution from the N-H bend is expected in a basic environment in this frequency region.[30] With PM6+SMD, the S and AS bands are very similar (1450 and 1490 cm^{-1}), the contribution of the C-H deformation and CH_2 wagging couples to both modes. In the case of PM6+PCM, the band gap is larger and the mixing with N-H bend replaces the contribution from the methylene group.

In the clusters including three explicit water molecules (bottom panel of Figure 2.5), in the case of ASP_1^{2-}, a weak intramolecular H-bond is formed between one of the amino H atoms and one O atom on C_α (distance 2.35 Å with DFT, 2.52 Å with PM6). The N atom is not directly accepting a hydrogen bond from water, while one water molecule bridges between the other two interacting directly with the carboxylate O atoms. In the ASP_2^{2-} cluster, the intramolecular hydrogen bond is formed between HN and the O attached to C_α (distance: 2.01 Å), whereas the N atom is accepting a hydrogen bond from one water molecule (distance: 1.93 Å with DFT, 2.09 Å with PM6) that is also involved in a bond with a carboxylate (1.89 Å- 1.83 Å) in a bridging position. No apparent differences between the parameters describing the carboxylate groups are observed compared with the calculations on ASP_1^{2-} and ASP_2^{2-} in the absence of explicit water.

Minor differences are observed in the IR spectra computed for the clusters to describe the carboxylate stretch band (bottom panel of Figure 2.6). Using B3LYP/6-311++G(d,p), we observe that the S stretch band is mostly affected by the modes of the methylene groups, whereas the AS stretch contains the coupling with the bend of the different water molecules. The higher frequency tail is due to the interaction with the NH bend. Using PM6, the coupling of the water bend mode with the carboxylate stretch is

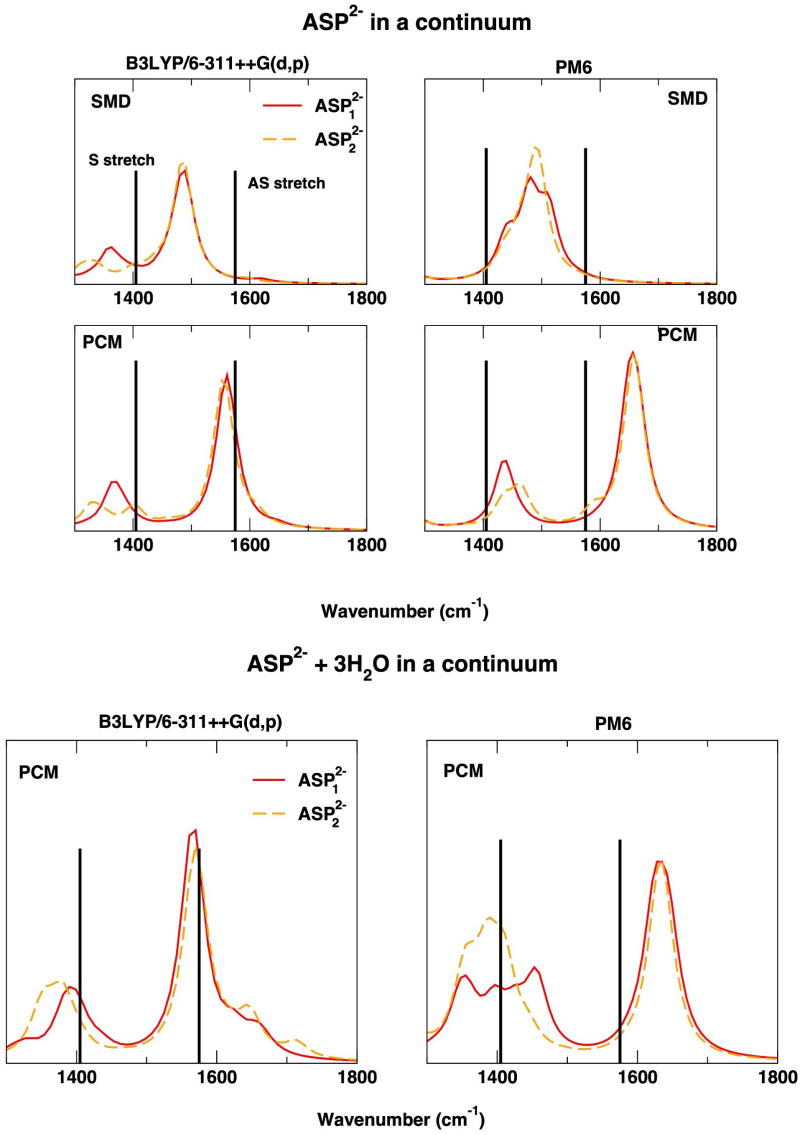


Figure 2.6: Infrared spectra for the two minima ASP_1^{2-} and ASP_2^{2-}. Top panel: results for the anions in a continuum solvent. Bottom panel: results for the anions surrounded by three water molecules next to the polar groups in PCM. Experimental frequencies estimated in Reference [30] for different modes are displayed as black lines. Intensities are expressed in arbitrary units.

possible for the S and the AS band, involving different water molecules. The contribution from the modes involving the methylene and amino groups is limited, compared with the results in a continuum. This leads to an improvement in the band shape since a band gap is now observed between the S and the AS stretch.

2.3.2 Results obtained using QM/MM molecular dynamics

We now move to a more realistic solvation model and consider a QM/MM simulation combining PM6 for the anions with TIP3P for the water molecules. We first focus on the conformational space and analyze the umbrella sampling profiles obtained by constraining the χ dihedral. Results are shown in Figure 2.7. For succinate, the global minimum is located at about 80°, corresponding to a distorted gauche conformation. Other, more shallow minima are present: one at -90° and the other corresponding to the anti conformation observed in SUC_2^{2-} and SUC_3^{2-}. The barriers to move from one basin to the other are quite small (of the order of 20-30 kJ/mol or smaller), we should therefore expect that the molecules explore different conformations along an MD trajectory. This is also true for aspartate, for both ASP^- and ASP^{2-}. However, the global minimum for ASP^- is located at -60° (corresponding to ASP_2^-), whereas the one for ASP^{2-} is located at about 55° (corresponding to ASP_2^{2-}). Given the limited sampling, we shall keep our analysis of the potential of mean force (PMF) to a qualitative description.

To show that our MD trajectory results are consistent with the PMF, we computed the distribution of the χ dihedrals measured along the dynamics. Results are reported in the Appendix (Figure A.1) and show that our analysis of the conformational space

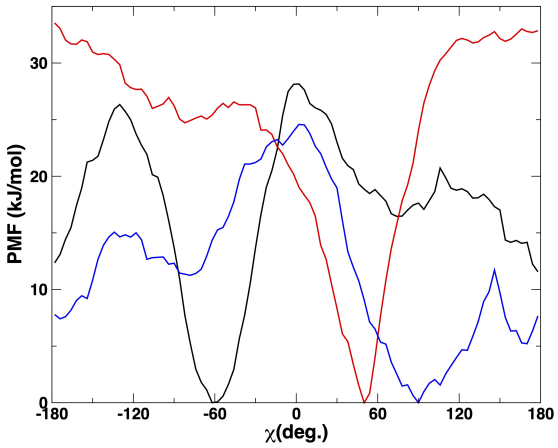


Figure 2.7: Potential of mean force obtained by umbrella sampling along the χ dihedral for the three molecules considered in this work: succinate (SUC^{2-}, blue curve), deprotonated aspartate (ASP^{-}, black curve), deprotonated aspartate in a basic environment (ASP^{2-}, red curve).

corresponds to basins explored along equilibrium MD.

The analysis of the solvation shell of the anions reveals interesting similarities with what has been obtained from electronic structure calculations. We shall start by describing the solute-solvent radial pair distribution functions (RDF, see Chapter 1). In the following analysis, we shall focus on the radial correlation function of pairs of atoms that characterize the interaction of polar groups, displayed in Figure 2.8. For succinate, the interaction between all O atoms from the carboxylate groups with the H atoms of water is remarkably similar, which illustrates that our trajectory is long enough to sample the different solvation structures. The first peak, corresponding to the H-bond distance, is located at 1.83 Å. Two more, quite well defined, peaks are present, corresponding to the second and third hydration shell, hinting at a strong interaction, which is reasonable due to the charge on the carboxylate groups. Similar results are obtained for the two aspartate anions, for which a slightly less structured third peak is observed. However, one of

these RDFs curves, in the case of ASP^-, displays a lower peak height for the first and second peaks, and this behavior is related to one of the oxygen atoms of the carboxylate (belonging to C_γ) forming an intramolecular H-bond with one (two) of the H atoms amino group. One or two HN atoms being involved is related to the rotational motion of the NH_3^+ moiety along the dynamics. A distribution of intramolecular O...HN distances is reported in the Appendix (Figure A.2). This observation is in accord with what was found for ASP_2^- using electronic structure calculations.

When considering the interactions of the amino group with water, major differences are observed between the two aspartate anions. When the amino group is protonated (i.e. ASP^-), the H atoms are weakly donating an H-bond (the first peak is located at 2 Å), while the N atom does not interact directly with the H atoms of water. On the other hand, no interaction is observed between the HN atoms and water for the aspartate bearing a deprotonated amino group (ASP^{2-}), whereas the N atom behaves as an acceptor of the H-bond. Again, this result is very well consistent with what was described for the quantum chemistry calculations including explicit water molecules.

Using to the definition of the coordination number $n_{\mathbf{r}_{min}}$ provided in Chapter 1 (Eq. 1.55), we can provide some insight into the number of water molecules surrounding the polar sites: For each O atom of the carboxylate groups in succinate, we obtain 7 water hydrogen atoms. We also computed the distribution of angles formed between the direction of the water dipole moment and the C-O axis. We found an average angle of 60° for the maximum of the orientational correlation. These calculations, the results of which are

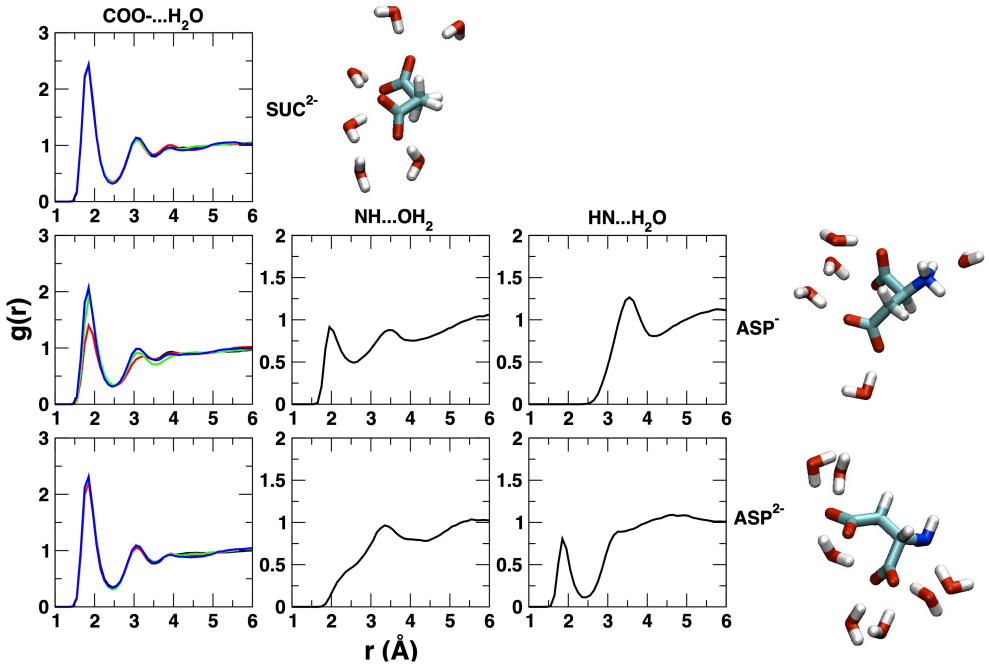


Figure 2.8: Relevant solute-water pair distribution functions. Top panel: interaction between oxygen atoms belonging to carboxylates in succinate (black, red, green, blue curves). Middle panel, from left to right: interaction between oxygen atoms belonging to carboxylates in ASP^- (black, red, green, blue curves); interaction with hydrogen atoms belonging to the amino group; interaction with the nitrogen atom belonging to the amino group. Bottom panel: same as for the middle panel, in the case of ASP^{2-}. For each panel, images illustrating the structure of the solvation shell created by extracting snapshots along the simulated trajectories are displayed. The software used to create the images is VMD.[38]

reported in the Appendix (Figure A.3) are based on the normalized function[39]

$$G_1(\mathbf{r}) = \frac{h^{110}(\mathbf{r})}{3g_{000}(\mathbf{r})} =< \overline{cos(\theta)_{12}} >, \tag{2.1}$$

where $h^{110}(\mathbf{r})$ is based on the orientational correlation function defined as

$$h^{110}(\mathbf{r}) = \frac{\int h(12)\Phi^{110}(12)d\Omega_1\Omega_2}{\int \left[\Phi^{110}(12)\right]^2 d\Omega_1\Omega_2}, \tag{2.2}$$

with

$$\Phi^{110}(12) = \hat{u}_1 \cdot \hat{u}_2. \tag{2.3}$$

$h(12) = g(12) - 1$ is the total pair correlation function between molecules 1 and 2, Φ^{110} is the rotational invariant function related with orientations, and Ω_i are the angular variables. We thus considered $\hat{u}_1$ as a unitary vector along one C-O axis and $\hat{u}_2$ as a unitary vector along the direction of the instantaneous molecular dipole of one solvent molecule.

This finding underlines that each O atom in one carboxylate group interacts with different water molecules, instead of interacting with the same water molecule, as in the optimized structures for the clusters involving SUC_1^{2-} and SUC_2^{2-} (for which the value of the angle is 45-50°). More importantly, in the latter cases, one would expect a population of H-bonded water molecules at shorter distances (namely about 1.7 Å) and a sharper peak in the orientational correlations. As a matter of fact, this finding is not surprising, since the configuration adopted in bulk water leads to a smaller perturbation of the local H-bond network in water compared with the one found for the clusters, and is compatible

with the action of molecular motions that are active when including temperature effects.

The orientation of water molecules with respect to the C-O axes of the carboxylate groups is only slightly different between succinate and the two aspartate anions considered since we obtained a value of 66° associated with the first peak of the G_1 function. From the RDFs, the number of water H atoms surrounding these groups is about 7, as in the case of succinate, with the exception of the C-O bond that is involved in an intramolecular hydrogen bond (in ASP^-), for which only two H atoms were found. The H-bond donor NH_3^+ moiety in ASP^- is surrounded by 5 H atoms.

We used two methods to compute the infrared spectrum numerically by using the instantaneous molecular dipole provided by the QM output in the QM/MM setup: a method based on the Wiener-Khinchin theorem and the maximum entropy method, implemented by using the routines provided in Reference [40]. The procedure has been summarized in Chapter 1.

The power spectra thus obtained are displayed in Figure 2.9.

The carboxylate AS stretch band peaks at 1670 cm^{-1}, compared with 1560 cm^{-1} in experiments. The intensity of the S band is quite low, and the band seems very broad, compared with experiments[37] (for which the peak is located at 1400 cm^{-1}). From our electronic structure PM6 calculations, we may notice that the AS stretch band is more sensitive to the solvation model used, whereas the S stretch band is located at about 1400 cm^{-1} in all cases. Peaks at lower frequencies than 1400 cm^{-1} are due to C-H bend motions, so we can safely assume that the peak on the lower frequency side of the top panel in Figure 2.9 is due to this mode.

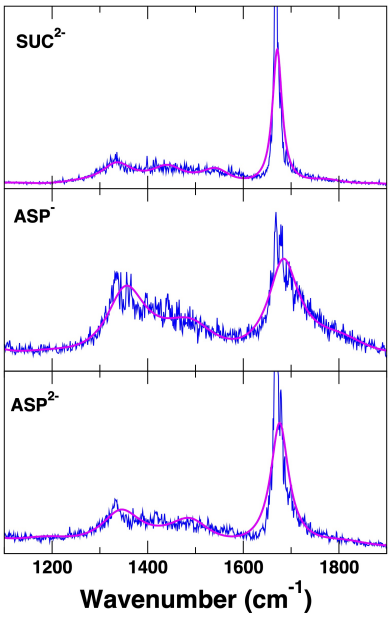


Figure 2.9: Infrared absorption spectra of the three anions considered. Magenta line: result based on the maximum entropy method. Blue line: result based on the Wiener-Khinchin theorem. Intensities are expressed in arbitrary units.

In the case of the computed band for ASP^{-}, the AS stretch band is larger but still peaking at 1670 cm^{-1}. As was already discussed, the experiments suggest that the AS stretch can couple with the bend modes of the NH_3^+ group, and also, in this case, its position is strongly dependent on the solvation model used. The lower frequency portion of the S stretch band that is observed in the spectrum is most likely due to the CH_2 modes. The results are very similar in the case of ASP^{2-}, but the AS band appears narrower, which is consistent with the experimental interpretation based on a lower coupling with the NH bend.[30] To conclude, it is worth noting that, in the case of aspartate, the QM/MM MD-based results display band shapes and positions that are very similar to the results obtained on the anions with PM6+PCM.

2.4 Conclusions

We provided a bottom-up computational approach to study the conformational space and the vibrational properties of three anions containing carboxylate groups: succinate, zwitterionic aspartate, and aspartate having a deprotonated amino group. More specifically, we focused on the carboxylate stretch band, which is sensitive to the environment and, therefore, can be used as a probe of local structure and of solvation and on an exploration of the conformational space of the molecules interacting with water. In the case of succinate and zwitterionic aspartate, some previous computational work in the literature provided important results to compare with.[4, 9] Similarly to what had been done there, we started from the anions in a continuum solvent, then we moved to a (simplified) microsolvation approach, in which we included one water molecule sitting next to each polar group and immersed the cluster thus obtained in a continuum solvent. In addition, we performed hybrid QM/MM MD simulations of the anions in bulk water, thus including finite temperature effects.

In the electronic structure calculations, two DFT methods and one semiempirical method were employed. The results on the predicted geometries, both for the anions and the clusters, were remarkably similar. This underlines the good performance of PM6 methods for predicting the structure of organic anions and their first hydration shell. This might be surprising, considering the artifacts stemming from applying semiempirical Hamiltonians to describe non-covalent interactions.[41] However, we can hypothesize that the strong electrostatic interactions between negatively charged anions (possessing polar groups) and water can be more easily modeled than the interaction between neutral

systems since the limitations of semiempirical methods are more apparent when weak non-covalent interactions are present.

When comparing the results obtained for the carboxylate stretch band based on normal mode analysis, we found that the overall agreement with experimental measurements is satisfying and in agreement with previous work.[4] However, one interesting conclusion that we can draw is that one observes larger differences when comparing results obtained with the same quantum chemistry method but a different continuum solvation model than when comparing two different quantum chemistry methods. Such findings should warn about carrying out calculations of such properties without careful benchmarking. Again, one has to stress here the good performance of PM6 to compute infrared bands.

A description of the local structuring of water molecules around the anions was achieved through the analysis of simulated trajectories in the PM6/MM hybrid scheme. In all treated systems, each carboxylate group interacts with different water molecules, a configuration that leads to a limited perturbation of the H-bond network in water. The H atoms belonging to the charged amino group in the zwitterionic aspartate act as H-bond donors and an intramolecular H-bond, already characterized through QM calculations, is also present. On the other hand, the amino group acts as an H-bond acceptor when it is deprotonated. An exploration of the conformational space of the anions by umbrella sampling along the C-C-C-C dihedral yielded a very similar description compared with the results from electronic structure calculation and it was consistent with the distribution of the dihedral along an equilibrium trajectory. The IR band simulated by computing the power spectrum of the QM dipole delivers the same picture as the IR band from

QM calculations. Further improvement on the band shape and (sub)band position would require specific treatment of quantum effects[42, 43] and/or more advanced techniques.[7]

In the next Chapter, we shall move to the molecular dynamics study of three carboxylates intercalated in layered double hydroxides. In addition to succinate and aspartate (the latter being in the ASP^{2-} protonation state due to the experimental conditions), we chose glutamate, the structure of which is similar to that of aspartate but contains a longer side chain (i.e. one additional methylene group). We, therefore, expect to observe a slightly different behavior of glutamate compared with aspartate under confinement, especially due to a different conformational space.

References

(1) G. Deacon and R. Phillips, *Coord. Chem. Rev.*, 1980, **33**, 227–250.

(2) Y. Lu and J. D. Miller, *J. Colloid Interface Sci.*, 2002, **256**, 41–52.

(3) V. Otero, D. Sanches, C. Montagner, M. Vilarigues, L. Carlyle, J. A. Lopes and M. J. Melo, *J. Raman Spectrosc.*, 2014, **45**, 1197–1206.

(4) C. C. Sutton, G. V. Franks and G. da Silva, *Spectrochim Acta A Mol Biomol Spectrosc*, 2015, **134**, 535–542.

(5) M. Nara, H. Torii and M. Tasumi, *J. Phys. Chem.*, 1996, **100**, 19812–19817.

(6) C. C. Sutton, G. da Silva and G. V. Franks, *Chem. Eur. J..*, 2015, **21**, 6801–6805.

(7) S. M. V. Pinto, N. Tasinato, V. Barone, A. Amadei, L. Zanetti-Polzi and I. Daidone, *Phys. Chem. Chem. Phys.*, 2020, **22**, 3008–3016.

(8) B. Grégoire, V. Erastova, D. L. Geatches, S. J. Clark, H. C. Greenwell and D. G. Fraser, *Geochim. Cosmochim. Acta*, 2016, **176**, 239–258.

(9) B. Hernández, F. Pflüger and M. Ghomi, *J. Comp. Chem.*, 2020, **41**, 1402–1410.

(10) J. J. Stewart, *J. Mol. Mod.*, 2007, **13**, 1173–1213.

(11) G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, 1977, **23**, 187–199.

(12) J. H. Jensen and M. S. Gordon, *J. Am. Chem. Soc.*, 1995, **117**, 8159–8170.

(13) D.-S. Ahn, S.-W. Park, I.-S. Jeon, M.-K. Lee, N.-H. Kim, Y.-H. Han and S. Lee, *J. Phys. Chem. B*, 2003, **107**, 14109–14118.

(14) C. O. da Silva, B. Mennucci and T. Vreven, *J. Phys. Chem. A*, 2003, **107**, 6630–6637.

(15) H.-S. Kim, D.-S. Ahn, S.-Y. Chung, S. K. Kim and S. Lee, *J. Phys. Chem. A*, 2007, **111**, 8007–8012.

(16) S. M. Bachrach, *J. Phys. Chem. A*, 2008, **112**, 3722–3730.

(17) E. Tang, D. Di Tommaso and N. H. de Leeuw, *Phys. Chem. Chem. Phys.*, 2010, **12**, 13804–13815.

(18) D. Scuderi, J. Bakker, S. Durand, P. Maitre, A. Sharma, J. Martens, E. Nicol, C. Clavaguera and G. Ohanessian, *Int. J. Mass Spectrom.*, 2011, **308**, 338–347.

(19) C. García-Iriepa, M. Zemmouche, M. Ponce-Vargas and I. Navizet, *Phys. Chem. Chem. Phys.*, 2019, **21**, 4613–4623.

(20) 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*, Gaussian Inc. Wallingford CT, 2016.

(21) R. Dennington, T. A. Keith and J. M. Millam, *GaussView Version 6*, Semichem Inc. Shawnee Mission KS, 2019.

(22) Y. Zhao, N. E. Schultz and D. G. Truhlar, *J. Chem. Theory Comput.*, 2006, **2**, 364–382.

(23) T. H. Dunning Jr, *J. Chem. Phys.*, 1989, **90**, 1007–1023.

(24) P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.

(25) A. McLean and G. Chandler, *J. Chem. Phys.*, 1980, **72**, 5639–5648.

(26) R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.*, 1980, **72**, 650–654.

(27) C. J. Cramer and D. G. Truhlar, *Acc. Chem. Res.*, 2008, **41**, 760–768.

(28) E. Cancès, B. Mennucci and J. Tomasi, *J. Chem. Phys.*, 1997, **107**, 3032–3041.

(29) B. Mennucci, E. Cances and J. Tomasi, *J. Phys. Chem. B*, 1997, **101**, 10506–10517.

(30) J. Pearson and M. Slifkin, *Spectrochim Acta A Mol Biomol Spectrosc*, 1972, **28**, 2403–2417.

(31) D. Case, R. Betz, D. Cerutti, T. Cheatham, III, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, D. Roe, A. Roitberg, C. Sagui, C. Simmerling, W. Botello-Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao and P. Kollman, *AMBER 2016*, University of California, San Francisco, 2016.

(32) H. C. Andersen, *J. Chem. Phys.*, 1980, **72**, 2384–2393.

(33) H. J. Berendsen, J. v. Postma, W. F. Van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

(34) K. Nam, J. Gao and D. M. York, *J. Chem. Theory Comput.*, 2005, **1**, 2–13.

(35) D. J. Price and C. L. Brooks III, *J. Chem. Phys.*, 2004, **121**, 10096–10103.

(36) S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *J. Comp. Chem.*, 1995, **16**, 1339–1350.

(37) N. J. Nicholas, G. V. Franks and W. A. Ducker, *Langmuir*, 2012, **28**, 7189–7196.

(38) W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.

(39) G. Stell, G. Patey and J. Høye, *Adv. Chem. Phys., Volume 48*, 1981, 183–328.

(40) W. H. Press, S. A. Teukolosky, W. R. Vetterling and B. P. Flannery, *Numerical Recipies in Fortran*, Cambridge University Press; Cambridge, UK, 1992.

(41) A. Marion, G. Monard, M. F. Ruiz-López and F. Ingrosso, *J. Chem. Phys.*, 2014, **141**, 034106.

(42) S. Egorov and J. Skinner, *Chem. Phys. Lett.*, 1998, **293**, 469–476.

(43) R. Ramirez, T. López-Ciudad, P. Kumar P and D. Marx, *J. Chem. Phys.*, 2004, **121**, 3973–3983.

Chapter 3

Molecular Dynamics Simulations of Nanoconfined Carboxylates in Layered Double Hydroxides

Contents

3.1 Introduction 96

3.2 Computational studies of LDH 99

3.3 Computational Setup 101

3.3.1 Simulation protocol 103

3.4 Results and discussion 103

3.4.1 Radial distribution function 104

3.4.2 Orientation, conformations, binding modes of the anions in the interlayer 112

3.4.3 Number density along the z-axis 118

3.4.4 Vibrational properties of intercalated anions 121

3.5 Conclusions 127

3.1 Introduction

Clays and clay-like layered minerals are fine-grained and poorly crystalline materials. The techniques of choice for characterizing these materials are powder X-ray diffraction, transmission electron spectroscopy for the structure, UV-VIS spectroscopy using diffuse reflectance, EXAFS and MAS-NMR for the analysis of the cations in the layers, FT-IR spectroscopy for the intercalated anions.[1] However, it is rather complex to achieve a molecular description of the system, especially with respect to the disordered inter-lamellar space (water structure, anions hydration, etc.). Unlike experimental methods, molecular modeling tools provide an atomistic perspective on the structure and behavior of clay minerals and layered double hydroxides (LDH). Thanks to the availability of reliable and transferable force fields, such as the widely used ClayFF[2], Monte Carlo and molecular dynamics methods have been successfully applied to analyze the interlayer structure and dynamics.[3] In the present study, we worked with hybrid materials synthesized and characterized by our experimental collaborators by bringing some deeper insights into the organization of the nanoconfined region between two inorganic layers in hybrid organic-inorganic LDH. Before moving to the results thus obtained, we shall provide some background information on the structure of these materials.

LDHs have been known for a long time, starting with the discovery of the mineral hydrotalcite and the publication of its structure (chemical formula $[Mg_6Al_2(OH)_{16}]CO_3{\cdot}4H_2O$), in 1915.[4] Minerals with closely related structures are usually known as the sjögrenite-hydrotalcite group. However, it was not until single crystal X-ray diffraction (XRD) studies were carried out by Allmann[5] and Taylor[6, 7] that the main structural features

of LDH were understood.

Generalizing the hydrotalcite formula to any LDHs, we may write:

$$\left[\mathrm{M}^{\mathrm{II}}_{1-x}\mathrm{M}^{\mathrm{III}}_{x}(\mathrm{OH})_2\right]^{+x}\left[A^{-\mathrm{n}}\right]_{\frac{x}{\mathrm{n}}} \tag{3.1}$$

M^{II} and $\mathrm{M}^{\mathrm{III}}$ represent the cations in the layer, and $\mathrm{A}^{-\mathrm{n}}$ is the exchangeable anion.[8] In the layers, the M^{+2} and M^{+3} cations are bridged by OH^{-} units coordinated at the octahedral positions, forming sheets, which are then stacked to give a layered structure similar to that of the mineral brucite $[\mathrm{Mg(OH)_2}]$.[9] In LDH, cations of similar size forming brucite-like sheets carry net positive charges that must be balanced by anions intercalated. Such anions coexist with structural water in the interlayer regions.

The interlayer spacing can expand/contract, allowing the inclusion of anions of different natures and sizes. An illustration of the LDH structure is presented below in Figure 3.1. In addition to magnesium and aluminum, LDH with other layer cations, including zinc, lithium, chromium, iron, calcium, gallium, nickel, indium, cobalt, manganese, and copper, have been synthesized.[10]

As shown in Figure 3.1, the perpendicular distance between the metal cations (in this case Mg^{+2} and Al^{+3}) from two different layers is called interlamellar space, and the area between the hydroxides from two different layers is called interlayer space.

A wide variety of anions can be incorporated into the interlayer region of these hosts using a range of methods. The dimensions and functional groups of the guest molecule are critical in determining the separation between the layers. The number (monolayer,

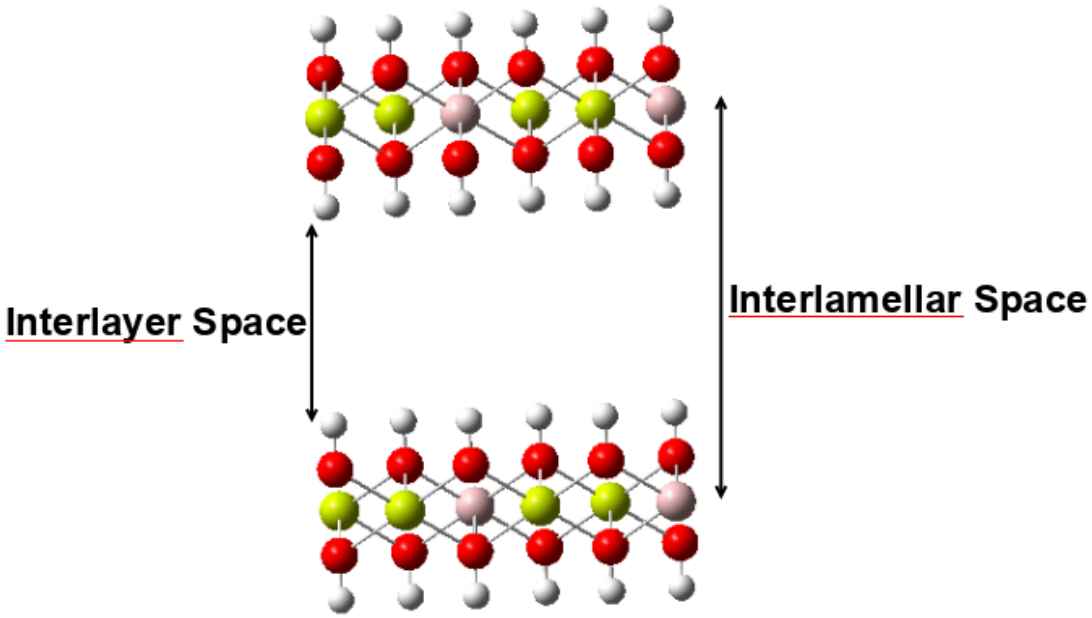


Figure 3.1: The structure of $[Mg_2Al_4(OH)_{12}]_n$ (White: H, Red: O, Yellow: Al, Pink: Mg).

bilayer), size, the orientation of the guest as well as the interactions between the negatively charged guest and positively charged host are all critical factors.[11] The utilization of simple inorganic anions such as the halides, hydroxide, carbonate, sulfate, silicate and chlorate is also well known.[12–15]

In this chapter, we shall describe the MD simulations that we ran for LDH materials having three different intercalated anions (aspartate (ASP^{2-}), succinate (SUC^{2-}) and glutamate (GLU^{2-})), each for different humidity conditions. Experiments were carried out under controlled humidity, giving access to the experimental number of water molecules per anion as well as to the interlamellar space to be used to build the starting configurations corresponding to the experimental conditions. In the case of aspartate and glutamate, the experimental analysis allowed us to know the protonation state of the amino acid molecules.

3.2 Computational studies of LDH

Since 2004, Cygan and co-workers developed a general force field, ClayFF[2], to be used in simulations of mineral systems and their interfaces with aqueous solutions. They evaluated and compared the bulk structures, relaxed surface structures, and intercalation processes to experimental and spectroscopic findings for validation. Simulations of clay, hydroxide, and oxyhydroxide phases and their interfaces with aqueous solutions successfully described the structure and behavior of water, hydroxyl, surface species, and intercalates in these systems.[2] The results obtained in the process demonstrated that ClayFF is a robust force field for molecular simulations of fluid interfaces with clays and other clay-related phases, as well as other inorganic materials characterized by complex structure and composition.[2]

Following the force field development, the number of theoretical works focusing on LDH received a major boost. A thorough discussion of all computational studies on LDH is beyond the scope of this Section. On the other hand, we shall discuss those results that were obtained on systems that are similar to those simulated in this work. Some interesting results on the role of the carboxylate groups in binding with the surface were obtained by Newman et al. for LDH intercalating phenylalanine and tyrosine.[16] The anions were arranged in a bilayer, with their longest molecular axis oriented in an approximately perpendicular direction with respect to the surface. Water molecules were mostly solvating the carboxylate groups and were excluded from the hydrophobic region of the phenyl rings of the anions. Structural differences between these two amino acids did not bring significant differences in the organization of the interlayer. Kumar et al.[17] worked

on an Mg/Al 3:1 LDH containing formate, acetate, and propanoate as intercalating anions. They provided some insights into the evolution of hydration energies as a function of water content. Their results suggested that the carboxylate-water interactions are more favorable than the carboxylate-surface interactions. MD simulations of one LDH sheet in contact with a mixture of anions of different natures (aspartate, glutamate, chloride, and bicarbonate) adsorbed from an aqueous salt solution showed that the carboxylate groups would interact directly with the inorganic surface, whereas the positively charged groups of the amino acids would allow either intercalate-intercalate interactions of solvation by water molecules.[18] The free energy of adsorption is larger for the amino acids than it is for the inorganic anions, thus displaying some degree of selectivity. Similar hypotheses were made based on MD results of LDH intercalating glutamate anions.[19] In a combined experimental-computational work, LDH intercalating aspartate anions were analyzed.[20] The authors focused on examining the possibility of aspartate-aspartate interactions that might lead to favorable situations for the formation of a peptide bond. Increasing the water content seems to prevent boosting such interactions, whereas the dry material is claimed to represent the best solution for concentration and reactivity. Some of the authors extended the MD study to a larger number of amino acids and peptides, confirming the hypothesis according to which drying would favor the interaction between amino acids and dehydration/rehydration cycles would promote the accumulation of these molecules at the surface in a way that the relevant groups for peptide bond formation would find each other in a favorable orientation.[21]

3.3 Computational Setup

In this work, LDH was described using the Clay force field (ClayFF)[2], which is based on an ionic (non-bonded) description of the metal-oxygen interactions and has a flexible force field framework that can be combined with other potentials. Amber force fields (gaff[22] and ff03.r1[23]) were employed to simulate the anions intercalated inside the hydrotalcite galleries. The SPC/E potential was used for water.[24] Details about the force fields are reported in the Appendix (Section B.2).

All MD simulations were conducted using Amber 16 and Amber Tools[25] and visualized with VMD.[26] The initial cell of Mg/Al 2:1 LDH, provided by Dr. E. André, had a three-layer structure to provide three layers and three interlamellar regions in the simulation box. After adjusting the interlamellar space to the experimental value for each set of conditions (anion + increasing amounts of water, see Table 3.1), we replicated the initial cell in the x,y directions (z being the axis orthogonal to the layers). Dr. E. André and Prof. Carteret also provided the number of water molecules per anion and the value of the interlamellar space based on their work on the material (refer to Figures B.1 and B.2). They also determined the protonation state of the aspartate anion based on the experimental conditions (see Section B.1). An illustration of the initial configurations is reported in the Appendix, Figure B.3.

Table 3.1: Parameters defining the box size and the water content of the systems considered.

Systems	**No. of water molecules / anion**	**Interlamellar space (Å)**	**Box size (Å)**		
			$\mathbf{L}_x$	$\mathbf{L}_y$	$\mathbf{L}_z$
LDH intercalating ASP^{2-}					
0_WAT	0 (0%)	9.06	31.707	54.918	27.180
1_WAT	1 (10%)	9.21	31.707	54.918	27.630
3.5_WAT	3.5 (30%)	11.21	31.707	54.918	33.630
6.5_WAT	6.5 (80%)	12.10	31.707	54.918	36.300
LDH intercalating GLU^{2-}					
0_WAT	0 (0%)	8.90	31.707	54.918	26.700
1_WAT	1 (10%)	9.30	31.707	54.918	27.900
4.25_WAT	4.25 (30%)	10.3	31.707	54.918	30.900
6.5_WAT	6.5 (80%)	10.4	31.707	54.918	31.200
LDH intercalating SUC^{2-}					
0_WAT	0 (0%)	8.87	31.707	54.918	26.610
1_WAT	1 (10%)	11.3	31.707	54.918	33.900
6.5_WAT	6.5 (30%)	12.1	31.707	54.918	36.300
8_WAT	8 (80%)	12.1	31.707	54.918	36.300

3.3.1 Simulation protocol

For each system and each considered hydration state, energy minimization was performed using the steepest descent method and conjugate gradient algorithms.[27] The Al^{3+} and Mg^{2+} cations were restrained with force constant 100 kcal/mol/Å^2 and the SHAKE algorithm was also used to keep the distance constant between hydrogen atoms and heavy atoms connected by covalent bonds.[28] A preliminary step for equilibration was achieved by simulated annealing: the system was heated (up to 2500K) and cooled (to 300K) for 5.5 ns by gradually changing the temperature in the NVT ensemble, using Andersen's temperature coupling scheme.[29] During this step, the positions of the OH^- groups, Al^{3+} and Mg^{2+} were constrained with a 500 kcal/mol/Å^2 force constant. A 1 ns trajectory was then run in the NPT ensemble at a temperature of 300 K and a pressure of 1 bar, using Berendsen's barostat.[30] Following this final equilibration step, the constraints on the hydroxyl groups were released, and a force constant of 10 kcal/mol/Å^2 was kept for Al^{3+} and Mg^{2+}. Data analysis was performed on a subsequent NPT simulation (10 ns) using the same parameters as those of the previous NPT step.

3.4 Results and discussion

In all systems considered, a preliminary analysis through visual inspection would clearly show dramatic changes in the orientation of the intercalates as the water content increases. We were interested in understanding how to relate such changes, already observed by other studies in the literature,[19, 20] with the ability of the intercalates to

interact with the surface and with surrounding water molecules, with particular attention to figure out whether the existence/absence of the amino group can play a significant role. We shall start the discussion by presenting the results obtained for the local structure in the interlamellar region.

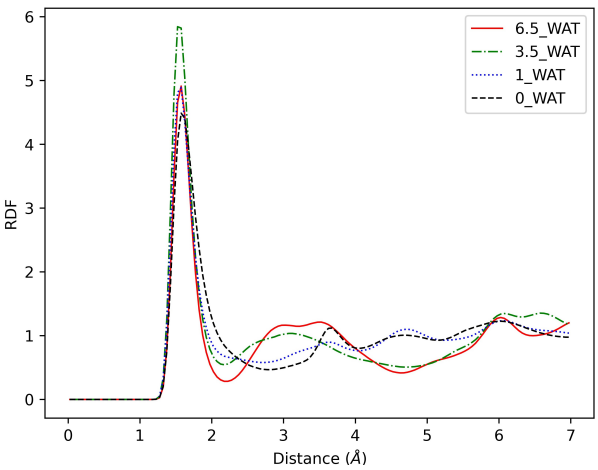
3.4.1 Radial distribution function

The investigated hybrid LDH are homogeneous along the x and y planes but not along the z direction. An analysis of radial distribution functions (RDF, see Eq. 1.54) beyond half of the interlamellar space will not be significant, and we shall therefore limit it to the first peak. The RDF analysis that follows is organized in three categories: interactions involving the layer (Section 3.4.1.1), interactions involving water (Section 3.4.1.2), and interactions among organic intercalates (Section 3.4.1.3).

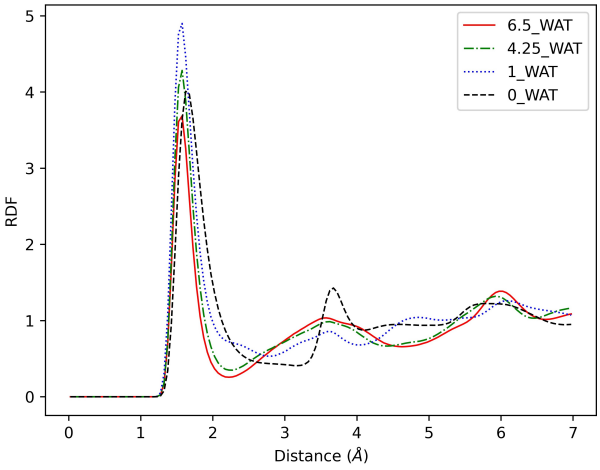
3.4.1.1 Interactions involving the layer

In Figures 3.2, 3.3, and 3.4, we present the RDF describing the interactions between the surface hydroxide groups and the intercalates. Let us start with the interactions describing hydrogen bonding between the carboxylate groups and the hydroxyl groups on the surface (Figure 3.2). The first peak is very similar in all cases, hinting at a strong anchoring interaction. Slightly different behavior is observed in the system intercalating succinate since the first peak becomes more narrow and occurs at a relatively shorter distance in the hydrated LDH compared with the anhydrous one.

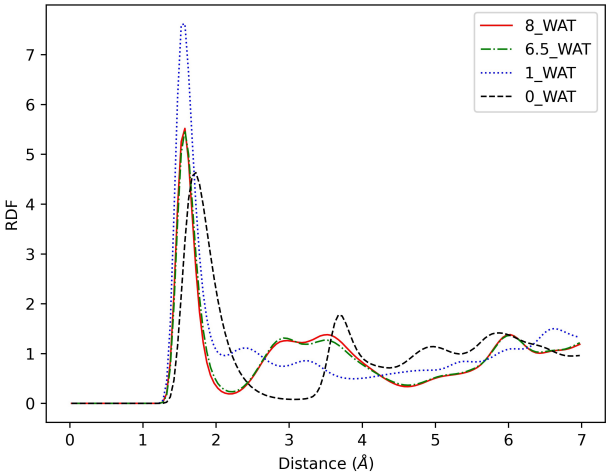
Figure 3.3 displays the OH...NH interaction between the layer and the two amino



(a) OH...OOC interaction, ASP^{2-}



(b) OH...OOC interaction, GLU^{2-}

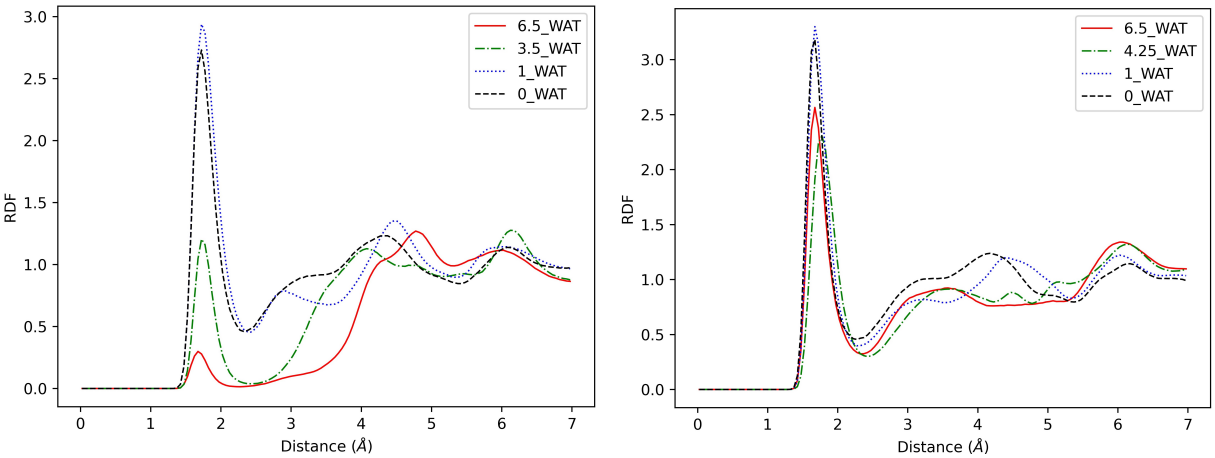


(c) OH...OOC interaction, SUC^{2-}

Figure 3.2: OH...OOC pair correlation functions between the layer and the organic intercalates for the three hybrid systems considered as a function of water content.

acids considered. When aspartate is intercalated (Figure 3.3a), this interaction decreases with increasing hydration, whereas when glutamate is intercalated (Figure 3.3b), we observe no significant modifications. Due to an extra -CH_2 group in the backbone, GLU anions have more degrees of freedom while being confined in LDH, and they could maintain a constant OH...NH interaction.

Finally, in Figure 3.4 we report the interactions between the layer and water molecules,



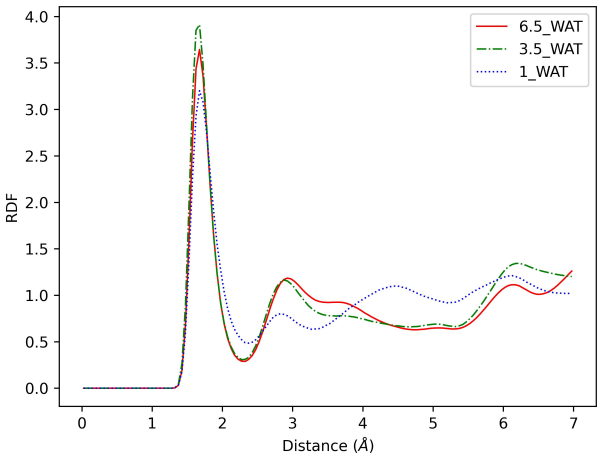
(a) OH...NH interaction, ASP^{2-} (b) OH...NH interaction, GLU^{2-}

Figure 3.3: OH...NH pair correlation functions between the layer and the organic intercalates for the hybrid system containing amino acids, as a function of water content.

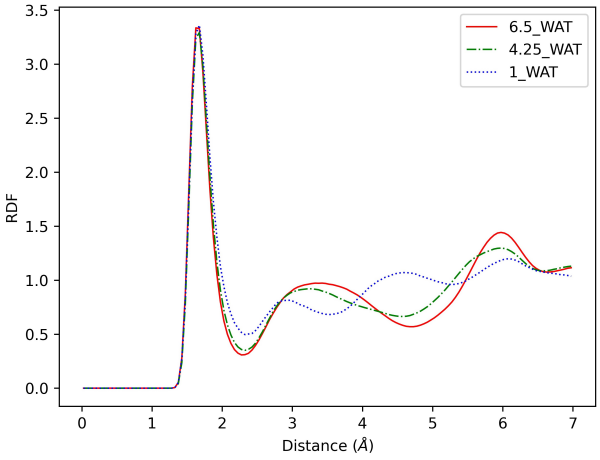
water acting as a hydrogen bond acceptor. The first peak is remarkably similar in all cases, however, some different shapes and trends are observed on the second peak, hinting at a strong interaction with the surface starting from lower hydration conditions and to a widening of the region hosting second neighbors (second peak). This will be further analyzed using the density distributions along the Z direction.

3.4.1.2 Interactions with water

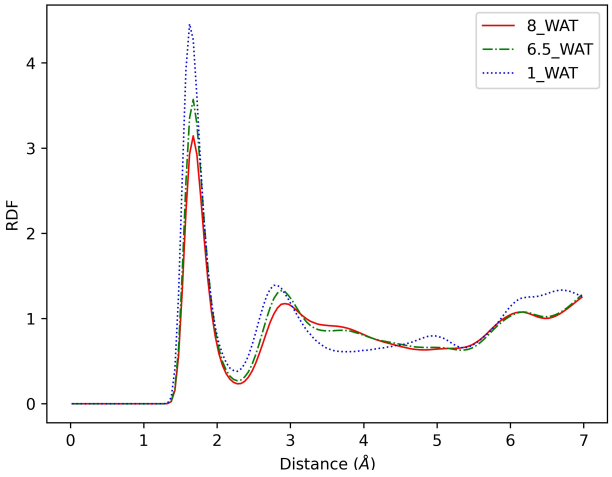
The relevant water-amino group pair correlations are reported in Figure 3.5. Results obtained for a simulation of the corresponding anions in bulk water are shown for a qualitative comparison. As a matter of fact, the water density is not the same in bulk water as in the LDH, thus making a comparison of the peak height meaningless. When aspartate is intercalated, the HOH...NH interactions increase with increasing water content (Figure 3.5a), and a very interesting effect is observed for the second peak, not present at the lowest hydration content but which is significant in the two higher hydration states. On



(a) OH...OH2 interaction, ASP^{2-}



(b) OH...OH2 interaction, GLU^{2-}

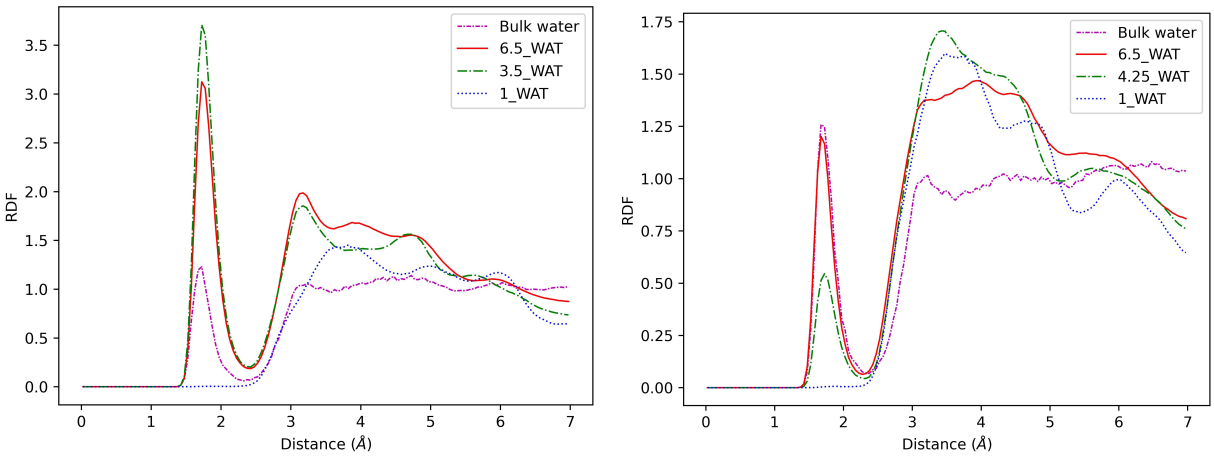


(c) OH...OH2 interaction, SUC^{2-}

Figure 3.4: RDF showing the OH...OH2 pair correlation functions between the layer and water for the three hybrid systems considered as a function of water content.

the other hand, In the case of the LDH intercalating glutamate, the larger effect is observed in the first peak while no further structuring is present at larger distances.

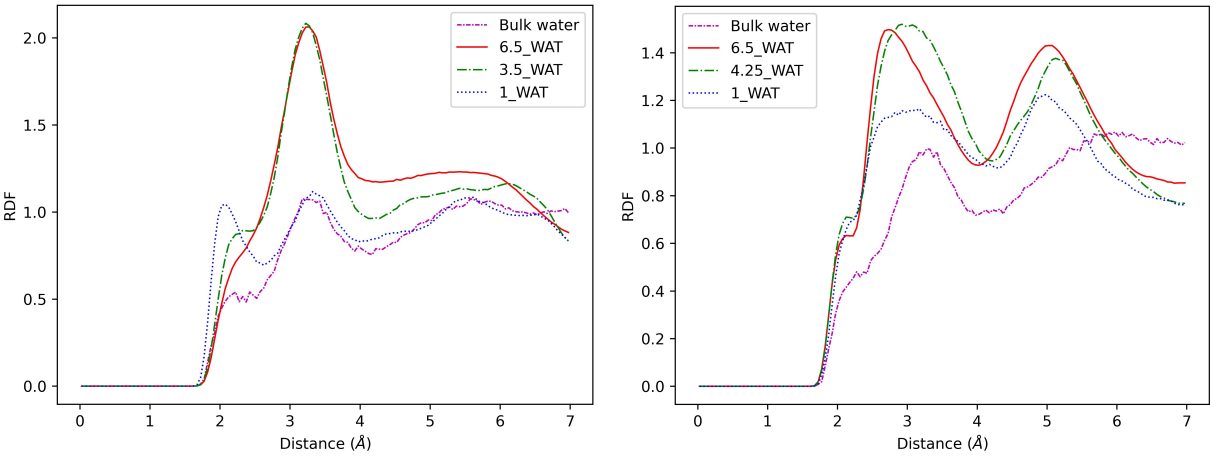
The results obtained in bulk water using the QM/MM strategy, reported in Chapter 2, are confirmed here, namely the absence of hydrogen bonds in which the H atoms of the amino group act as hydrogen bond donors on water (Figure 3.6), with the only exception being the system containing one water molecule per aspartate, in which some structuring at the first peak is observed, a possible consequence of larger confinement effects in a



(a) HOH...HNH interaction, ASP^{2-} (b) HOH...HNH interaction, GLU^{2-}

Figure 3.5: HOH...HNH pair correlation functions between water molecules and the organic intercalates.

smaller interlayer.

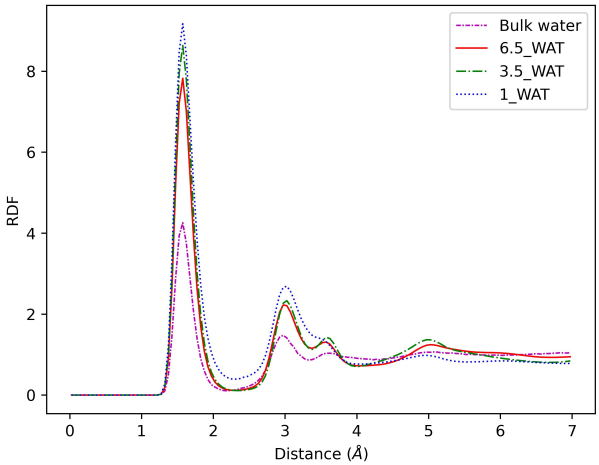


(a) H2O...HN interaction, ASP^{2-} (b) H2O...HN interaction, GLU^{2-}

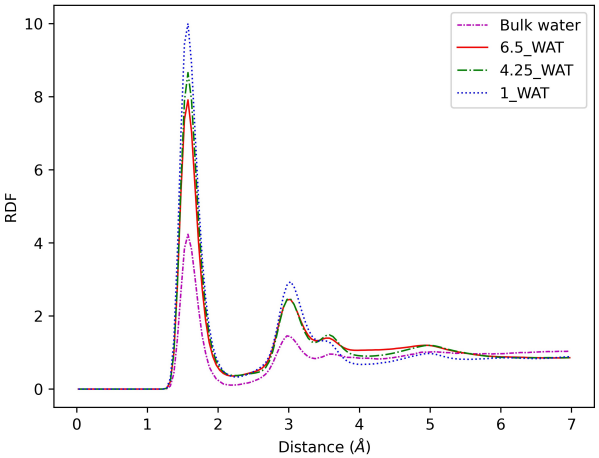
Figure 3.6: H2O...HN pair correlation functions between water molecules and the amino groups of organic intercalates.

Figure 3.7 shows the interactions between water molecules and the carboxylate groups of the organic intercalates. Our results are consistent with a very strong interaction between water molecules and the carboxylate groups and it is clear from the inspection of the first peaks that the structure of water around those groups is very similar (peaks shape,

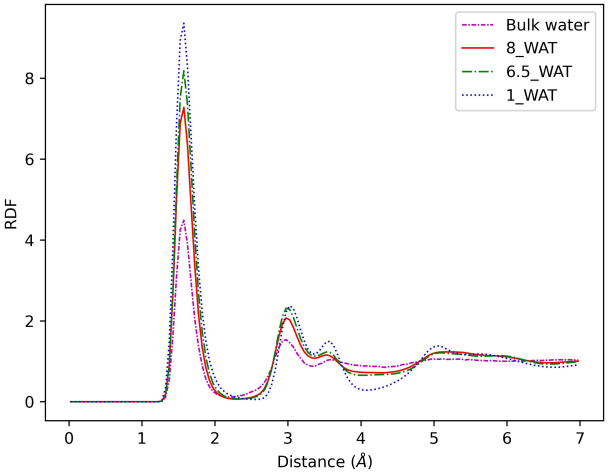
position) to that in bulk water, even at lower hydration. This can be safely interpreted as a fingerprint of a local hydrogen bond network in water that is well established in all situations.



(a) HOH...OOC interaction, ASP^{2-}



(b) HOH...OOC interaction, GLU^{2-}

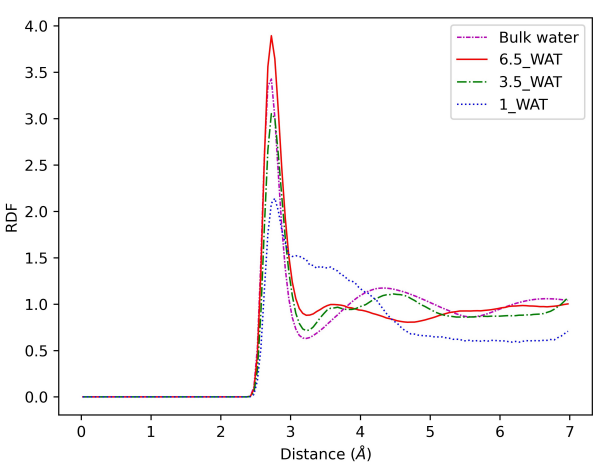


(c) HOH...OOC interaction, SUC^{2-}

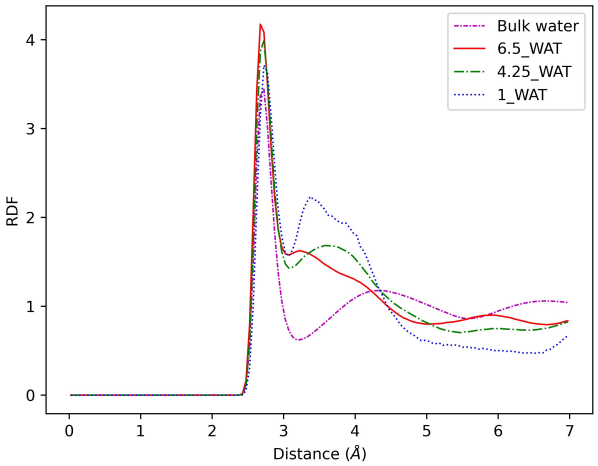
Figure 3.7: HOH...COO pair correlation functions between water molecules and the carboxylic groups of organic intercalates.

We turn now to water-water interactions in the interlamellar region. The first peak of the RDF is very similar in all cases (including bulk water) but one exception exists, namely the LDH intercalating one water per succinate. Two very structured but low

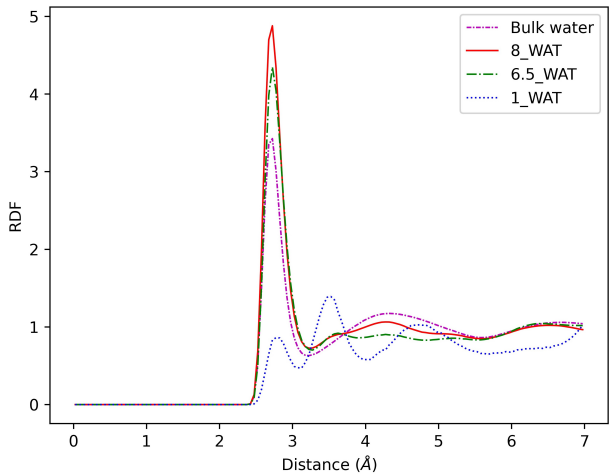
peaks are observed here within the distances corresponding to one interlamellar space (up to 5 Å). As for the second peak, none of the systems tend to have a peak shape similar to that in bulk water, and the minimum after the first peak is less pronounced at lower hydration.



(a) H2O...HOH interaction, ASP^{2-}



(b) H2O...HOH interaction, GLU^{2-}

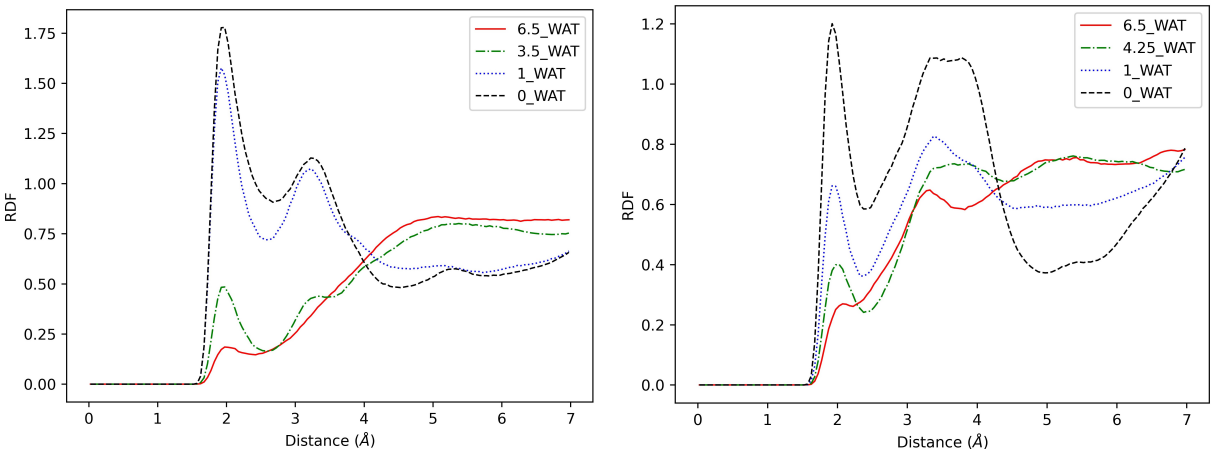


(c) H2O...HOH interaction, SUC^{2-}

Figure 3.8: H2O...HOH pair correlation function for the water-water interactions.

3.4.1.3 Interaction among amino acid intercalates

Our final RDF results focus on the interactions between closest neighbors along a path that has been considered the key precursor to the synthesis of peptide bonds in LDH.[31–33] The relevant RDF are thus the pair interactions between the oxygens of carboxylates on one unit and the hydrogens of the amino group of the other (Figure 3.9). Our results are well consistent with the hypothesis that lower hydration leads to a much stronger COO...HN interactions, bringing those units to the proximity that might finally lead to chemical reactivity (see first peaks for both amino acids). Differences between the behavior of the aspartate and glutamate upon increasing hydration are mostly observed in the shape and in the evolution of the second, much broader peak. This peak disappears at higher hydration conditions in the case of the LDH intercalating aspartate while it not only stays but becomes more structured in the case of glutamate. However, in both cases, the evolution of the RDF shapes can be interpreted as a sign of weaker proximity along the COO...HN direction, which is more sudden in the case of aspartate, and more regular in the case of glutamate (especially for the first peak). We recall here that this interpretation needs to be kept qualitative since we are referring to a non-homogeneous system.



(a) COO...HN interaction, ASP^{2-} (b) COO...HN interaction, GLU^{2-}

Figure 3.9: Intermolecular COO...HN pair correlation functions for the anion-anion interactions in the hybrid systems intercalating amino acids.

3.4.2 Orientation, conformations, binding modes of the anions in the interlayer

Once the local ordering has been illustrated employing RDF, we turn to the results describing the evolution of the orientation of the anions with increasing hydration. This is intimately related to the confinement effect and the increase of the size of the interlamellar region when more water is adsorbed.[19, 31]

We monitored the θ angle formed by the vector joining the two carbon bearing oxygen atoms (carboxylate groups) and the surface along the simulated trajectory. An illustration is reported in Figure 3.10. As a matter of fact, when we move from succinate to aspartate and glutamate, the configurational space becomes increasingly complex, and it is harder to define a clear descriptor of the orientation of the molecule. However, we showed that this angle still gives qualitative information on the changes experienced by the anions.

To rationalize our results, we present them in terms of populations of orientational

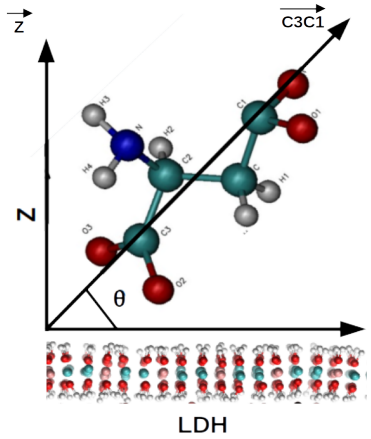


Figure 3.10: Definition of the θ angle used to describe the orientation of the anions in the case of ASP^{2-}.

states, computed as percentages. There are three orientational states of organic intercalates (1) parallel (-10.0° – 10.0°), (2) intermediate orientation (-80.0° – -10.0°and 10.0° – 80.0°), and (3) perpendicular (-90.0° – -80.1°and 80.1° – 90.0°).

Table 3.2: Orientational states for the anions as a function of increasing hydration.

Systems	**Parallel (%)**	**Intermediate orientation (%)**	**Perpendicular (%)**
ASP^{2-}			
0_WAT	55	45	0
1_WAT	39	61	0
3.5_WAT	9	69	22
6.5_WAT	32	37	31
GLU^{2-}			
0_WAT	96	4	0
1_WAT	38	62	0
4.25_WAT	30	70	0
6.5_WAT	25	74	1
SUC^{2-}			
0_WAT	100	0	0
1_WAT	0	41	59
6.5_WAT	15	33	52
8_WAT	26	33	41

The most remarkable difference is observed between succinate and the two amino

acids, moving from the anhydrous system to the system having one water molecule per anion. Succinate anions move from a situation in which they are lying horizontally with respect to the surface to being either perpendicular or in an intermediate state. Aspartate and glutamate, on the other hand, do not have enough space to move to a perpendicular orientation. However, further differences are observed between aspartate and glutamate. In the former case, anions can be either horizontal or in an intermediate state even in the anhydrous system, which is not possible for glutamate, probably due to the larger structure of the molecule. As soon as the interlamellar space increases, both amino acids are in comparable situations.

Let us switch now to the two higher hydration states. Aspartate can be found in different orientation states (meaning that perpendicular molecules co-exist with intermediate orientations or even parallel molecules), whereas it would seem as if glutamate can either be parallel or in an intermediate state. However, at this stage, the picture might become blurry, given the more complicated structure and conformational space of glutamate. As for succinate, in higher hydration states all orientations are possible. On one hand, our results are in agreement with the very few studies that addressed changes in the orientation of intercalates with increasing water content.[31] On the other hand, in our systematic study based on the experimental adsorption isotherms, we learn that some molecules can still be found in a horizontal orientation, and this population is not negligible.

Some additional qualitative information on the effect of the interlayer evolution upon hydration can be provided by observing the corresponding evolution of the conformational

space of the intercalates. The dihedral distributions of the C-C-C-C dihedral in SUC^{2-} and the C_α-C_β-C_γ in the amino acids are reported in the Appendix Section B.3, Figures B.6, B.7, B.8. In the case of aspartate, gauche/distorted gauche configurations are preferred with respect to anti in the anhydrous system. Such a configuration is very much reduced in the lowest hydration state but becomes important again in the two higher hydration states. For the latter, results are very well consistent with the bulk results reported in Chapter 2. An illustration of the different conformations observed for aspartate is displayed in Figure 3.11. When the interlamellar space is reduced as in the anhydrous system, there is not enough room for the system to adopt an anti configuration. One consequence is that we can observe different kinds of binding to the surface. In the case

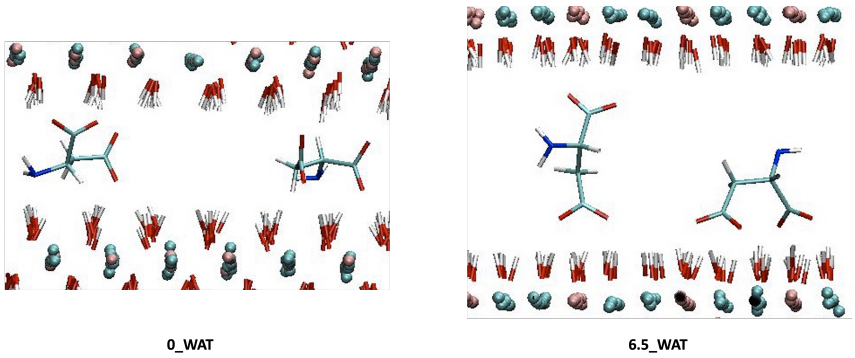


Figure 3.11: Snapshot representing conformations of ASP^{2-} intercalated in LDH at two different hydration states.

of glutamate, the anhydrous system is the only case where the anti conformation is the prevalent one, whereas when water is present, both anti and gauche are present (see Figure 3.12). It is important to stress a major difference here. In the case of aspartate, the proximity of the amino group with the carboxylate group at the C_β carbon and the resulting steric hindrance only allow one gauche/distorted gauche orientation ($\chi \sim$ -60°). On the other hand, the presence of a longer side chain in glutamate allows recovering the

two possibilities (χ = -60° and χ = 60°).

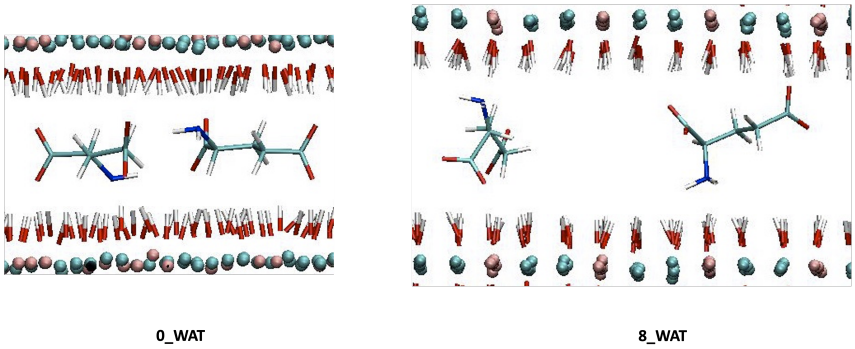


Figure 3.12: Snapshot representing conformations of GLU^{2-} intercalated in LDH at two different hydration states.

As for SUC^{2-}, the gauche configuration is not observed in the anhydrous system and in the lowest hydration state, where it gets slightly populated in the highest hydration states (an illustration is given in Figure 3.13), which is quite different with respect to the bulk behavior (Chapter 2).

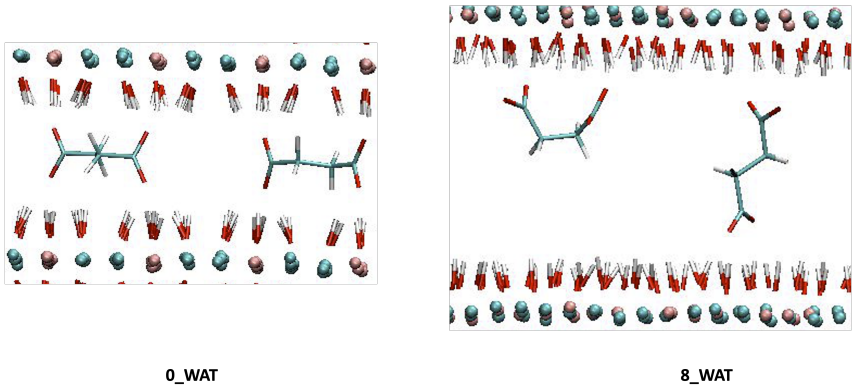


Figure 3.13: Snapshot representing conformations of SUC^{2-} intercalated in LDH at two different hydration states.

The analysis of the conformations adopted by the anions shows that each anion has its specific ability to interact with the surface and react to a situation under confinement. We found an analogy of the binding modes of the carboxylate groups with the binding modes

of the same groups when coordinating to metal, and indeed the different conformations adopted in the interlamellar region lead to differences in such binding modes. This analogy is presented in Figure 3.14. For the anhydrous systems, when the anions only have enough room to lie horizontally and the favored conformation is anti for SUC^{2-} and GLU^{2-}, such anions are found in a bidentate bridging-like conformation. For one carboxylate group, each oxygen atom interacts with a different surface (by analogy with the bidentate bridging arrangement with a metal). In the case of ASP^{2-}, given the prevalent gauche conformation, when one carboxylate group adopts this mode, the other one cannot and can be found either in a monodentate interaction mode or in a bidentate chelating-like mode, in which both oxygen atoms interact with the same surface. When the water amount increases, the interlamellar space becomes larger. However, in the case of succinate, this happens quite abruptly. Already in the case of the 1_WAT system, all the observed molecules are found in a vertical arrangement and this population stays in all hydration states. Molecules belonging to this population can only be found in a bidentate chelating-like binding mode. Let us move now to ASP^{2-}. When the water content increases, the reorientation of the anions occurs smoothly and the population of molecules lying horizontally is present in all hydration states. The monodentate and bidentate chelating-like binding modes are observed in all such states. Finally, in the case of GLU^{2-}, the anion never completely reorients, and a larger distribution of accessible orientational states is related to larger conformational mobility. Therefore, the binding modes of the anions do not experience large changes with increasing hydration.

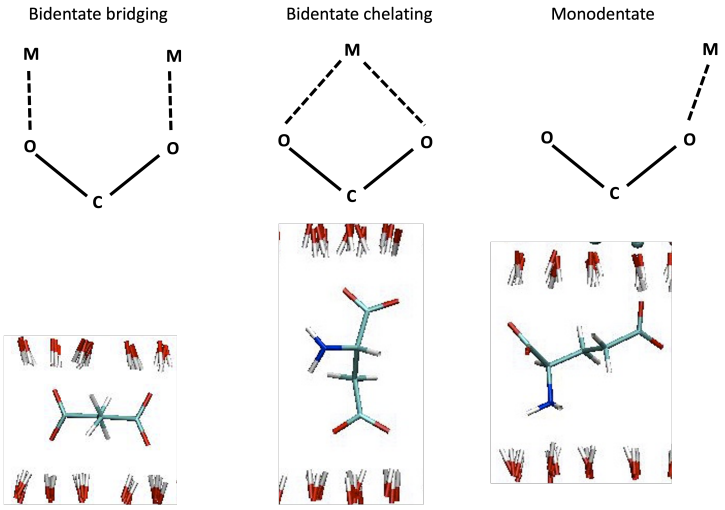
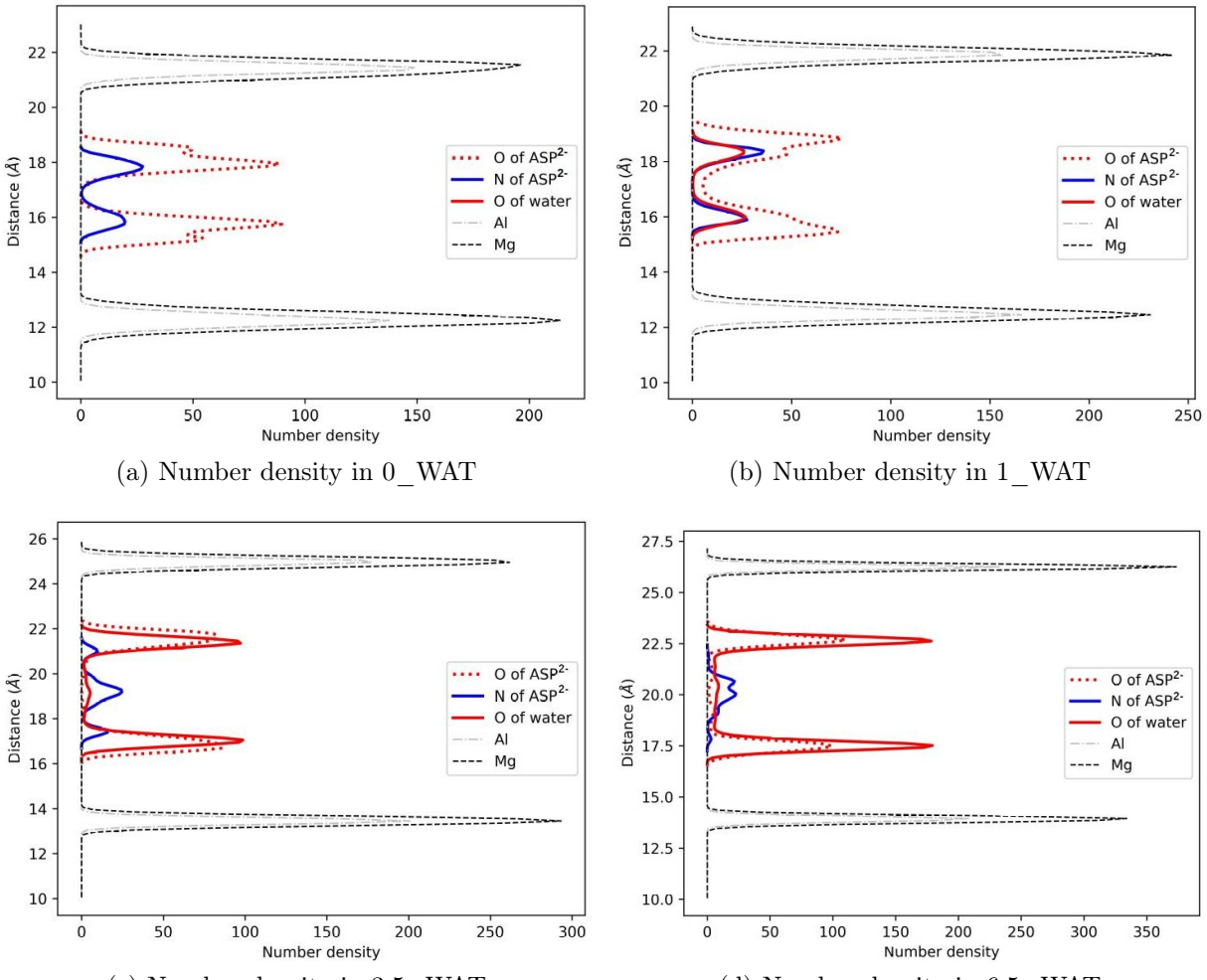


Figure 3.14: Binding modes of carboxylates coordinated with a metal M and analogy with binding modes of the confined anions with the surface.

3.4.3 Number density along the z-axis

The number density distribution of some relevant atoms along the axis perpendicular to the surface (z-axis) provides complementary information about structuring in the interlamellar space compared with the RDF. In Figure 3.15 we report the results obtained for the LDH intercalating aspartate. The distributions of the z positions of the cations are displayed as a reference. In the absence of water, the carboxylate oxygen atoms and the nitrogen atoms from the amino groups sit next to the surface. This situation stays quite similar in the system where one water molecule per anion is present. This is consistent with similar distributions of orientational states observed in the previous Section. The z position of the water oxygen atoms matches one of the N atoms, which is consistent with water molecules solvating the amino group. On the other hand, moving to higher hydration states results in N atoms being also present in the middle of the layer and progressively moving 'away' from the water O atoms. Such atoms are found in close prox-

imity with the carboxylate O atoms, instead, a finding that is fully consistent with the reported affinity of water for the carboxylate groups in LDH.[16, 17] A small population of water molecules also appears in the middle of the interlayer, possibly sitting next to the carboxylate groups of the anions oriented horizontally with respect to the surface.



(a) Number density in 0_WAT

(b) Number density in 1_WAT

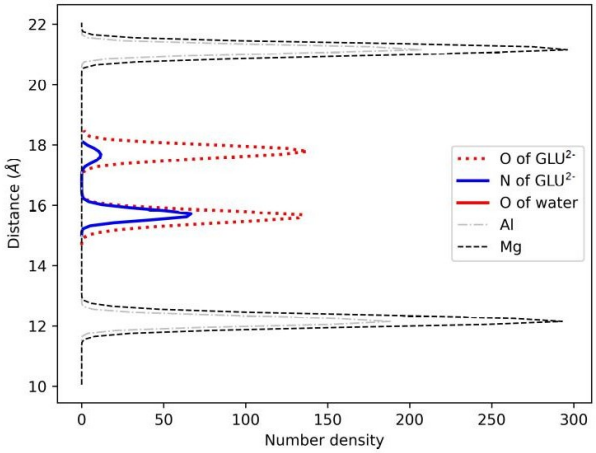
(c) Number density in 3.5_WAT

(d) Number density in 6.5_WAT

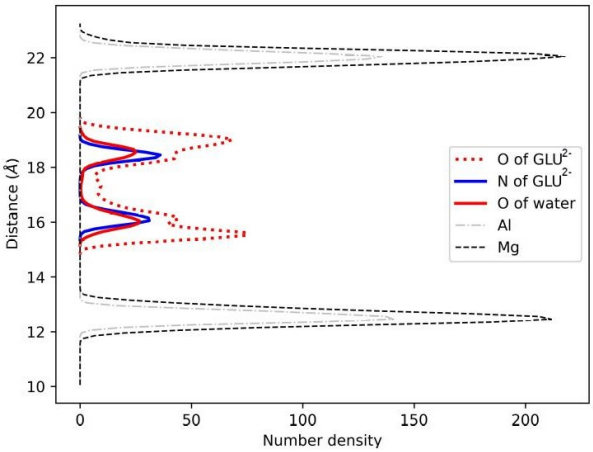
Figure 3.15: Number density along the z-axis for increasing water content in the system intercalating aspartate.

A similar trend is observed for the LDH intercalating glutamate (Figure 3.16). However, we observe here an important population of O atoms from the carboxylate groups in the middle of the interlayer, especially in the two highest hydration states. The other

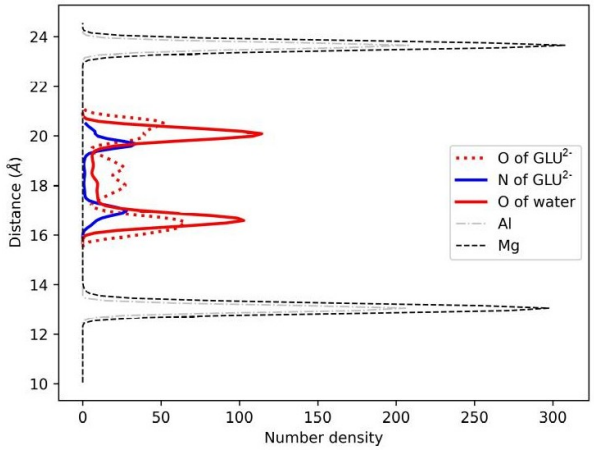
major difference is observed in the distribution of the N atoms: in the case of glutamate, a population corresponding to N atoms in the middle of the interlayer is not seen. An artifact is observed for the anhydrous systems, for which the two peaks corresponding to N atoms next to the two surfaces have different heights, possibly due to insufficient sampling.



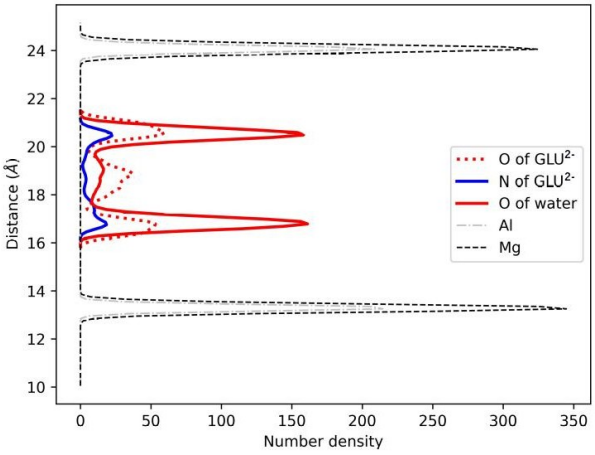
(a) Number density in 0_WAT



(b) Number density in 1_WAT



(c) Number density in 4.25_WAT

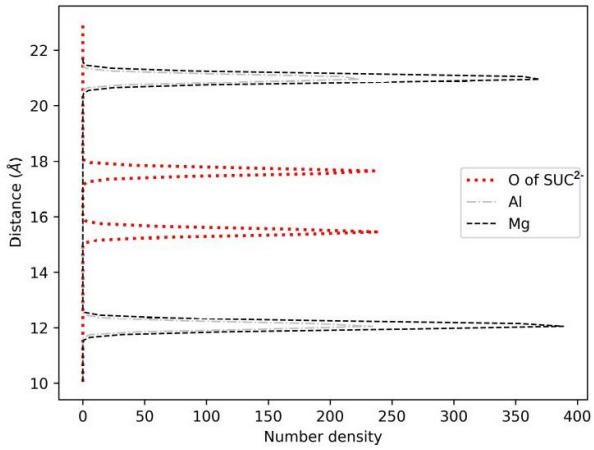


(d) Number density in 6.5_WAT

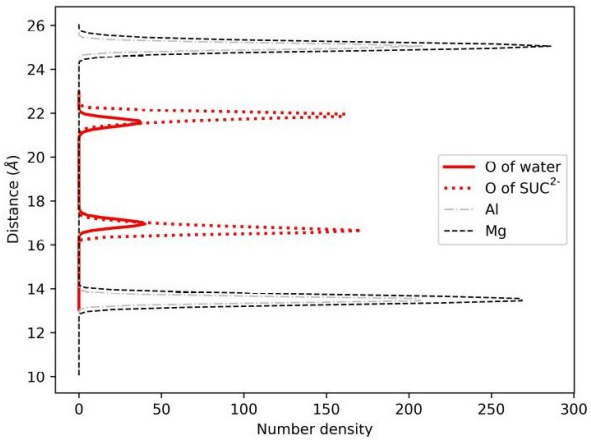
Figure 3.16: Number density along the z-axis for increasing water content in the system intercalating glutamate.

We conclude this discussion by presenting the number density along z in the case of the system intercalating succinate (Figure 3.17). The situation is much simpler here than

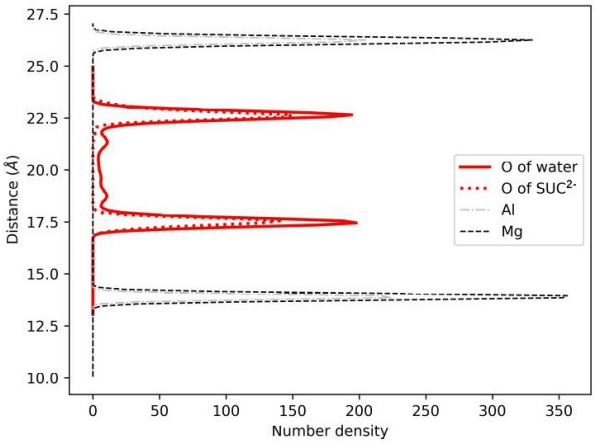
in the case of the amino acids, since we exclusively observe the peaks of the carboxylates next to the surface, and water sitting next to them in all hydration states considered, with a small population appearing in the middle of the interlayer at higher hydration states.



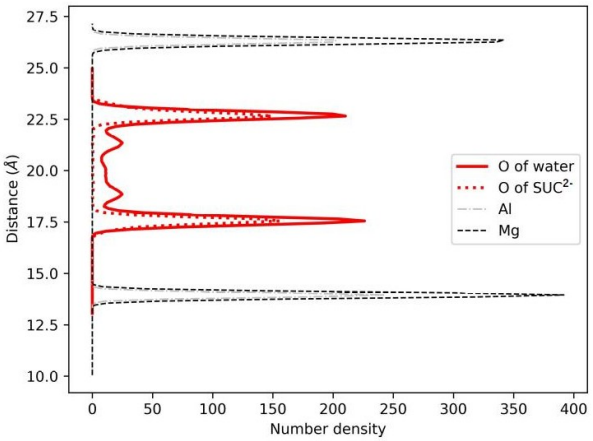
(a) Number density in 0_WAT



(b) Number density in 1_WAT



(c) Number density in 6.5_WAT



(d) Number density in 8_WAT

Figure 3.17: Number density along the z-axis for increasing water content in the system intercalating succinate.

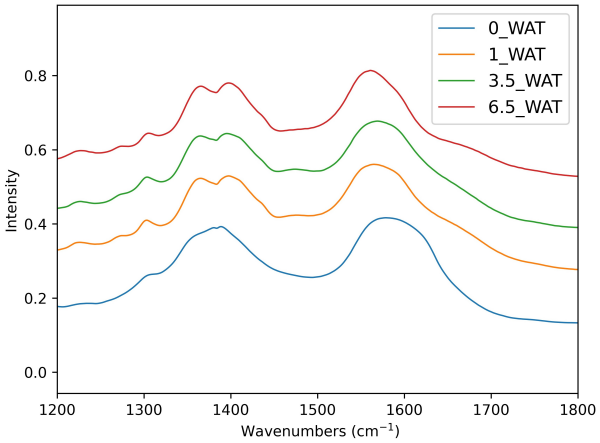
3.4.4 Vibrational properties of intercalated anions

In this Section, we shall provide a description of the evolution of the carboxyl stretch band of the three intercalated carboxylates with increasing hydration, to compare with

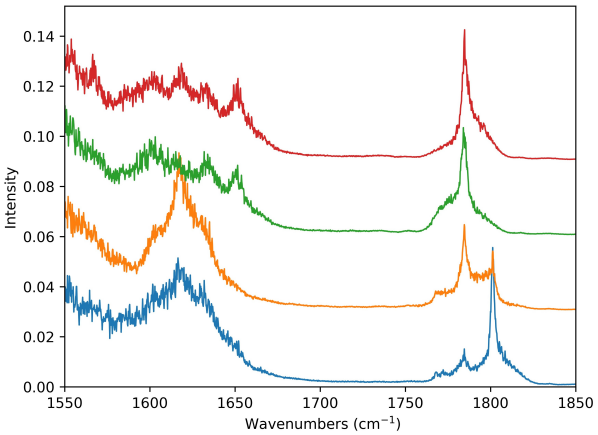
experimental measurements of the IR spectra and bring a molecular interpretation. Two different methods were used to compute the vibrational spectra, one based on the vibrational density of states (VDOS) and the other on the molecular dipole time correlation function (TCF). We recall that these methods are briefly described in Chapter 1.

In Figure 3.18, we report the collection of our results and a comparison with the corresponding experimental IR measurements for the symmetric (S) and antisymmetric (AS) carboxyl stretch band of the aspartate anion intercalated into LDH at various hydration contents. The assignment of the bands was confirmed by analyzing the Fourier transform of atomic velocities of the C and O atoms of the carboxyl groups. This preliminary analysis is not reported here for the sake of simplicity. The position of the computed band is strongly blue-shifted (by about 200 cm^{-1}) compared with the experimental band. The value of the shift is thus higher for the force field based dynamics than it is for the PM6/MM dynamics presented in Chapter 2. In addition, the band shape is not well reproduced and we already commented on this in that Chapter. However, the AS-S stretch band gap is in quite good agreement with the experimental one. In addition, the trend with increasing hydration nicely follows the experimental trend. We see the AS band progressively red shifting and the S band slightly becoming broader. The differences among the results obtained with different computational methods are minor, especially if we consider that it is not possible to propose a clear interpretation of the different (sub)peaks appearing. For instance, in the AS, one would see two to three peaks, but we do not have a convincing argument not to consider them just as noise.

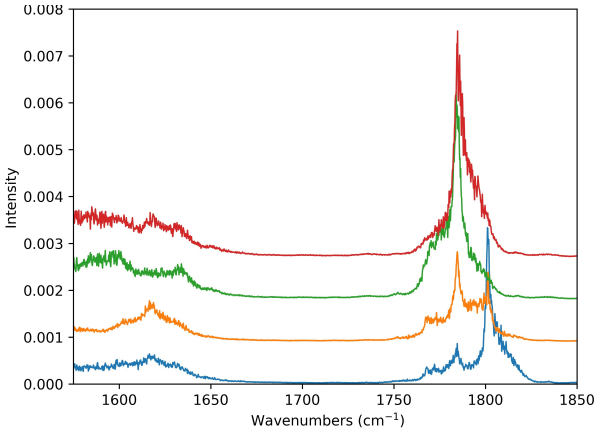
A similar analysis can be made for the carboxylate stretch band of intercalated glu-



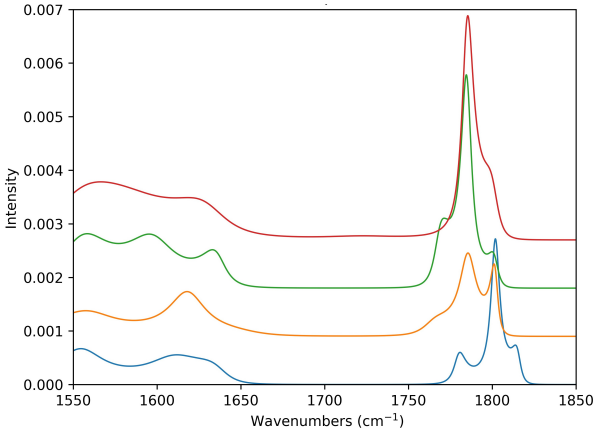
(a) Carboxylate stretch band of ASP^{2-} in LDH from experiments.



(b) Carboxylate stretch band of ASP^{2-} in LDH as computed from VDOS.



(c) Carboxylate stretch band of ASP^{2-} in LDH as computed from the dipole moment TCF, WK-FFT method.

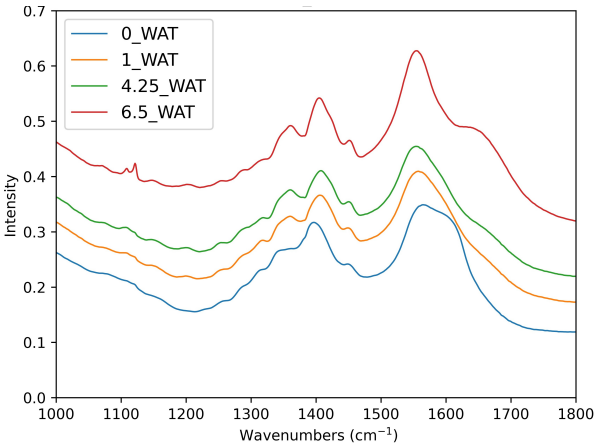


(d) Carboxylate stretch band of ASP^{2-} in LDH from the dipole moment TCF, MEM method.

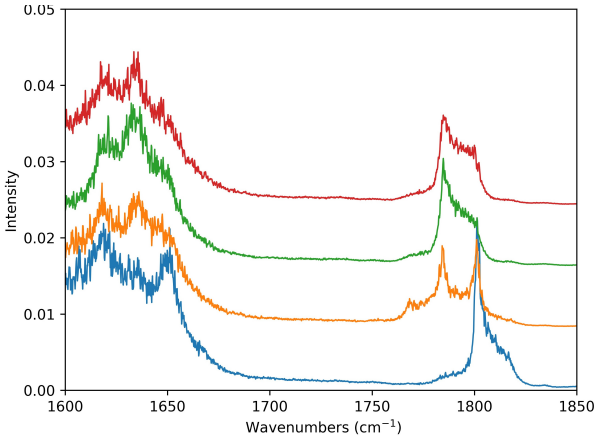
Figure 3.18: Experimental and simulated carboxylate stretch band of ASP^{2-} intercalated into LDH.

tamate, reported in Figure 3.19. The experimental S band becomes lightly blue shifted with increasing hydration, which is also present in the computed band. In addition to becoming red shifted, the experimental AS band presents a shoulder at higher hydration. In the computed band, this might correspond to the observed broadening.

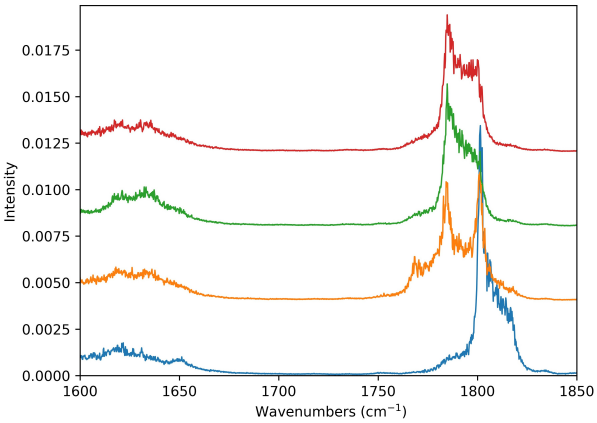
The evolution of the band upon hydration measured by experiments is quite striking. We observe a moderate blue shift of the S band and a red shift accompanied by a pro-



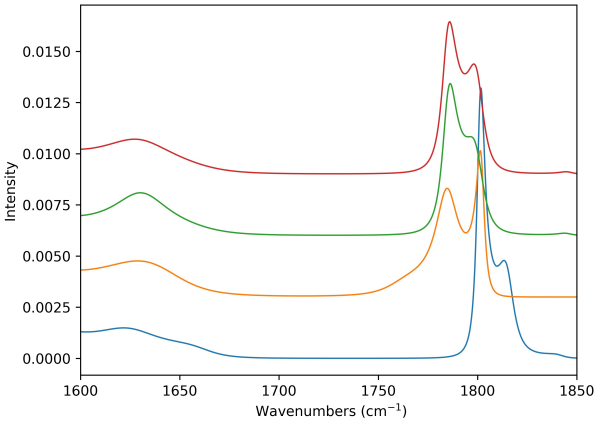
(a) Carboxylate stretch band of GLU^{2-} in LDH from experiments.



(b) Carboxylate stretch band of GLU^{2-} in LDH as computed from VDOS.



(c) Carboxylate stretch band of GLU^{2-} in LDH as computed from the dipole moment TCF, WK-FFT method

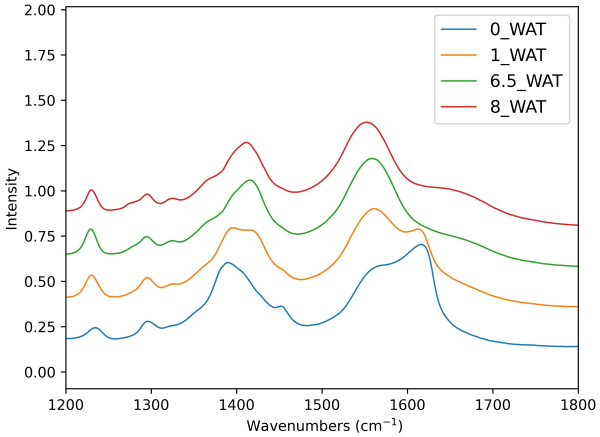


(d) Carboxylate stretch band of GLU^{2-} in LDH as computed from the dipole moment TCF, MEM method.

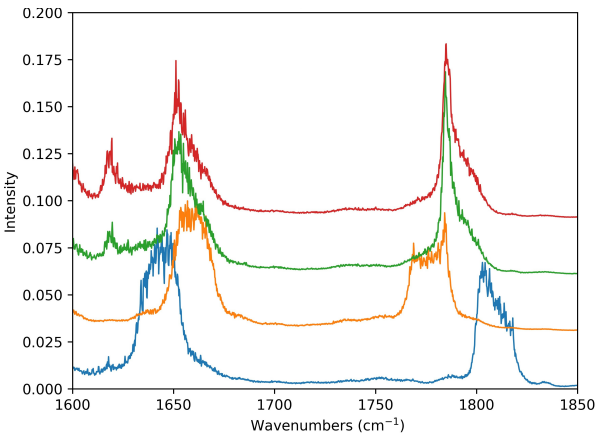
Figure 3.19: Experimental and simulated carboxylate stretch band of GLU^{2-} intercalated into LDH.

nounced change in the band shape. The qualitative trend of the two bands is also seen in the computed spectra, the effect being overestimated in the case of the 1_WAT system.

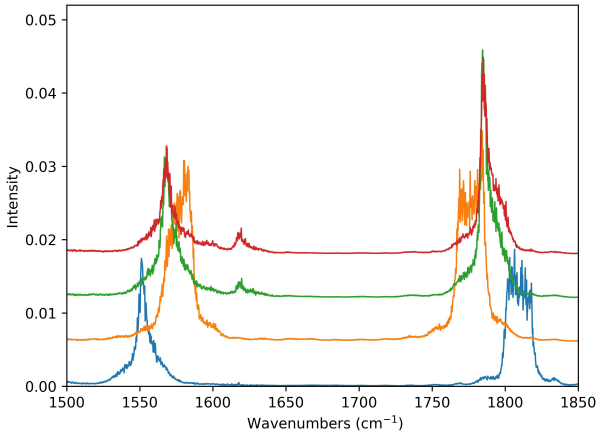
Our interpretation of the observed trend for the vibrational spectra and of the differences observed in the three systems considered is based on the structural analysis provided in terms of reorientation of the anions, accessible conformations, and binding modes of the carboxylate units. We recall that, in the case of SUC^{2-}, the anions migrate from a



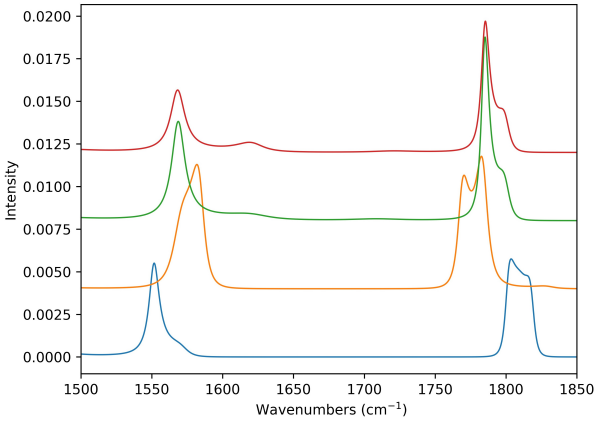
(a) Carboxylate stretch band of SUC^{2-} in LDH from experiments.



(b) Carboxylate stretch band of SUC^{2-} in LDH as computed from VDOS.



(c) Carboxylate stretch band of SUC^{2-} in LDH as computed from the dipole moment TCF, WK-FFT method.



(d) Carboxylate stretch band of SUC^{2-} in LDH as computed from the dipole moment TCF, MEM method.

Figure 3.20: Experimental and simulated Carboxylate stretch band of SUC^{2-} intercalated into LDH.

situation in which they lie horizontally and interact with the surface through bidentate bridging-like modes to a situation in which they are vertical with respect to the surface, and they interact with bidentate chelating-like modes.

In the 1980s, Deacon and Phillips provided an extensive investigation of the relative position of the anti-symmetric and symmetric stretch bands for complexes based on acetate and trifluoroacetate complexes.[34] The AS-S band gap was correlated with the

type of binding: compared with the free anion, a larger band gap band would be observed for monodentate binding, a smaller band gap for bidentate chelating and a similar band gap would be observed for bidentate bridging modes. Those criteria have been extensively used in the literature to assign the type of binding of a wide variety of compounds. However, the authors themselves warned against applying this classification without caution, especially in the case of complex systems and for different experimental conditions. Among more recent studies, we chose to resume two that represent the kind of controversy that is still related with the correlation between the band gap and the binding mode. In 2015, Sutton et al.[35] performed a computational study of complexes of acetate and formate with different cations in an aqueous environment. For these carboxylates, the predictions by Deacon and Phillips were fulfilled but the main correlation was between the band gap and the geometrical parameters of the carboxylate groups. As a matter of fact, the distortion of the C-O bond and of the O-C-O angle upon complexation would explain the evolution of the band gap with the different modes of binding. On the other hand, in 2010 Martinez et al.[36] performed an experimental investigation of the AS-S band gap for manganese complexes of carboxylates having a more complicated structure and their conclusions do not support the prediction of the evolution of the band gap with binding modes, whereas they do prove that the experimental conditions are key to correctly assign the C-O stretch band since a change in the coordination sphere of the metal can be observed.

Although a straightforward assignment does not seem universally accepted, we can safely assume that the AS-S stretch band gap is sensitive to the biding modes. In the

following, we shall describe the interpretation that we propose for the evolution of this region of the IR spectrum based on our structural analysis that we performed by means of MD simulations upon water uptake and for the three different intercalated systems. When the hydration increases, some anions are found lying horizontally, but they do not come back to an environment that leads to bidentate bridging-like modes since those anions are sitting next to one surface. The carboxylate units become increasingly hydrated, and the shoulder that appears on the higher frequency side of the AS stretch band might be due to these two populations of anions experiencing different local environments. For ASP^{2-} and GLU^{2-}, the evolution of the binding modes is much smoother, corresponding to a smoother evolution of the bands upon hydration.

3.5 Conclusions

In this Chapter, we presented a molecular dynamics study of the properties of LDH hybrid materials intercalating aspartate, succinate, and glutamate anions. We employed the ClayFF for modeling the LDH, GAFF for organic intercalates, and SPC/E for water molecules. In particular, we modeled the anhydrous systems and three different hydration states corresponding to the experimental conditions under which the measurements of the interlamellar space, the number of water molecules per anion, and the infrared spectra were obtained by our experimental collaborators.

The analysis of radial distribution functions, restricted to distances corresponding to the size of the interlamellar space, as well as the number distribution of significant

atoms along the z-axis (perpendicular to the surface) provided information on the most important intermolecular interactions that are active in the confined systems, on the position of the polar groups, and on how they evolve with increasing water content. The interaction of carboxylate units with the hydroxyl groups on the surface is strong but water also has a strong affinity to form COO...HOH bonds, which results in an accumulation of water molecules at the surface. The amino group of the amino acids can act as a hydrogen bond acceptor, from the surface, and from water, but in anhydrous conditions, it acts as a hydrogen bond donor to the carboxylate group of a nearby anion. The latter situation becomes increasingly less favorable with increasing hydration, a phenomenon that was already mentioned in the literature and claimed to be a possible path to the synthesis of peptide bonds in abiotic conditions and low water content.[31, 32]

As a matter of fact, the interlamellar space increases in all systems to accommodate an increasing number of water molecules. However, some major differences are observed in the system intercalating succinate, where the small size of the molecule allows a sudden reorientation (from parallel to the surface to vertical orientation) as soon as some water is adsorbed. On the other hand, this reorientation occurs more smoothly for aspartate and there is no complete reorientation for glutamate in the interlayer. Another descriptor that is sensitive to the changes in the confined region is the dihedral describing the conformational space of the molecules. For SUC^{2-}, only an anti conformation is possible when the interlamellar region is the smallest (i.e. in the absence of water) and when the water content increases some molecules adopt a gauche conformation as well. In the case of aspartate, evolution is observed toward a conformational equilibrium similar to that

observed in bulk QM/MM dynamics (Chapter 2). Finally, a tendency to a more complex conformational space is observed in glutamate, for which the side chain is longer.

The computation of the infrared carboxylate stretch band using two different methods (power spectra based on VDOS and on the molecular dipole moment) brought a very similar picture. Compared with the experimental ones, though the band position and shape are poorly reproduced, the S and AS band gap and its evolution with increasing hydration are nicely accounted for. This allows us to make some hypotheses on the differences observed among the intercalates. The very strong effect measured for succinate is related to the drastic reorientation effects and their consequences on the binding modes on the surface. For aspartate and succinate, the reorganization follows a smoother path and the spectra evolve more regularly. The observed shifts of the (sub)bands and some broadening can be explained in terms of a more flexible conformational space and of increasing solvation of the carboxylate groups.

References

(1) V. Rives, *Mater. Chem. Phys.*, 2002, **75**, 19–25.

(2) R. T. Cygan, J.-J. Liang and A. G. Kalinichev, *J. Phys. Chem. B*, 2004, **108**, 1255–1266.

(3) A. G. Kalinichev, X. Liu and R. T. Cygan, *Clays and Clay Minerals*, 2016, **64**, 335–336.

(4) E. Manasse, *Atti Soc. Toscana Sci. Nat*, 1915, **24**, 92.

(5) R. Allmann, *Acta Crystallogr B Struct Sci*, 1968, **24**, 972–977.

(6) H. F. W. Taylor, *Mineral. mag.*, 1969, **37**, 338–342.

(7) H. F. W. Taylor, *Mineral. mag.*, 1973, **39**, 377–389.

(8) A. J. Marchi and C. R. Apesteguía, *Applied Clay Science*, 1998, **13**, 35–48.

(9) *Layered Double Hydroxides*, ed. X. Duan and D. G. Evans, Springer, Berlin ; New York, 2005.

(10) F. L. Theiss, G. A. Ayoko and R. L. Frost, *Appl. Surf. Sci.*, 2016, **383**, 200–213.

(11) F. Cavani, F. Trifirò and A. Vaccari, *Catal*, 1991, **11**, 173–301.

(12) I. C. Chisem and W. Jones, *J. Mater. Chem.*, 1994, **4**, 1737–1744.

(13) G. W. Brindley and S. Kikkawa, *Clays Clay Miner.*, 1980, **28**, 87–91.

(14) S. Miyata, *Clays Clay Miner.*, 1983, **31**, 305–311.

(15) C. Depège, F.-Z. El Metoui, C. Forano, A. de Roy, J. Dupuis and J.-P. Besse, *Chem. Mater.*, 1996, **8**, 952–960.

(16) S. P. Newman, T. Di Cristina, P. V. Coveney and W. Jones, *Langmuir*, 2002, **18**, 2933–2939.

(17) P. P. Kumar, A. G. Kalinichev and R. J. Kirkpatrick, *J. Phys. Chem. C*, 2007, **111**, 13517–13523.

(18) A. Tsukanov and S. Psakhie, *Sci. Rep.*, 2016, **6**, 1–8.

(19) A. G. Kalinichev, P. Padma Kumar and R. James Kirkpatrick, *Philos. Mag.*, 2010, **90**, 2475–2488.

(20) B. Grégoire, V. Erastova, D. L. Geatches, S. J. Clark, H. C. Greenwell and D. G. Fraser, *Geochim. Cosmochim. Acta*, 2016, **176**, 239–258.

(21) V. Erastova, M. T. Degiacomi, D. G. Fraser and H. C. Greenwell, *Nat Commun*, 2017, **8**, 2033.

(22) J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem*, 2004, **25**, 1157–1174.

(23) Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang and P. Kollman, *J. Comput. Chem*, 2003, **24**, 1999–2012.

(24) H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren and J. Hermans, in *Intermolecular Forces*, ed. B. Pullman, Springer Netherlands, Dordrecht, 1981, vol. 14, pp. 331–342.

(25) D. Case, R. Betz, D. Cerutti, T. Cheatham, III, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, D. Roe, A. Roitberg, C. Sagui, C. Simmerling, W. Botello-Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao and P. Kollman, *AMBER 2016*, University of California, San Francisco, 2016.

(26) W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.

(27) A. R. Leach, *Molecular Modelling: Principles and Applications*, Prentice Hall, Harlow, England ; New York, 2nd ed, 2001.

(28) V. Kräutler, W. F. van Gunsteren and P. H. Hünenberger, *J. Comput. Chem*, 2001, **22**, 501–508.

(29) T. A. Andrea, W. C. Swope and H. C. Andersen, *J. Chem. Phys.*, 1983, **79**, 4576–4584.

(30) H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

(31) B. Grégoire, V. Erastova, D. L. Geatches, S. J. Clark, H. C. Greenwell and D. G. Fraser, *Geochim. Cosmochim. Acta*, 2016, **176**, 239–258.

(32) V. Erastova, M. T. Degiacomi, D. G. Fraser and H. C. Greenwell, *Nat. Commun.*, 2017, **8**, 2033.

(33) B. Grégoire, H. C. Greenwell and D. G. Fraser, *ACS Earth Space Chem.*, 2018, **2**, 852–862.

(34) G. Deacon and R. Phillips, *Coord. Chem. Rev.*, 1980, **33**, 227–250.

(35) C. C. Sutton, G. da Silva and G. V. Franks, *Chem. Eur. J..*, 2015, **21**, 6801–6805.

(36) D. Martıḿnez, M. Motevalli and M. Watkinson, *Dalton Trans.*, 2010, **39**, 446–455.

Chapter 4

Modeling the Optical Properties of 2,2'-Bipyridine-3,3'-Diol in Water

Contents

4.1 Introduction . . . 136

4.2 Integrated multilevel QM/MM computational approach . . . 138

4.2.1 QM calculations . . . 139

4.2.2 Intramolecular FF parametrization with Joyce . . . 139

4.2.3 MD simulations . . . 140

4.2.4 Snapshot extraction and calculation of the UV-VIS spectra . . . 141

4.3 Results and discussion . . . 142

4.3.1 QM calculations . . . 142

4.3.2 FF parametrization . . . 145

4.3.3 MD simulations . . . 148

4.3.4 Absorption spectra . . . 151

4.4 Preliminary conclusions . . . 154

4.1 Introduction

2,2'-bipyridine-3,3'-diol (BPOH2) is a planar molecule containing two aromatic units, as has been revealed by crystal structure analysis.[1, 2] It has two intramolecular hydrogen bonds in its diol tautomeric form. Upon photo-excitation, excited-state intramolecular double proton transfer (ESIDPT) takes place to generate a di-zwitterion (diketo tautomer).[3] Femtosecond fluorescence up-conversion and transient absorption experiments provide the mechanistic details of ESIDPT of BPOH2,[4, 5] and the excited-state proton transfer dynamics of BPOH2 has been extensively investigated from experimental and theoretical points of view.[3, 6–10]

In this project, we focused on the exploration and parametrization of the potential energy surface of BPOH2. From now on, we shall thus discuss the ground state properties of the molecule exclusively.

In 1995, Barone and co-workers showed that three energy minima might exist on the potential energy surface of BPOH2, corresponding to dienol (diol), keto-enol, and diketo forms (Figure 4.1). However, based on the computed relative energies with respect to dienol (diol), the diketo and keto-enol tautomers are supposed to be transient forms, thus preventing their experimental characterization.[11]

The absorption maximum of BPOH2 is centered at 340 nm and is assigned to the $\pi - \pi^*$ transition of its diol tautomeric form. In neat water, some less intense absorption bands appear between 400 and 450 nm, and they have been claimed to be due to the absorption from the diketo tautomer ground state since water would stabilize this form.[8,

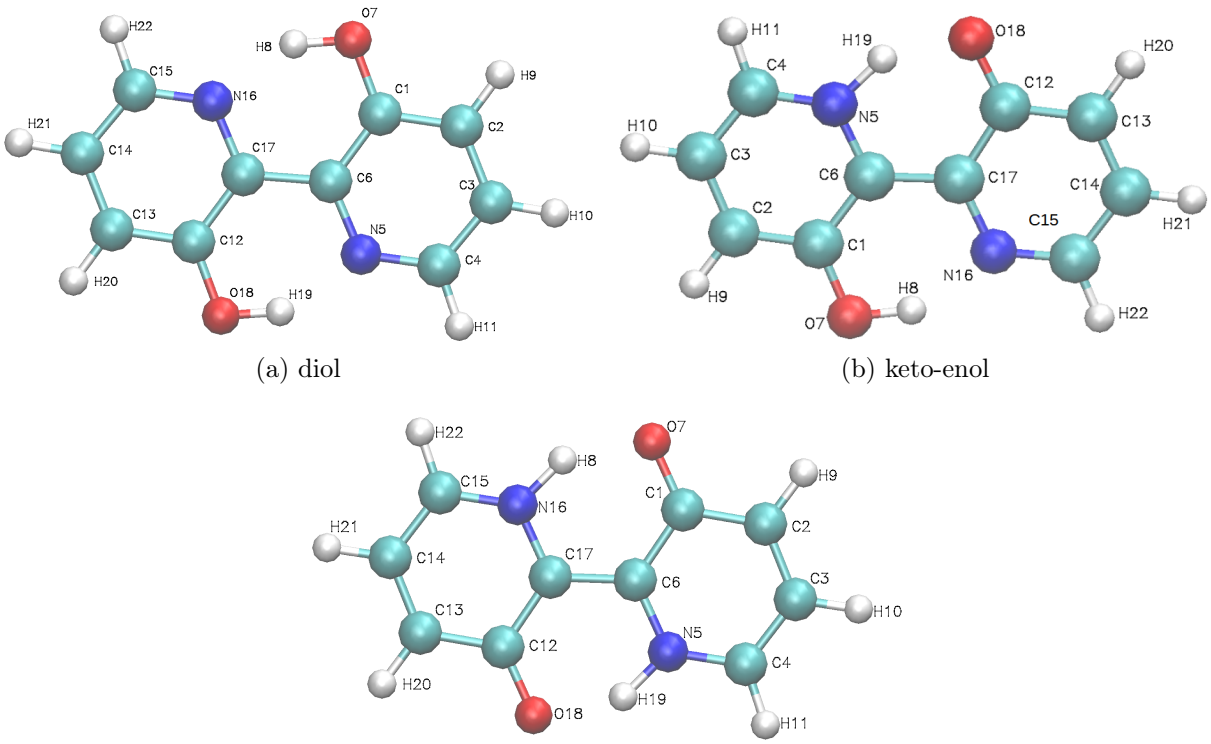


Figure 4.1: Structures and atom numbering of the three tautomers of BPOH2.

9, 12]

Experimental work using BPOH2 as a probe showed that its absorption and fluorescence spectra in the presence of nanocavities (cyclodextrins) in an aqueous solution could provide information about the degree of hydrophobicity of the host and the different mechanisms of encapsulation.[12] However, so far, this interpretation is simply based on conjectures. Due to the presence of both inter- and intra-molecular hydrogen bonds in water, calculations of spectra without the explicit inclusion of the environment would fail to model such systems. On the other hand, DFT-based dynamics or QM/MM methods would be too expensive, given the size of the system. We thus chose to employ a method that we recently applied to an organometallic compound, following a sequential classical-

QM approach based on Joyce.[13, 14] The computational details of such an approach will be presented in the next Section.

4.2 Integrated multilevel QM/MM computational approach

The computational strategy that we have followed is developed in collaboration with Dr. G. Prampolini, who is one of the co-authors of the Joyce protocol (see Chapter 1).[15, 16] The basic idea is to use quantum chemistry calculations on the target molecule to determine a QM profile for the potential energy surface (PES), to be used to feed the Joyce procedure. The output is an MM PES that is very close to the QM one and provides force field (FF) parameters to be used in MD simulations in which the molecule can be immersed in a complex environment. If the new FF is validated according to rigorously fixed criteria, the MD trajectory is used to extract snapshots to be used in a sequential QM/MM scheme.

The set-up of the Joyce procedure requires advanced modeling skills since it strongly depends on a deep knowledge of the target molecule and on how different modes can be described with model potential functions (for the intramolecular part) and on a realistic charge distribution (for the intermolecular Coulombic term). Regarding the intramolecular van den Waals term, the parameters are taken from the parametrization provided by the force field in the literature (typically OPLS[17]). The validation step of the intramolecular part in order to be able to faithfully reproduce the QM potential energy

surface is particularly crucial. We shall give, in the following sections, some computational details about each step of the integrated approach. We recall the illustration of this approach provided in Chapter 1, Figure 1.1.

4.2.1 QM calculations

All the QM calculations were performed using the Gaussian16[18] suite of code and visualized using GaussView6.[19] The structures of BPOH2 were first optimized at the CAM-B3LYP/6-311G(2d,p)[20, 21] level. The absorption spectra were computed using TD-DFT, an extension of DFT used to treat vertical excitation energies.[22] The same calculations were performed in the presence of a conductor-like polarizable continuum model (CPCM)[23] to take into account the solvation effects in water. Considering the importance of the torsional degrees of freedom associated with the N5-C6-C17-N16 = δ_{Ph} and C6-C1-O7-H8 = δ_{OH} dihedrals (see Figure 4.1 for atom numbering), we performed a relaxed scan of torsional angles between 0°and 180°in the gas phase. The point charges to be used in the Coulombic intermolecular potential were computed using localization schemes (Merz-Singh-Kollman scheme[24], CHelpG scheme[25], and CM5[26]) on the isolated molecules in their optimized geometry, accounting for solvent polarization through the CPCM scheme.

4.2.2 Intramolecular FF parametrization with Joyce

For each considered molecule, the data contained in the QM calculations were used to run Joyce. The code reads the optimized geometries and their energies, gradients, and

Hessian matrices. The process of parametrization started with an approximated topology file of the BPOH2 tautomers in a format that is compatible with the code that is used for MD simulations, GROMACS.[27] The process was divided into several sequential steps in which it is possible to define the internal coordinates manually to be optimized and fix some other coordinates until the criteria to be fulfilled to assure convergence was met. A final topology file is thus obtained, which includes fitted parameters. The theoretical aspects of FF parametrization using Joyce are available in Section 1.5. In our case, further validation of the fitted parameters was carried out by performing an MD simulation of the isolated molecule as well as an MM scan of δ_{Ph} and δ_{OH}. The simulation in vacuum was performed for 500 ps at room temperature with Berendsen thermostat.[28]

4.2.3 MD simulations

The BPOH2 molecule was included in a box containing 1198 water molecules for diol and 1199 water molecules for diketo, modeled using the flexible variation of the TIP3P water model.[29] The RESP fitting[30] was performed in antechamber[31] on the charges calculated using QM calculations to finally get the charges from the Merz-Singh-Kollman scheme on BPOH2. A cubic box (33.2396 Å $\times$ 33.2396 Å $\times$ 33.2396 Å) of TIP3P-flex[29] water molecules was created, and the center of mass of the solute BPOH2 was put at the center of the box.

The equilibration of the system started with the relaxation of solvent molecules around the frozen BPOH2 molecule. An NPT calculation was performed for 250 ps with the velocity rescaling (v-rescaling) technique[32] and the Parrinello-Rahman baro-

stat.[33] We then performed an NVT simulation for 1 ns with the same thermostat. MD production was performed along a 10 ns trajectory in the NPT ensemble.

4.2.4 Snapshot extraction and calculation of the UV-VIS spectra

100 independent configurations of the solute were extracted at regular intervals along the simulated trajectories using the built-in GROMACS tools. For each frame, TD-DFT calculations were carried out at the CAM-B3LYP/6-311G(2d,p) level in CPCM water to compute the absorption spectrum. Such a scheme is usually referred to as mechanical embedding. A convolution using Gaussian functions centered on each predicted transition was used to reconstruct the spectrum resulting from the different snapshots. The convolution $G(E)$ was calculated using the expression:[34]

$$G(E) = \sum_i g_i(E) = \sum_i g_i^{max} exp\Big[- \Big(\frac{E - E_i}{\sigma} \Big)^2 \Big], \tag{4.1}$$

where $g_i(E)$ is a Gaussian function centered on one value of the computed transition energies E_i. σ (= 0.1 eV) is the half-width of the Gaussian band at the maxima of $g_i(E)$. In the next steps (currently underway), we shall include water molecules from the first solvation shell in the extracted snapshots and apply the same procedure to this larger system to compute the spectrum, thus explicitly taking into account surrounding molecules.

4.3 Results and discussion

The focus of the study is to be able to simulate and provide a molecular description of the experimental absorption spectra in water (Figure 4.2).[12] Two peaks around 330 nm and 430 nm are observed in water and the absorption maximum is related to a $\pi - \pi^*$ transition.[3]

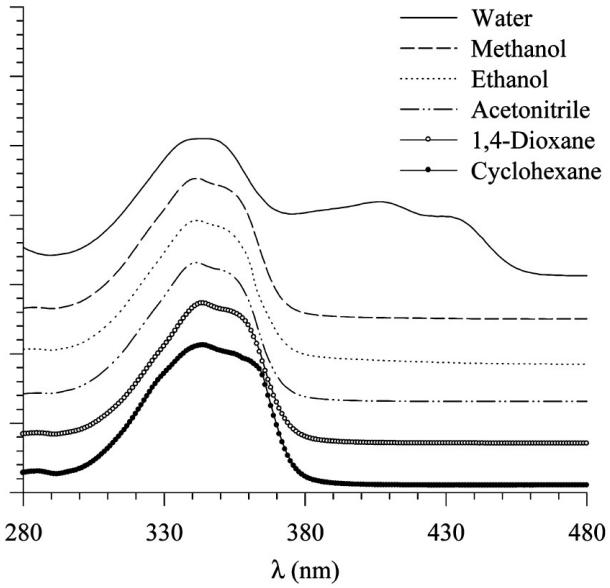


Figure 4.2: The experimental absorption spectra of BPOH2, Figure extracted from Ref.[12].

We shall start the discussion by describing the results obtained from electronic structure calculations.

4.3.1 QM calculations

We started with performing QM calculations to explore the potential energy surface of the diol tautomer. We ran a relaxed scan of two relevant torsional angles (N5-C6-C17-N16 = δ_{Ph} and C6-C1-O7-H8 = δ_{OH}), which affect the aromaticity of the molecule and

might have an important effect on its absorption spectrum. Results are shown in Figure 4.3.

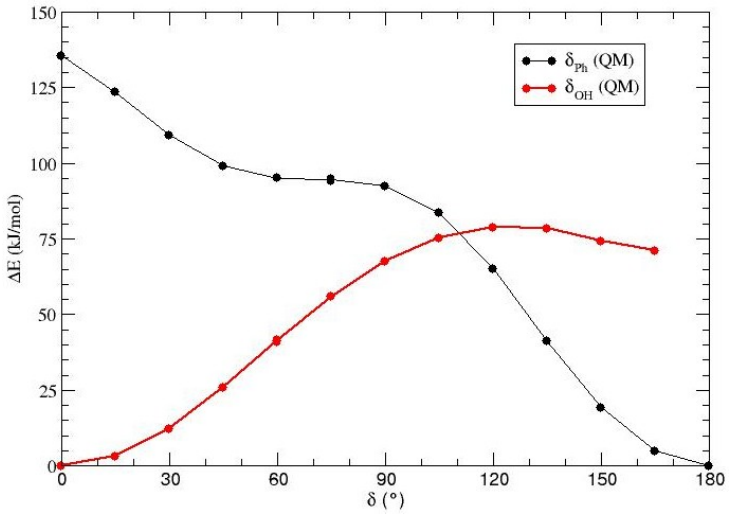


Figure 4.3: Relaxed scan of δ_{Ph} and δ_{OH} at a QM level for diol.

From the relaxed scan, we identified three conformers of BPOH2 based on the orientation of the phenolic -OH groups with respect to the aromatic ring. The global minimum for the diol was observed when the aromatic rings were co-planar, *i.e.* $\delta_{Ph} = 180°$, and the two H atoms of -OH groups were involved in hydrogen bonds with the N atoms (corresponding to $\delta_{OH} = 0°$). In addition, we obtained two local minima, both having co-planar rings but differing in the orientation of the hydroxyl groups. In one case, one -OH group forms a hydrogen bond with N ($\delta_{OH} = 0°$), while the other is pointing away from the (intermolecular) hydrogen bond direction ($\delta_{OH} = 180°$). We would refer to this as the 'one-open' conformer ($\delta_{OH} = 0°$). In the other case, both -OH groups point away, and we would refer to this conformer as the 'both-open' one. The structure of these conformers are reported in the Appendix, Figures C.1b and C.1a. The relative energies of the 'both-open' and of the 'one-open' conformers with respect to the global minimum were +136.0

kJ/mol and +70.0 kJ/mol, respectively.

In the case of the diketo tautomer, we only needed to perform the torsional scan of the dihedral N5-C6-C17-N16 = δ_{Ph} (see Figure 4.4). The global minimum of diketo was observed when δ_{Ph} was equal to 180°, which implies that the two aromatic rings are planar. The relative energy of the diketo tautomer with respect to the diol form was +36.2 kJ/mol in water (CPCM). In addition to the diketo tautomer, we also found a stable keto-enol tautomer (Figure 4.1b). In the gas phase, the keto-enol structure was obtained by performing a relaxed scan along the N..H..O proton transfer coordinate. The relative energy of the keto-enol tautomer with respect to the diol form was +21.3 kJ/mol in water (CPCM). While performing the relaxed scan of δ_{Ph} and δ_{OH}, the keto-enol was not stable and converted into the diketo or the diol forms, an issue that prevents finalizing the construction of the force field for this tautomer and on which we are presently working.

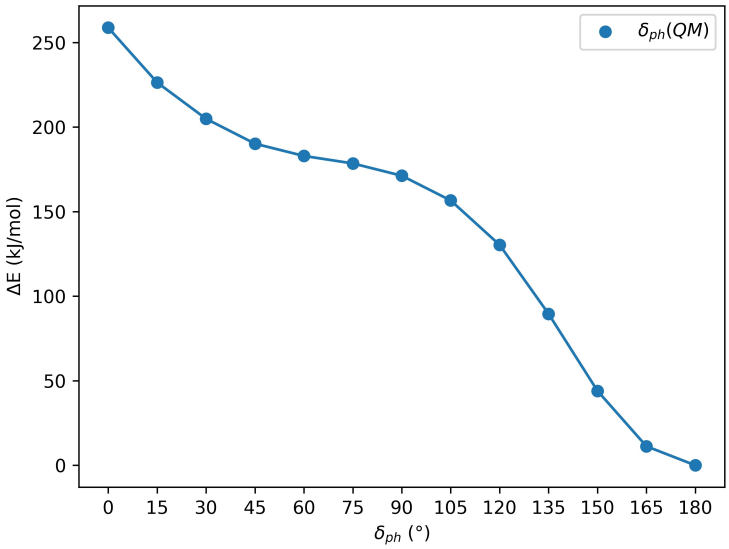


Figure 4.4: Relaxed scan of δ_{Ph} in the diketo tautomer at the QM level.

After obtaining the optimized geometries, we performed the TD-DFT calculations to simulate the electronic spectra in CPCM water. Figure 4.5 shows the absorption

spectra thus obtained. We observed λ_{max} at 308 nm and at 405 nm for the diol and the diketo forms, respectively. The absorption spectrum of keto-enol peaked at 366 nm. The position of the band for the keto-enol and diketo forms, appearing in a lower energy region compared with the diol, would be consistent with these species being present in an aqueous environment and giving rise to the smaller bands in the experimental spectrum.

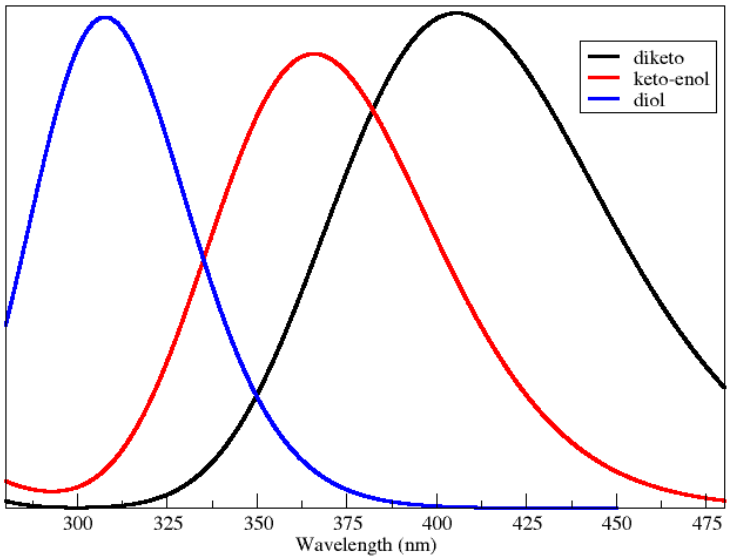


Figure 4.5: The absorption spectra of diol, keto-enol, and diketo were obtained through QM calculations.

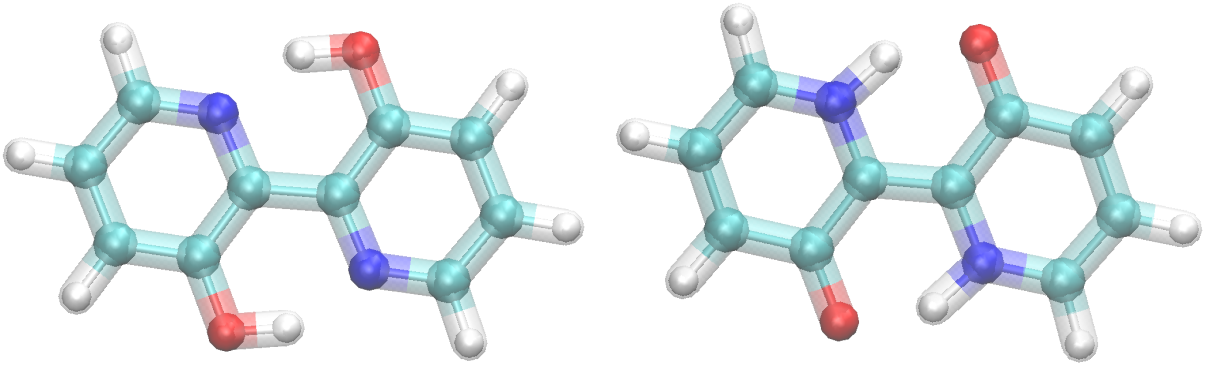
4.3.2 FF parametrization

The Joyce protocol applied to the diol and the diketo forms provided parameters that were considered satisfying with respect to the criteria to be met for the procedure to be successful. Table 4.1 collects the RMSD on different internal coordinates obtained using the parameterized FF with respect to QM geometries, showing a very good agreement, as it can also be visualized through Figure 4.6.

Some further tests were carried out by performing an MM scan of δ_{Ph} and δ_{OH} for the

Table 4.1: RMSD to compare the MM structures obtained based on the FF development, with the QM ones.

Internal coordinates	**diol**	**diketo**
Bond length (Å)	1×10^{-3}	1×10^{-3}
Angle (°)	0.1	0.1
Dihedral (°)	0.1	0.1



(a) diol (b) diketo

Figure 4.6: Comparison of the optimized structures obtained from QM (in the ball and stick representation) and from MM (in licorice representation) of the diol and diketo tautomers.

diol tautomer, the behavior of which showed that some further readjusting was needed and some additional LJ parameters between non-bonded pairs (N16...H8, N5...H19, H19...H8, and N16...N5 in diol) had to be introduced. In Figure 4.7, we show the final comparison between the relaxed scan of δ_{OH} and δ_{Ph} calculated from QM and from the parametrized FF for diol after readjustment. The two curves are in very good agreement, except for the region in which δ_{Ph} is between 0° and 30°, in which a strong non-bonding interaction, mainly electrostatic repulsion, occurs between H8 and H19. This interaction is only possible at higher temperatures, and hence the current form of the FF is acceptable for MD simulations at room temperature.

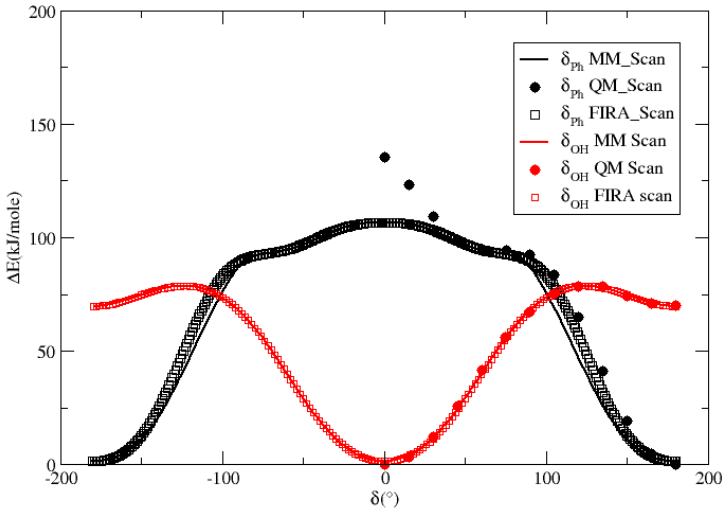


Figure 4.7: diol: relaxed scans of δ_{OH} and δ_{Ph} from QM and from the parametrized FF. Results are also shown for the Frozen Internal Rotation Approximation described in Chapter 1.

In the case of the diketo tautomer, we also performed an MM scan for δ_{Ph}, which was in good agreement with the QM scan (Figure 4.8) after adjusting the non-bonding LJ parameters for O7...H8, H18...O19, H8...H19, and O7...O18. The final optimized FF

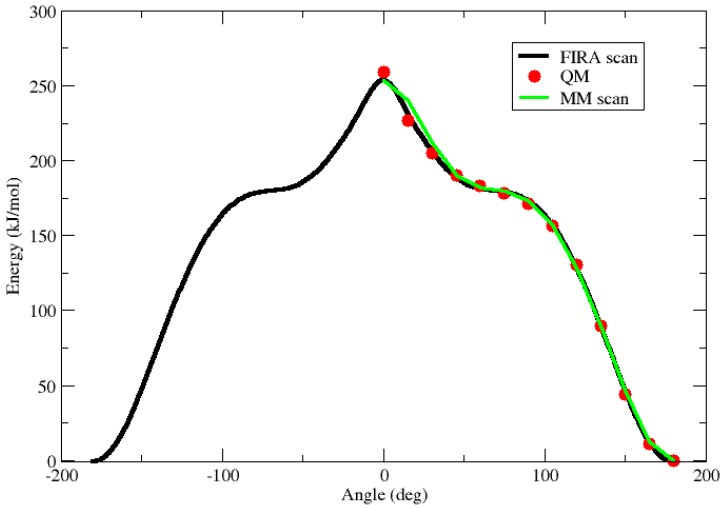
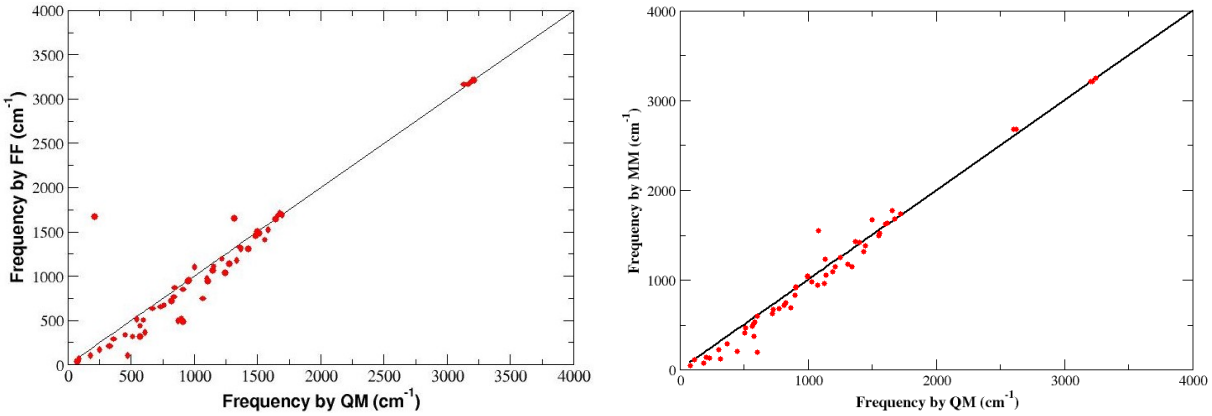


Figure 4.8: diketo: relaxed scans of δ_{OH} and δ_{Ph} from QM and from the parametrized FF. Results are also shown for the Frozen Internal Rotation Approximation described in Chapter 1.

parameters for the diol and the diketo tautomers are reported in the Appendix (Tables

C.1-C.8).

A good agreement is observed when comparing the frequencies of normal modes calculated from QM calculations and the developed force field (Figure 4.9). The small deviation in the case of diol (Figure 4.9a) is due to erroneous ordering of normal modes while comparing them and does not affect the physical behavior of the molecular vibrations.



(a) diol (b) diketo

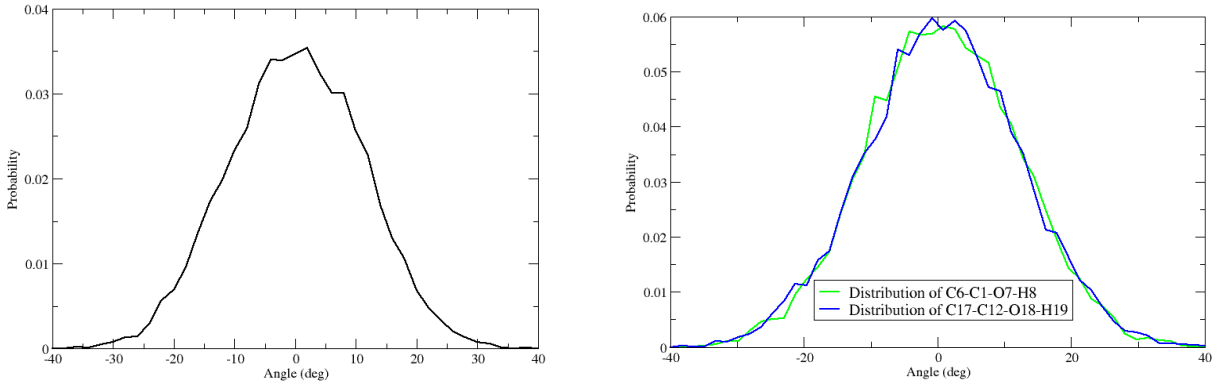
Figure 4.9: Comparison of the frequencies of normal modes of two tautomers obtained from QM and MM.

For the keto-enol form, the FF parameterization will require some adapted solution, since, as we already mentioned, the QM relaxed scan could not provide the necessary information due to the conversion to the other tautomers. From now on, we will restrict our presentation to the diol and diketo forms.

4.3.3 MD simulations

After obtaining the FF, we carried out MD simulations for the diol and the diketo conformers in water. We also simulated the 'both-open' and 'one-open' conformers of the

diol, but they switched back to the structure of the closed conformer (global minimum) during the course of the simulation. Hence, we can safely assume that they would not contribute to the final absorption spectrum of the diol. The distributions of the dihedrals measured along the simulation are reported in Figure 4.10 for the diol and in Figure 4.11 for the diketo forms. An analog of δ_{Ph}, the N5-C6-C17-C12 dihedral, is reported. In the optimized structure, when N5-C6-C17-N16 (δ_{Ph}) is 180°, N5-C6-C17-C12 is 0°. The advantage of showing the N5-C6-C17-C12 distribution is that it tells us about the rings' co-planarity during the simulation. For the diketo tautomer, the distribution of such dihedral was broad compared to the diol, which alerted us of the presence of a distorted structure and an unwanted non-planarity of the aromatic rings. For the diol, the results are consistent with an average situation in which the two rings are co-planar, and the -OH groups belong to the same plane (thus preserving the intramolecular hydrogen bond).



(a) Distribution of the N5-C6-C17-C12 dihedral

(b) Distribution of δ_{OH}

Figure 4.10: Distribution of flexible dihedrals in diol.

For the diol tautomer, the distributions of N...H and of O-H distances are shown in Figure 4.12. The phenolic -OH bond lengths do not vary much (around 1.0 Å), whereas the N...H distance displays a broad distribution, as expected. The O atoms of the phenolic

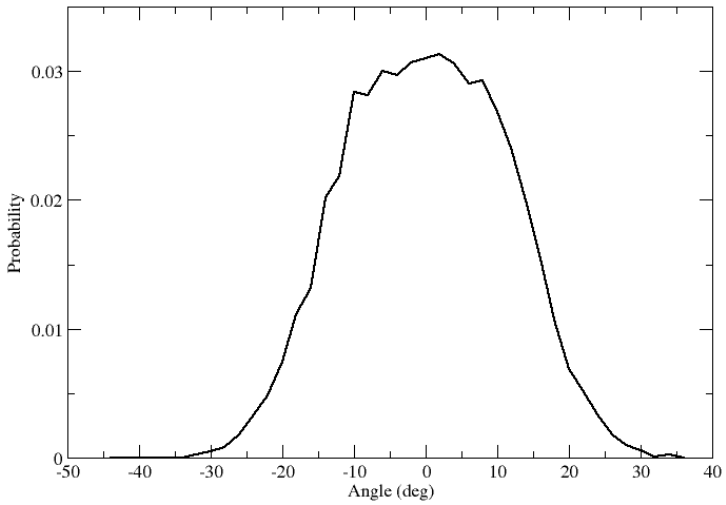


Figure 4.11: Diketo tautomer: the distribution of N5-C6-C17-C12 dihedral.

-OH group form weak hydrogen bonds with the H atoms of water molecules, as can be observed in the pair distribution function for the HO...HOH interaction, reported in Figure 4.13.

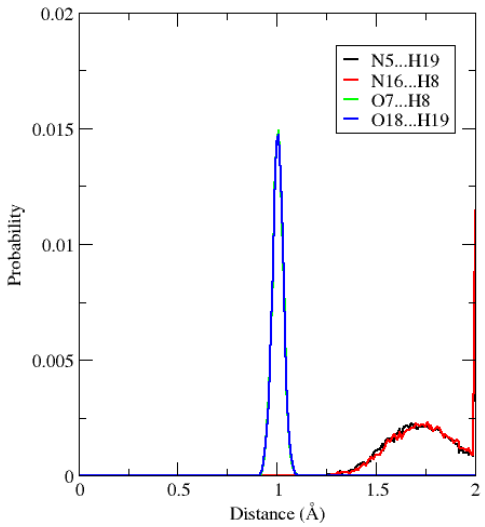


Figure 4.12: Diol tautomer: distribution of N...H and O-H distances.

In the MD simulation of the diketo tautomer in water, the N-H distance does not vary much, which is again consistent with what is expected for a covalent bond (see Figure 4.14). The O atoms of the keto group form strong hydrogen bonds with the neighboring

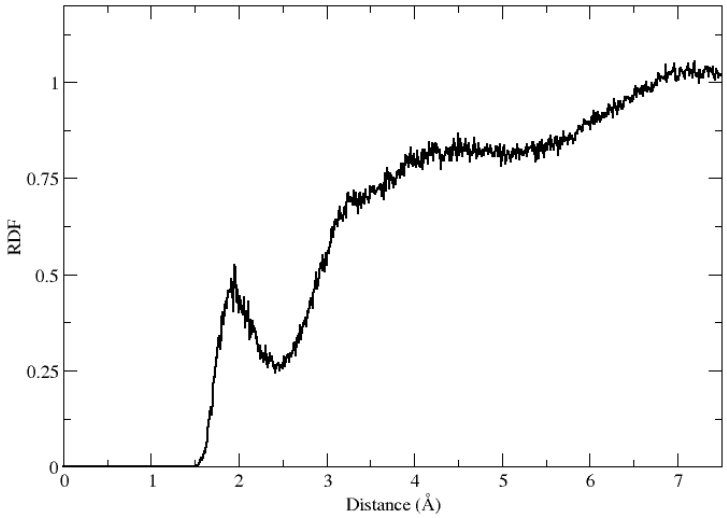


Figure 4.13: Diol tautomer: radial distribution function describing the HO...HOH interactions.

water molecules (see pair correlations in Figure 4.15).

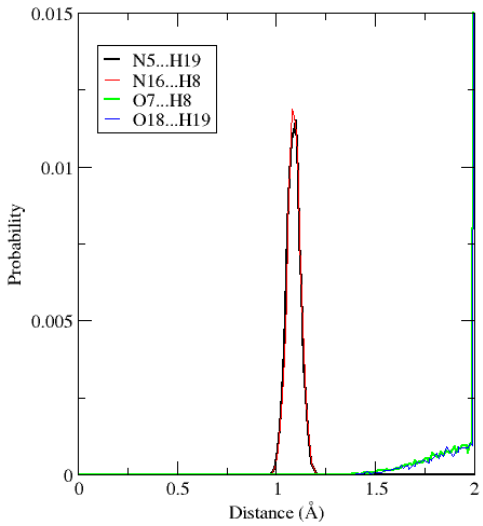


Figure 4.14: Diketo tautomer: distribution of N-H and O...H distances.

4.3.4 Absorption spectra

To incorporate the polarization effects of the environment on the structure of the solute in the absorption spectra, we used the mechanical embedding method. Some tests

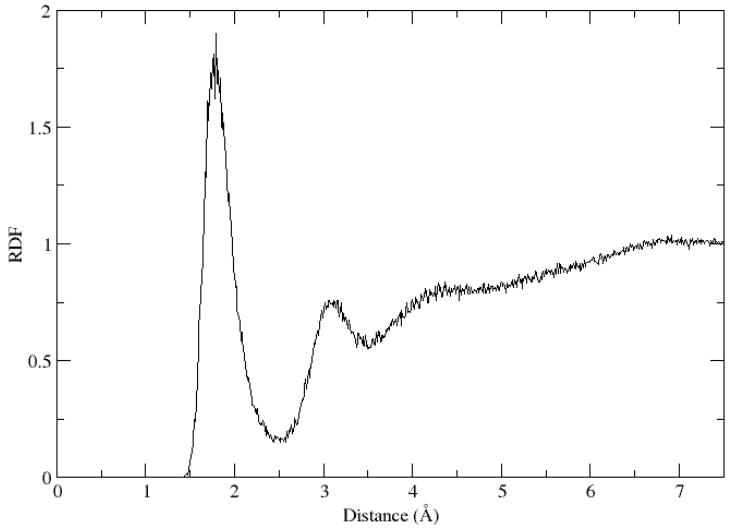
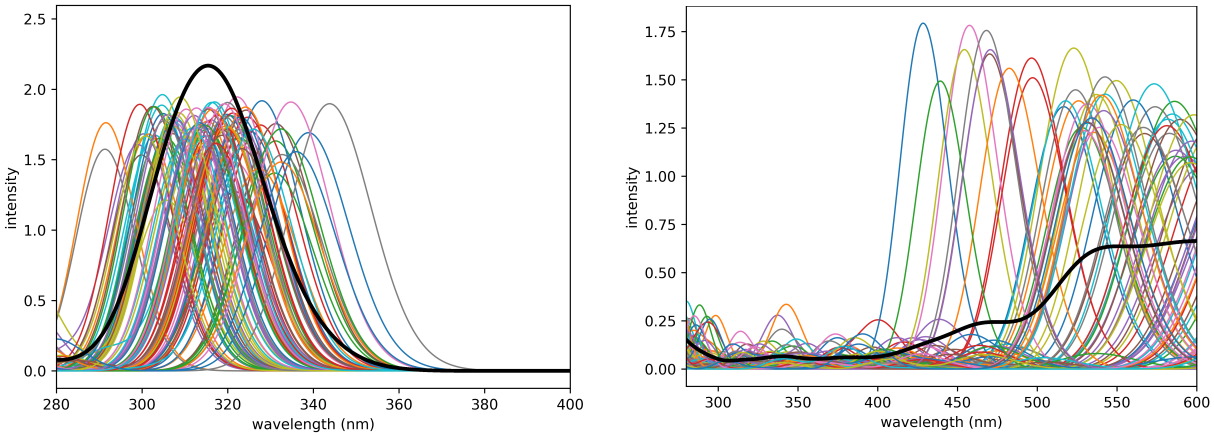


Figure 4.15: Diketo tautomer: radial distribution function describing the O...HOH interactions.

performed by taking into account different polarization schemes and different ways to include water molecules in the snapshots were carried out by Dr. Prampolini (see Appendix Figure C.2). His results show that the inclusion of explicit water molecules in the computation of the absorption spectrum of the diol form does not allow us to describe the low energy bands of the experimental spectrum, further support to our focus on the existence of the other two tautomers.

In Figure 4.16, we display the results thus obtained for the absorption spectra of the diol and diketo tautomers. The black line represents the convolution of the spectra computed for the snapshots (colored lines). The line shape and the position of the maximum (λ_{max} about 320 nm) are in good agreement with experiments and the peaks at higher wavelengths are still not observed (thus supporting our hypothesis on the origin stemming from other tautomers). However, our results were inconsistent for the diketo tautomers and pointed to an issue in the parametrization of the force field. An analysis of the results

based on single snapshots showed that most of the results were red-shifted with respect to the vertical excitation energy (≈ 3 eV) of the optimized diketo (see Appendix, Figure C.3). The structures of those snapshots revealed that the planarity of the rings was not correctly accounted for, as we already suspected from the distribution of the dihedral describing the co-planarity. As a matter of fact, the parametrized force constants ruling the planarity of the aromatic rings were way too low (≈ 4 kJ/mol/rad^2) as compared to those obtained for the diol (≈ 50 kJ/mol/rad^2).



(a) diol (b) diketo

Figure 4.16: Absorption spectra of two tautomers obtained at the QM level by using mechanical embedding (the colored lines are the spectra computed for the snapshots and their convolution is shown as the black line).

An improvement of the diketo FF is presently underway, and the preliminary results obtained by adjusting the force constant related to the rings' planarity are encouraging.

4.4 Preliminary conclusions

Although our initial goal was to develop a force field for BPOH2 in order to measure absorption and fluorescence in different macrocavities immersed in water, the parametrization steps and the MD simulation in the water bulk revealed a very interesting potential energy surface of this molecule in its ground state. Two more tautomers were shown to be stable in water, whose spectra peak at higher wavelengths and are very likely to be responsible for the lower intensity bands in this region.

As a matter of fact, our results for the ground state absorption spectrum of BPOH2 in water (mechanical embedding) did not show such bands. Some preliminary tests run by including surrounding water molecules in the snapshots on which QM calculations are run did not improve this feature, further supporting our hypothesis on the presence of one/two more tautomers. The next steps of our work, still underway, will first address the finalization of the FF development for the diketo and the keto-enol tautomers and the MD simulations in water, leading to the snapshot extraction. Additional refinements of the absorption spectrum calculation, including water molecules and point charges, will be also carried out.[13]

References

(1) J. Lipkowski, A. Grabowska, J. Waluk, G. Calestani and B. Hess, *J. crystallogr. spectrosc. res.*, 1992, **22**, 563–572.

(2) H. Bulska, A. Grabowska and Z. R. Grabowski, *J. Lumin.*, 1986, **35**, 189–197.

(3) A. L. Sobolewski and L. Adamowicz, *Chem. Phys. Lett.*, 1996, **252**, 33–41.

(4) P. Prosposito, D. Marks, H. Zhang and M. Glasbeek, *J. Phys. Chem. A*, 1998, **102**, 8894–8902.

(5) F. V. Neuwahl, P. Foggi and R. G. Brown, *Chem. Phys. Lett.*, 2000, **319**, 157–163.

(6) H. Bulska, *Chem. Phys. Lett.*, 1983, **98**, 398–402.

(7) Y. Su and S. Chai, *J. Photochem. Photobiol. A: Chem.*, 2014, **290**, 109–115.

(8) J. Sepioł, H. Bulska and A. Grabowska, *Chem. Phys. Lett.*, 1987, **140**, 607–610.

(9) P. Borowicz, A. Grabowska, R. Wortmann and W. Liptay, *J. Lumin.*, 1992, **52**, 265–273.

(10) R. Gelabert, M. Moreno and J. M. Lluch, *ChemPhysChem*, 2004, **5**, 1372–1378.

(11) V. Barone and C. Adamo, *Chem. Phys. Lett.*, 1995, **241**, 1–6.

(12) O. K. Abou-Zied, *J. Phys. Chem. B*, 2010, **114**, 1069–1076.

(13) G. Prampolini, F. Ingrosso, A. Segalina, S. Caramori, P. Foggi and M. Pastore, *J. Chem. Theory Comput.*, 2019, **15**, 529–545.

(14) G. Prampolini, F. Ingrosso, J. Cerezo, A. Iagatti, P. Foggi and M. Pastore, *J. Phys. Chem. Lett.*, 2019, **10**, 2885–2891.

(15) I. Cacelli and G. Prampolini, *J. Chem. Theory Comput.*, 2007, **3**, 1803–1817.

(16) V. Barone, I. Cacelli, N. D. Mitri, D. Licari, S. Monti and G. Prampolini, *Phys. Chem. Chem. Phys.*, 2013, **15**, 3736–3751.

(17) W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.

(18) 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*, Gaussian Inc. Wallingford CT, 2016.

(19) R. Dennington, T. A. Keith and J. M. Millam, *GaussView Version 6*, Semichem Inc. Shawnee Mission KS, 2019.

(20) T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51–57.

(21) R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.*, 1980, **72**, 650–654.

(22) M. E. Casida and M. Huix-Rotllant, *Annu. Rev. Phys. Chem.*, 2012, **63**, 287–323.

(23) S. Miertuš, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117–129.

(24) U. C. Singh and P. A. Kollman, *J. Comput. Chem.*, 1984, **5**, 129–145.

(25) C. M. Breneman and K. B. Wiberg, *J. Comput. Chem.*, 1990, **11**, 361–373.

(26) A. V. Marenich, S. V. Jerome, C. J. Cramer and D. G. Truhlar, *J. Chem. Theory Comput.*, 2012, **8**, 527–541.

(27) M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1**, 19–25.

(28) H. J. Berendsen, J. v. Postma, W. F. Van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

(29) W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

(30) R. Woods and R. Chappelle, *Theochem*, 2000, **527**, 149–156.

(31) J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graph.*, 2006, **25**, 247–260.

(32) L. V. Woodcock, *Chem. Phys. Lett.*, 1971, **10**, 257–261.

(33) M. Parrinello and A. Rahman, *Phys. Rev. Lett.*, 1980, **45**, 1196–1199.

(34) J. B. Foresman and A. Frisch, *Exploring chemistry with electronic structure methods*, Gaussian, Inc, Wallingford, CT USA, Third edition, 2015.

Conclusions and Future Perspectives

This Ph.D. project was devoted to developing theoretical and computational tools to describe solvation and confinement effects in complex systems. Particular attention was devoted to the vibrational properties of organic molecules intercalated in clay-like systems and to the optical properties of chromophores that are sensitive to confinement effects. These two parts of the project have benefited from collaborations with other theoretical and experimental groups established by the laboratory where the project has been run. Important changes to the original project, which was supposed to include important mobility to work in partner laboratories abroad, had to be made due to the strong limitations to international mobility dictated by the Coronavirus pandemic. However, we can now safely state that such modifications allowed us to successfully keep the original idea of the project, i.e., to develop and combine computational strategies to be used in a wide variety of applications when the molecular system of interest strongly interacts with its environment and is subject to confinement effects.

Our molecular dynamics study of Mg/Al 2:1 layered double hydroxides, intercalated with organic anions such as succinate, aspartate, and glutamate, was motivated by some questioning arising when observing the experimental water adsorption isotherms as well

as the evolution of the carboxylate stretch band of the infrared spectrum with increasing hydration of the material (see Appendix). Our detailed study of the modifications induced by an increase in the interlamellar space to the structure of the interlamellar region shed some light on the adsorption mechanism of these anions, which is mediated by anchoring of the carboxylate groups to the surface via the formation of hydrogen bonds with the hydroxyl groups, and further stabilized by water molecules, displaying a strong affinity with such groups.[1, 2]

The ability of carboxylate groups to accept hydrogen bonds strongly depends on the structural features of the molecule bearing such groups as well as its surroundings. We were interested in having a comparison between the very complex environment of the LDH and a much simpler situation in which we would progressively describe the role of surrounding molecules in a bulk solvent. Two previous studies in the literature inspired us to set up a bottom-up approach to analyze the behavior of two model carboxylates in water.[3, 4] We chose succinate and aspartate, the structure of which only differs because of the amino group in aspartate. For the latter, we considered both the zwitterionic protonated form, which is stable in neutral solutions and the deprotonated form, which is stable in basic environments (e.g., in LDH). The anions were modeled using quantum chemistry calculations in a continuum solvent as well as with a microsolvation approach. Results on the potential energy surface and on the carboxyl stretch band of the infrared spectrum showed that one should expect larger differences when using the same level of quantum chemistry but different solvation models than when the same solvation model is used with different methods (different DFT functionals and basis sets, semiempirical

methods). Since the PM6 level was in agreement with higher-level results, we went one step further in the description of solvation effects by using QM/MM molecular dynamics. This allowed us to provide a deep description of the solvation shell of the anions at finite temperature, of the differences observed close to the amino group, and, in particular, when different protonation states of the latter are taken into account.

In LDH, the picture is indeed more complex. As a matter of fact, the reorientation of the anion in the interlamellar region plays a very important role in the adsorption mechanism, which was already suggested by other studies in the literature[5] but, to the best of our knowledge, not systematically analyzed yet. Each anion reacts differently to the increased interlamellar space. Succinate reorients quite abruptly (which is consistent with the steep rise in the interlamellar space and in the trend of the adsorption isotherm) and moves from a situation in which it binds to the surface in a bidentate bridging-like arrangement to a situation in which the binding modes are mostly in a bidentate chelating-like arrangement. In higher hydration states, anions can be found either in a vertical orientation with respect to the surface or lying horizontally next to the surface. The evolution of different populations of anions, interacting differently with the surface, is consistent with the evolution of the carboxylate stretch band. Aspartate and glutamate undergo smoother transitions, and in the case of glutamate, a complete reorientation is not observed since the anion is bulkier and, at the same time, presents larger conformational mobility. In the amino acid systems, the amino group can participate in binding to the surface, though this happens mostly in the anhydrous systems or at low hydration. In those conditions, neighbor amino acids can also interact along a hydrogen bond path

(COO...HNH) that has been claimed to be propitious toward chemical reactivity leading to peptide bond formation, and thus polymerization, to form polypeptides in abiotic conditions.[5–7]

To move to even more challenging projects, some interesting extensions of the present work would be to treat the properties of confined chromophores in LDH and, eventually, chemical reactivity with possible applications in the domain of photocatalysis.[8, 9] Some effort is currently being made in the laboratory to parametrize through Joyce the ground and excited states of a ruthenium polypyridine complex intercalated into LDH. Some interesting experimental results hinted at a different mechanism for the chemiluminescence decay under confinement compared with the bulk.[10] Interestingly, the substituents at the pyridine rings are carboxylate groups, and our experience in analyzing their binding modes will be key in assessing the role of the surface in the reorganization induced by the electronic transition. As for systems undergoing chemical reactivity, we may be confronted with the necessity of implementing different multiscale schemes or adapting our parametrization procedure to develop reactive force fields.[11]

In the second part of the Ph.D. project, we worked to build a sophisticated force field of an organic chromophore used as a model of photoinduced intramolecular proton transfer, 2,2'-bipyridine-3,3'-diol. Our final goal is to give a molecular interpretation of the differences measured in the absorption and fluorescence spectra when the molecule is confined in a cyclodextrin microcavity of different sizes and natures.[12] The proposed methodology is based on a refined sequential classical-QM approach, in which an ad hoc parametrized force field for the chromophore allows to run MD simulations using a

potential energy surface the reproduces all the features of a QM-determined PES. MD is used to sample configurations of the molecule (and of increasingly larger portions of the surrounding medium) along the trajectory and to perform QM calculations of the desired properties (in our case, the UV/VIS and fluorescence spectra) on each snapshot. The convolution of the data thus obtained would provide more realistic spectra: absorption spectra when the sampling is done in a ground state trajectory and fluorescence spectra when such sampling is done in an excited state trajectory.[13, 14]

During the parametrization of the ground state molecule, we became aware of the importance of two more tautomers that are stabilized in water: an intermediate keto-enol form and a diketo form. In particular, they seem to be responsible for low energy bands in the absorption spectrum that are particularly sensitive to the environment and might provide information on how hydrated (and then hydrophilic) the confined space within the cyclodextrin cavities is.[12, 15] As a matter of fact, we showed using QM calculations that these two forms are stable in water and that their absorption spectra peak at wavelengths corresponding to those bands, the agreement being very satisfying. This discovery, though extremely promising, forced us to provide a different force field for each tautomer, thus making our efforts on the description of the ground state system stronger than expected. We are now in the process of finalizing this step. Some natural, further perspectives arising from this work are the obtention of the absorption spectrum for the three tautomers using the sequential approach in water and in other solvents. We shall then move to the complexation of the chromophore into the cyclodextrin cavity, a subject for which the laboratory has long-standing experience.[16–19] Finally, as in

the case of LDH, we shall move to model chemical reactivity, in particular to describing intramolecular proton transfer in the excited state, a particularly challenging question that attracts a great deal of attention in the scientific community.[20–25]

References

(1) S. P. Newman, T. Di Cristina, P. V. Coveney and W. Jones, *Langmuir*, 2002, **18**, 2933–2939.

(2) P. P. Kumar, A. G. Kalinichev and R. J. Kirkpatrick, *J. Phys. Chem. C*, 2007, **111**, 13517–13523.

(3) C. C. Sutton, G. V. Franks and G. da Silva, *Spectrochim Acta A Mol Biomol Spectrosc*, 2015, **134**, 535–542.

(4) B. Hernández, F. Pflüger and M. Ghomi, *J. Comp. Chem.*, 2020, **41**, 1402–1410.

(5) B. Grégoire, V. Erastova, D. L. Geatches, S. J. Clark, H. C. Greenwell and D. G. Fraser, *Geochim. Cosmochim. Acta*, 2016, **176**, 239–258.

(6) B. Grégoire, H. C. Greenwell and D. G. Fraser, *ACS Earth Space Chem.*, 2018, **2**, 852–862.

(7) V. Erastova, M. T. Degiacomi, D. G. Fraser and H. C. Greenwell, *Nat. Commun.*, 2017, **8**, 2033.

(8) K. Teramura, S. Iguchi, Y. Mizuno, T. Shishido and T. Tanaka, *Angew. Chem. Int. Ed.*, 2012, **51**, 8008–8011.

(9) L. Mohapatra and K. Parida, *J. Mater. Chem. A*, 2016, **4**, 10744–10766.

(10) F.-N. Xiao, K. Wang, F.-B. Wang and X.-H. Xia, *Anal. Chem.*, 2015, **87**, 4530–4537.

(11) T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga et al., *Npj Comput. Mater.*, 2016, **2**, 1–14.

(12) O. K. Abou-Zied, *J. Phys. Chem. B*, 2010, **114**, 1069–1076.

(13) G. Prampolini, F. Ingrosso, A. Segalina, S. Caramori, P. Foggi and M. Pastore, *J. Chem. Theory Comput.*, 2019, **15**, 529–545.

(14) G. Prampolini, F. Ingrosso, J. Cerezo, A. Iagatti, P. Foggi and M. Pastore, *J. Phys. Chem. Lett.*, 2019, **10**, 2885–2891.

(15) K. Gavvala, A. Sengupta, R. K. Koninti and P. Hazra, *Phys. Chem. Chem. Phys.*, 2014, **16**, 933–939.

(16) M. Altarsha, V. Yeguas, F. Ingrosso, R. López and M. F. Ruiz-López, *J. Phys. Chem. B*, 2013, **117**, 3091–3097.

(17) F. Ingrosso, M. Altarsha, F. Dumarçay, G. Kevern, D. Barth, A. Marsura and M. F. Ruiz-López, *Chem. Eur. J.*, 2016, **22**, 2972–2979.

(18) I. Yakavets, H.-P. Lassalle, I. Yankovsky, F. Ingrosso, A. Monari, L. Bezdetnaya and V. Zorin, *J. Photochem. Photobiol. A*, 2018, **367**, 13–21.

(19) B. Aslanoglu, I. Yakavets, V. Zorin, H.-P. Lassalle, F. Ingrosso, A. Monari and S. Catak, *Phys. Chem. Chem. Phys.*, 2020, **22**, 16956–16964.

(20) A. S. Klymchenko and A. P. Demchenko, *J. Am. Chem. Soc.*, 2002, **124**, 12372–12379.

(21) J. Seo, S. Kim and S. Y. Park, *J. Am. Chem. Soc.*, 2004, **126**, 11154–11155.

(22) J. Zhao, S. Ji, Y. Chen, H. Guo and P. Yang, *Phys. Chem. Chem. Phys.*, 2012, **14**, 8803–8817.

(23) A. C. Sedgwick, L. Wu, H.-H. Han, S. D. Bull, X.-P. He, T. D. James, J. L. Sessler, B. Z. Tang, H. Tian and J. Yoon, *Chem. Soc. Rev.*, 2018, **47**, 8842–8880.

(24) W. Wang, M. Marshall, E. Collins, S. Marquez, C. Mu, K. H. Bowen and X. Zhang, *Nat. Commun*, 2019, **10**, 1–7.

(25) Y. Li, D. Dahal, C. S. Abeywickrama and Y. Pang, *ACS Omega*, 2021, **6**, 6547–6553.

Appendix A

First appendix

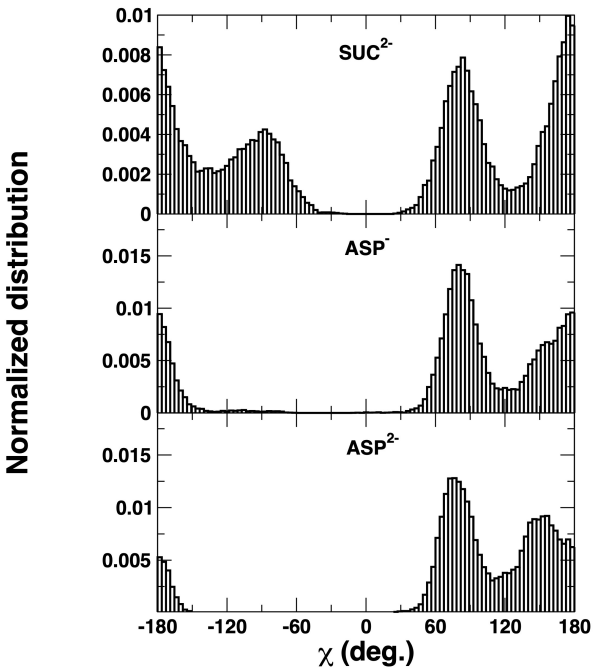


Figure A.1: Normalized distribution of two relevant dihedrals characterizing the conformations of the three anions considered in water.

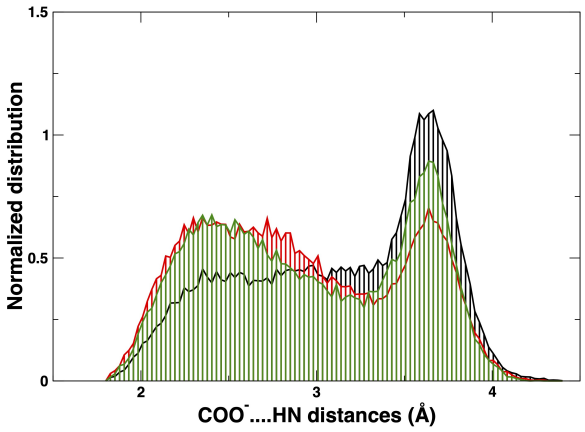


Figure A.2: Normalized distribution of the $C_\gamma OO^-...HNH_2^+$ distances correspond to the formation of an intramolecular hydrogen bond in aspartate (ASP^-).

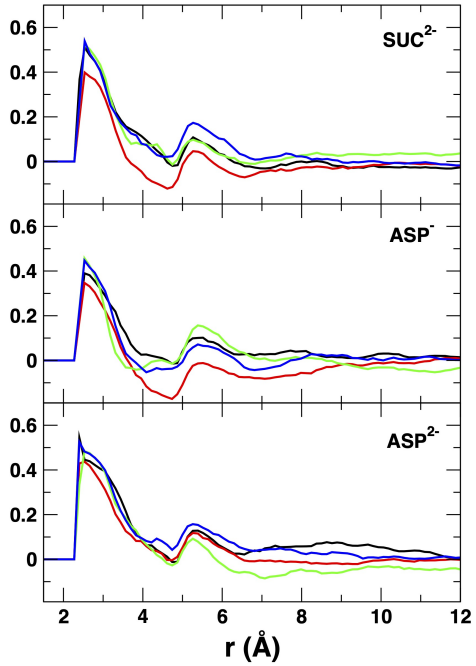


Figure A.3: Orientational correlations based on the calculation of the $G_1(r)$) normalized function. The different colors refer to the four different O atoms of the carboxylate groups.

Appendix B

Second appendix

Contents

B.1 Experimental measurements for LDH intercalating organic anions 172

B.2 Force fields used in MD simulations 173

B.2.1 Clay force field 173

B.2.2 Generalized amber force field 176

B.2.3 Force fields for water 177

B.3 Supporting information on the MD simulations of LDH intercalating organic anions 179

B.1 Experimental measurements for LDH intercalating organic anions

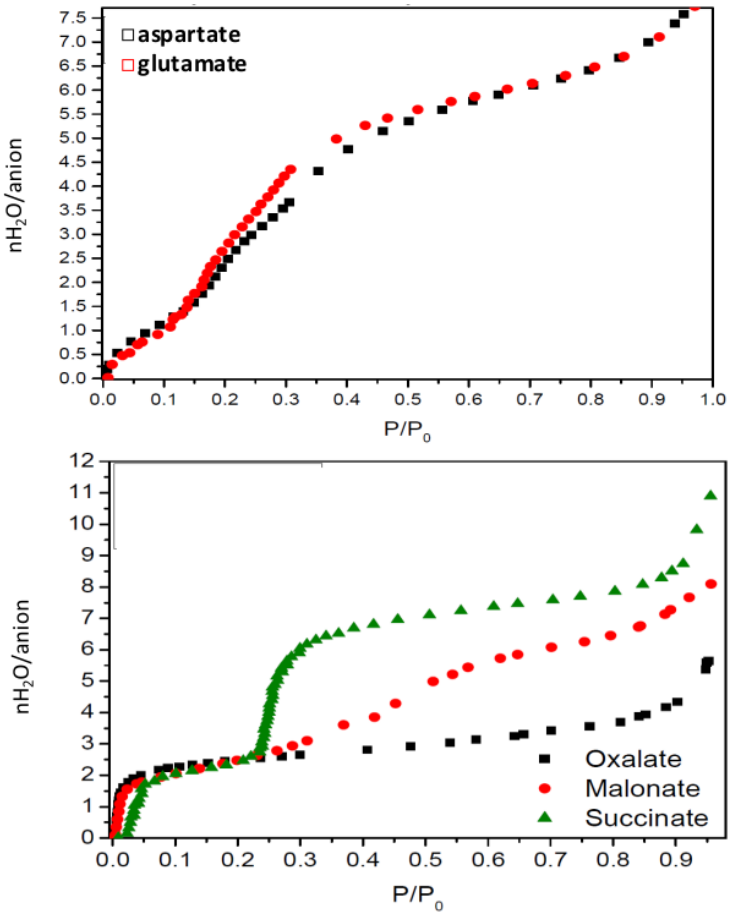
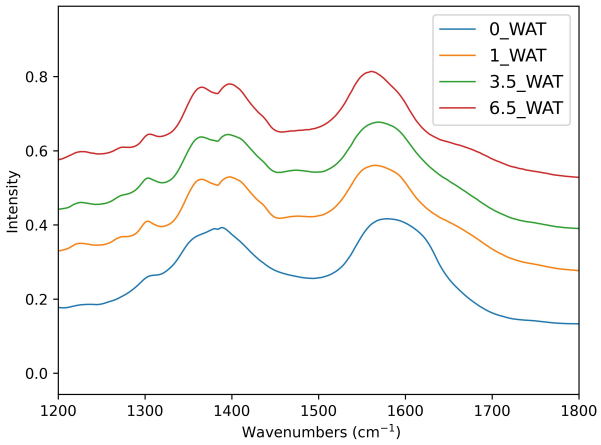
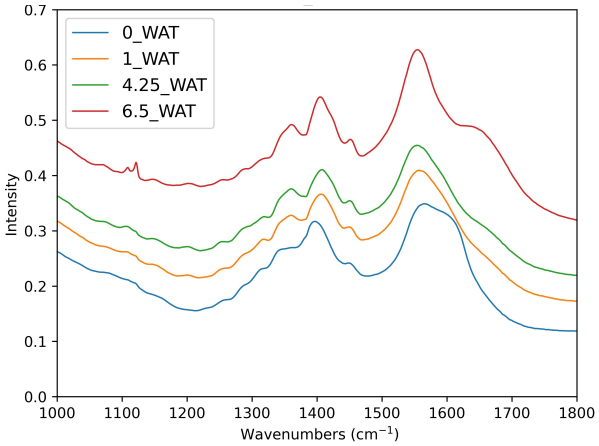


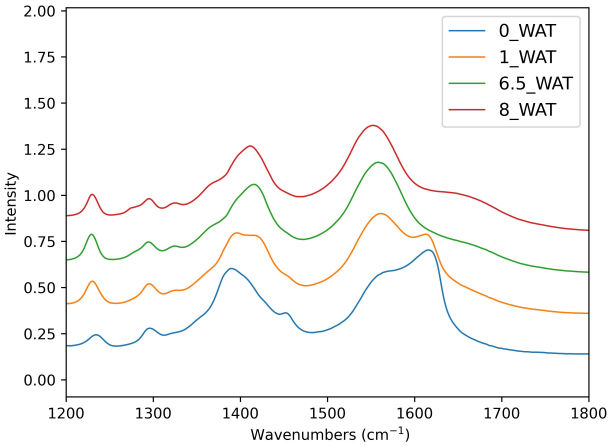
Figure B.1: Adsorption isotherms of LDH intercalated with organic intercalates.



(a) IR spectra of ASP



(b) IR spectra of GLU



(c) IR spectra of SUC

Figure B.2: Experimental infra-red of intercalates.

B.2 Force fields used in MD simulations

B.2.1 Clay force field

Clay force field (ClayFF)[1] relies on the single point charge (SPC) water model and was created by Berendsen et al.[2] to represent the water molecules, hydroxyl groups, and oxygen-oxygen interactions. In ClayFF, the hydroxyl group's non-bonding interactions

include the incorporation of partial charges for Coulombic terms and the short-range interactions represented by a simple Lennard-Jones term. O-H Bond stretch and M-O-H bond angle terms are introduced like the SPC water model using the expressions determined by Teleman et al.[3] to incorporate the flexibility for the hydroxides. The total energy terms are composed of four terms:

$$E_{total} = E_{Coul} + E_{VDW} + E_{bondstretch} + E_{anglebend} \tag{B.1}$$

In the Equation B.1, E_{Coul}, E_{VDW}, $E_{bondstretch}$, and $E_{anglebend}$ are the contribution of Coulombic (electrostatic) interactions, the short-range non-bonded interactions, bond stretching and bending, respectively. The Coulombic interaction energy is

$$E_{Coul} = \frac{e^2}{4\pi\epsilon_0} \sum_{i \neq j} \frac{q_i q_j}{r_{ij}} \tag{B.2}$$

q_i and q_j are the partial charges calculated from quantum mechanics calculations, e is the charge of the electron, and ϵ_0 is the dielectric permittivity of vacuum (8.85419×10^{-12} F/m). Lennard-Jones (12-6) function (LJ function) was used for short-range non-bonded interactions which are

$$E_{VDW} = \sum_{i \neq j} D_{0,ij} \left[\left(\frac{R_{0,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{0,ij}}{r_{ij}} \right)^{6} \right] \tag{B.3}$$

$D_{0,ij}$ and $R_{0,ij}$ are the empirical parameters of the pair of atoms i and j calculated from the fitting of a large number of experimental structures of simple oxides, hydroxides, and oxyhydroxides. We adopted an ionic description for metal-oxygen interactions associated

with hydrated phases and the O-H bond of the SPC water to model the O-H bond. The bond stretching term is

$$E_{bondstretch} = \sum_{ij} k_1(r_{ij} - r_0)^2 \tag{B.4}$$

where k_1 is the force constant, and r_0 represents the equilibrium bond length and both of these values were borrowed from the SPC water model.[3] The bending of the metal-hydroxyl (metal-oxygen-hydrogen) bond is

$$E_{anglebend} = \sum_{ij} k_2(\theta_{ijk} - \theta_0)^2 \tag{B.5}$$

where k_2 is the force constant, θ_{ijk} is the bond angle for the metal-oxygen-hydrogen, and θ_0 represents the equilibrium bond angle. The optimized parameters used in the context of this chapter for Al, Mg, and hydroxides are shown in Table B.1.[1] We did not use the metal-oxygen-hydrogen terms and restricted bending motion of metal-hydroxide bond angle to obtain a better equilibrated system during simulation.

Table B.1: The optimized parameters of ClayFF

Species	**Charge**	$D_{0,ij}$ **(kcal/mol)**	$R_{0,ij}$**(Å)**
octahedral Al	1.5750	1.3295×10^{-6}	4.7943
octahedral Mg	1.3600	9.0298×10^{-7}	5.9090
hydroxyl H	0.4250	–	–
hydroxyl O	-0.9742	–	–
Bonding Parameters for the ClayFF Force Field			
Bond stretch		k_1 (kcal/mol-Å^2)	r_0(Å)
hydroxyl O	hydroxyl H	554.1349	1.0000

B.2.2 Generalized amber force field

We employed the generalized amber force field (GAFF)[4] for modeling the organic intercalates. The functional form of the force field is similar to that of Equation 1.29. In GAFF, the van der Waals parameters are inherited from the traditional Amber force fields, the partial charges are assigned using a restrained electrostatic potential fit (RESP) model. For the implementation of bond stretching, bending, and dihedral terms, parameterizations are performed on bond lengths and bond angles that are weakly coupled to other parts in the energy function and influenced by the experimental observations or *ab initio* calculations. Force constants are estimated through an empirical approach to reproduce the *ab initio* calculated or experimentally observed vibrational frequencies.

Table B.2: The optimized parameters of GAFF

Bonding Parameters for the GAFF Force Field			
Bond stretch		r_0 (Å)	ln K_{ij}
C	C	1.526	7.643
C	H	1.090	6.217
C	O	1.440	7.347
C	N	1.470	7.504
C	N	1.010	6.057

B.2.3 Force fields for water

The simple point charge (SPC)[2] and extended simple point charge (SPC/E)[5] are the two models for water molecules that are consistent with the system of water molecules intercalated in LDH.[1] These models are three-site point models where atoms are replaced by point masses with partial charges. The 12-6 LJ function was used for non-bonded interactions between water molecules, which is written as:

$$E_{LJ} = \sum_{ij} \left(\frac{A}{r_{ij}}\right)^{12} - \left(\frac{B}{r_{ij}}\right)^{6} \tag{B.6}$$

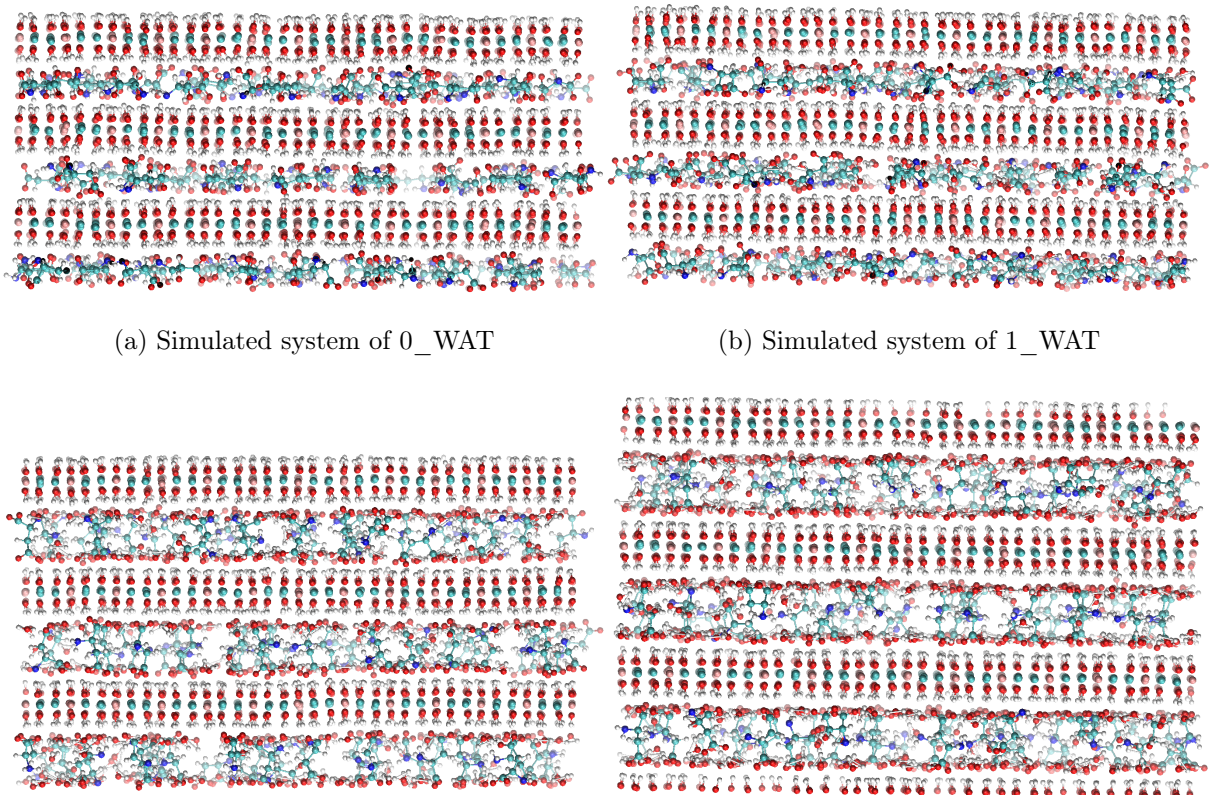
where the r_{OO} is the distance between the oxygen atoms of two water molecules, and A and B are the system dependent constants that are proportional to bond dissociation energy and equilibrium bond length. The electrostatic interactions are represented by Equation 1.36. The bond stretching and bending are represented by Equations 1.32 and

1.33, respectively. The optimized parameters are shown in Table B.3.

Table B.3: The optimized parameters of SPC and SPC/E

Parameters	**SPC**	**SPC/E**
Bonding Parameters		
OH bond length (Å)	1.0	1.0
HOH bond angle (deg)	109.47	109.47
Non-Bonding Parameters		
A (10^3 kcal Å^{12}/mol)	629.4	629.4
B (kcal Å^6/mol)	625.5	625.5
Partial charge O	-0.82	-0.8476
Partial charge H	+0.41	+0.4238

B.3 Supporting information on the MD simulations of LDH intercalating organic anions



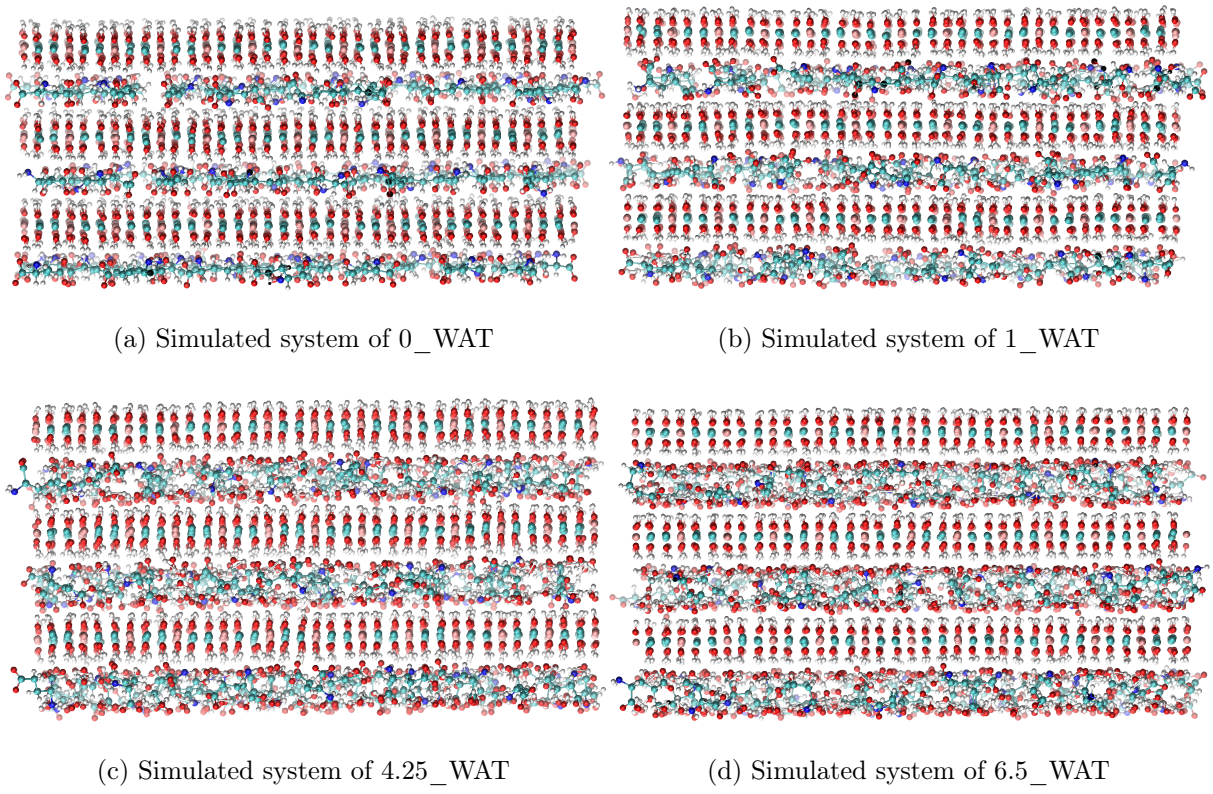
(a) Simulated system of 0_WAT

(b) Simulated system of 1_WAT

(c) Simulated system of 3.5_WAT

(d) Simulated system of 6.5_WAT

Figure B.3: Simulated systems of ASP (Blue: N, White: H, Red: O, Cyan: C, Pink: Al, Turquoise: Mg).



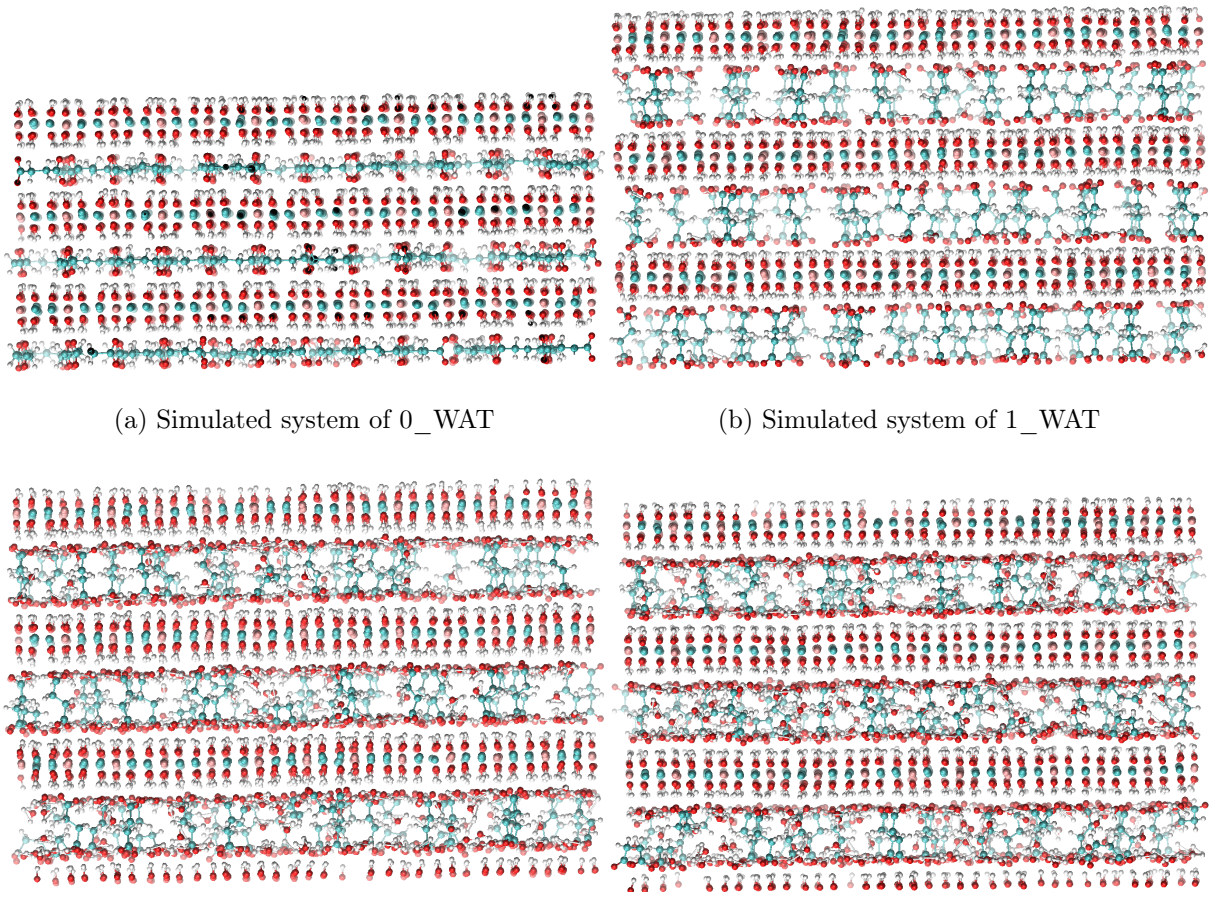
(a) Simulated system of 0_WAT

(b) Simulated system of 1_WAT

(c) Simulated system of 4.25_WAT

(d) Simulated system of 6.5_WAT

Figure B.4: Simulated systems of GLU (Blue: N, White: H, Red: O, Cyan: C, Pink: Al, Turquoise: Mg).



(a) Simulated system of 0_WAT

(b) Simulated system of 1_WAT

(c) Simulated system of 6.5_WAT

(d) Simulated system of 8_WAT

Figure B.5: Simulated systems of SUC (Blue: N, White: H, Red: O, Cyan: C, Pink: Al, Turquoise: Mg).

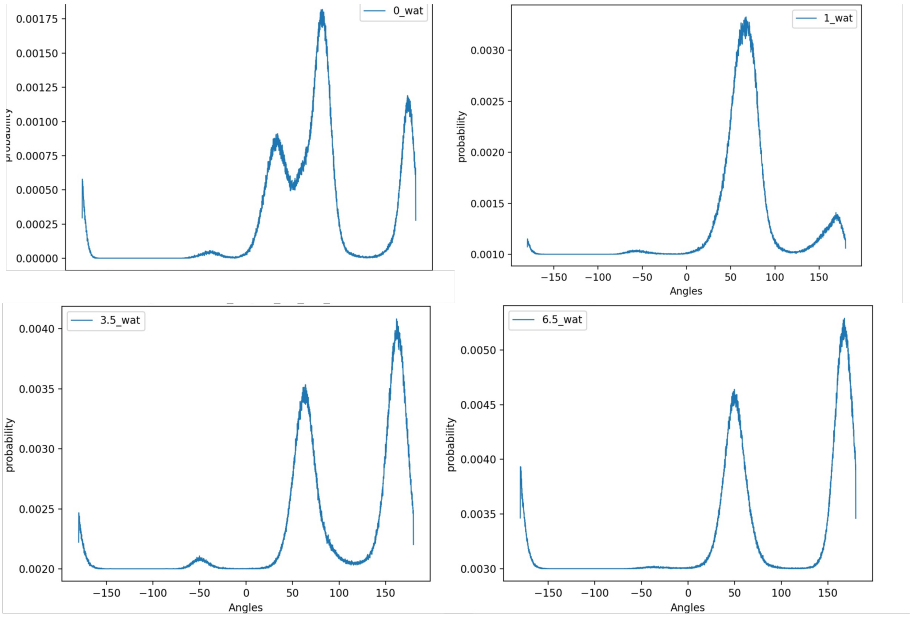


Figure B.6: Distribution of the C_α-C_β-C_γ dihedral in the system intercalating aspartate with increasing water content.

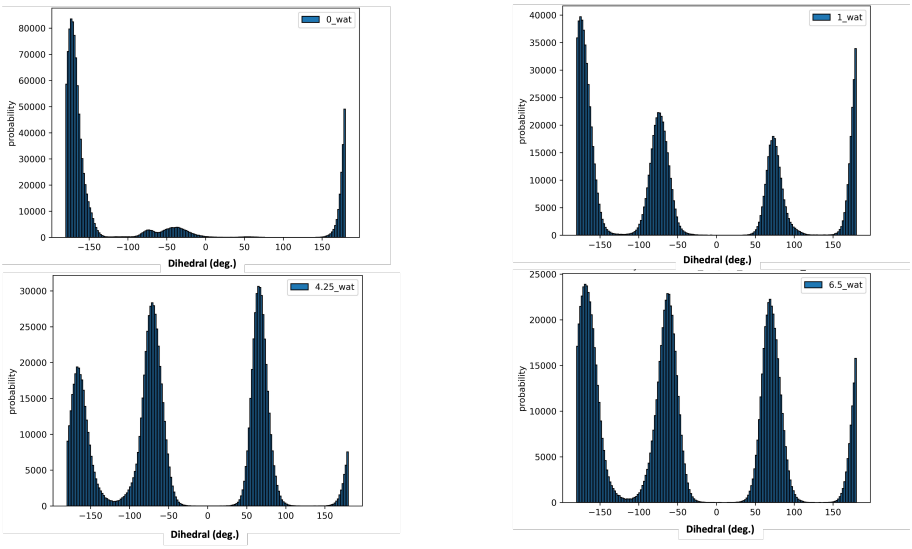


Figure B.7: Distribution of the C_α-C_β-C_γ dihedral in the system intercalating glutamate with increasing water content.

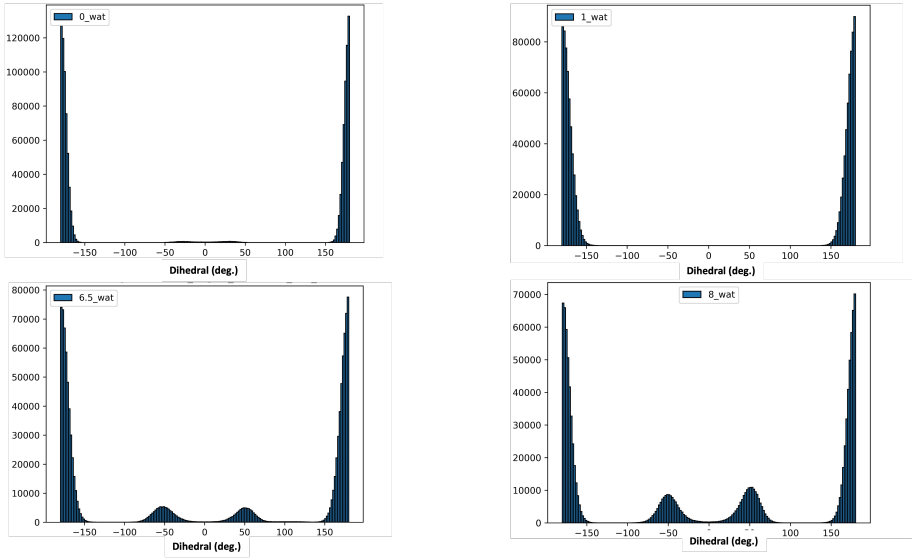


Figure B.8: Distribution of the C-C-C-C dihedral in the system intercalating succinate with increasing water content.

References

(1) R. T. Cygan, J.-J. Liang and A. G. Kalinichev, *J. Phys. Chem. B*, 2004, **108**, 1255–1266.

(2) H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren and J. Hermans, in *Intermolecular Forces*, ed. B. Pullman, Springer Netherlands, Dordrecht, 1981, vol. 14, pp. 331–342.

(3) O. Teleman, B. Jönsson and S. Engström, *Molecular Physics*, 1987, **60**, 193–203.

(4) J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem*, 2004, **25**, 1157–1174.

(5) H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269–6271.

Appendix C

Third appendix

Contents

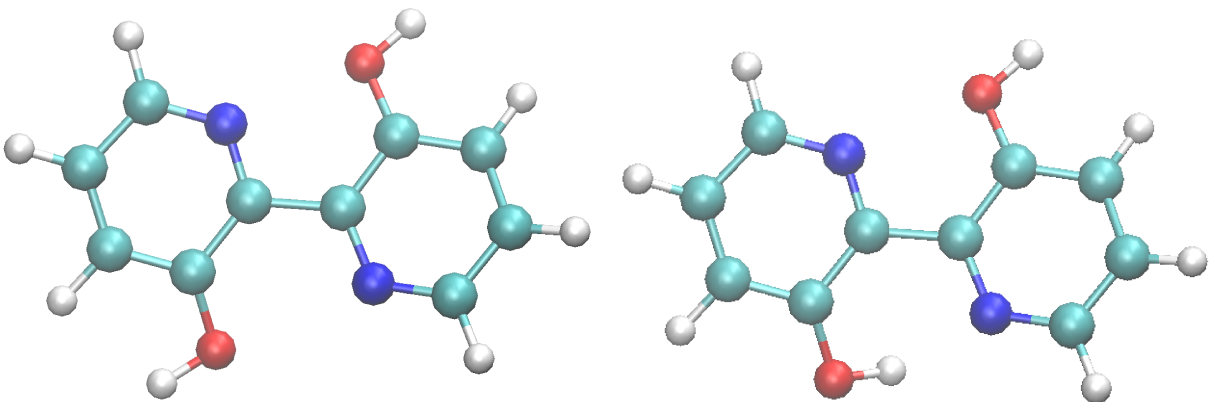
C.1 Conformers of BPOH2 187

C.2 FF parameters of diol 188

C.3 FF parameters of diketo 191

C.4 MD simulations, calculation of absorption spectra 194

C.1 Conformers of BPOH2



(a) open (b) one-open

Figure C.1: Conformers of BPOH2.

C.2 FF parameters of diol

Table C.1: Intermolecular parameters of the target molecule σ (Å), ϵ (kJ/mol) and q (e). Atom types are shown in Figure 4.1

Atom type	σ	ϵ	q
O	3.070	0.7113	-0.480078
H1	0.000	0.0000	0.278598
C1	3.550	0.2929	0.205145
N	3.250	0.7113	-0.195544
C2	3.550	0.2929	-0.067979
C3	3.550	0.2929	-0.169786
C4	3.550	0.2929	-0.116980
H2	2.420	0.1255	0.142100
H3	2.420	0.1255	0.162928
H4	2.420	0.1255	0.151536
C5	3.550	0.2929	0.090058
O_W	3.151	0.6364	-0.8340
H_W	0.000	0.0000	0.4170

Table C.2: Intramolecular parameters for stiff torsions: equilibrium dihedral angles ϕ^0 are in degrees and force constants k^t in kJ/mol rad^{-2}.

dihedral	ϕ^0	$\mathbf{k^t}$	**dihedral**	ϕ^0	$\mathbf{k^t}$
O_W-H_W-H_W-C2	0.0	52.48	H_W -H_W -C2 -N	0.0	52.48
H_W-C2-N-C5	0.0	52.48	C2 -N -C5 -O_W	0.0	52.48
H_W-O_W-C5-N	0.0	52.48	C5 -O_W -H_W -H_W	0.0	52.48
C1-C4-C3-C2	0.0	52.48	C4 -C3 -C2 -N	0.0	52.48
C3-C2-N-C5	0.0	52.48	C2 -N -C5 -C1	0.0	52.48
C4-C1-C5-N	0.0	52.48	C5 -C1 -C4 -C3	0.0	52.48
O-O_W-C5-C5	0.0	14.72	O -C1 -C5 -C5	0.0	14.72
H3-H_W-C2-H2	0.0	16.82	H4 -H_W -H_W -H3	0.0	14.02
O-O_W-H_W-H4	0.0	85.65	H3 -C3 -C2 -H2	0.0	16.82
H4-C4-C3-H3	0.0	14.02	O -C1 -C4 -H4	0.0	85.65
O-H_W-C5-O_W	0.0	182.77	H4 -O_W -H_W -H_W	0.0	286.72
H3-H_W-C2-H_W	0.0	420.22	H2 -H_W -N -C2	0.0	499.12
C5-O_W-N-C5	0.0	533.50	O -C4 -C5 -C1	0.0	182.77
H4-C1-C3-C4	0.0	286.72	H3 -C4 -C2 -C3	0.0	420.22
H2-C3-N-C2	0.0	499.12	C5 -C5 -C1 -N	0.0	562.48

Table C.3: FF parameters for non-bonded intramolecular interactions: σ^{intra} (Å) and ϵ^{intra} (kJ/mol).

Pair (ij)	σ_{ij}^{intra}	ϵ_{ij}^{intra}
N-H1	1.490	0.1000
H1-N	1.490	0.1000
H1-H1	1.800	0.0100

Table C.4: Intramolecular parameters for flexible torsions: number of cosines n, γ (°) and force constants k^d in kJ/mol.

dihedral	N_{cos}	n	k^d (kJ/mol)	γ (°)
C5-O_W-O-H1	6	0	28.734	0.00
		1	-36.083	0.00
		2	-16.166	0.00
		3	1.563	0.00
		4	0.438	0.00
		5	0.396	0.00
C5-C1-O-H1	6	0	28.734	0.00
		1	-36.083	0.00
		2	-16.166	0.00
		3	1.563	0.00
		4	0.438	0.00
		5	0.396	0.00
O_W-C5-C5-C1	6	0	28.734	0.00
		1	24.925	0.00
		2	-8.831	0.00
		3	2.482	0.00
		4	1.614	0.00
		5	-1.126	0.00
N-C5-C5-N	6	0	28.734	0.00
		1	24.925	0.00
		2	-8.831	0.00
		3	2.482	0.00
		4	1.614	0.00
		5	-1.126	0.00

C.3 FF parameters of diketo

Table C.5: Intermolecular parameters of the target molecule σ (Å), ϵ (kJ/mol) and q (e) Atom types are shown in Figure 4.1

Atom type	σ	ϵ	q
O	3.070	0.7113	-0.621401
H1	0.000	0.0000	0.300972
C1	3.550	0.2929	0.290275
N	3.250	0.7113	-0.067672
C2	3.550	0.2929	-0.118444
C3	3.550	0.2929	-0.190359
C4	3.550	0.2929	-0.126379
H2	2.420	0.1255	0.180509
H3	2.420	0.1255	0.180351
H4	2.420	0.1255	0.137252
C5	3.550	0.2929	0.034895
O_W	3.151	0.6364	-0.8340
H_W	0.000	0.0000	0.4170

Table C.6: Intramolecular parameters for stiff torsions: equilibrium dihedral angles ϕ^0 are in degrees and force constants k^t in kJ/mol rad^{-2}.

dihedral	ϕ^0	$\mathbf{k^t}$	**dihedral**	ϕ^0	$\mathbf{k^t}$
O_W-H_W-H_W-C2	0.0	4.11	C5 -O_W -H_W -H_W	0.0	4.11
H_W-O_W-C5-N	0.0	4.11	O_W -H_W -C2 -N	0.0	4.11
H_W-C2-N-C5	0.0	4.11	C2 -N -C5 -O_W	0.0	4.11
C5-C1-C4-C3	0.0	4.11	C4 -C1 -C5 -N	0.0	4.11
C1-C4-C3-C2	0.0	4.11	C4 -C3 -C2 -N	0.0	4.11
C3-C2-N-C5	0.0	4.11	C2 -N -C5 -C1	0.0	4.11
O-O_W-C5-C5	0.0	100.62	O -C1 -C5 -C5	0.0	100.62
H4-H_W-H_W-H3	0.0	23.01	H4 -C4 -C3 -H3	0.0	23.01
H3-H_W-C2-H2	0.0	23.01	H3 -C3 -C2 -H2	0.0	23.01
H2-C2-N-H1	0.0	23.01	O -O_W -H_W -H4	0.0	51.23
O-C1-C4-H4	0.0	51.23	O -H_W -C5 -O_W	0.0	773.80
H4-O_W-H_W-H_W	0.0	327.66	H3 -H_W -C2 -H_W	0.0	426.48
H2-H_W-N-C2	0.0	358.98	H1 -C2 -C5 -N	0.0	788.35
C5-O_W-N-C5	0.0	298.42	O -C4 -C5 -C1	0.0	773.80
H4-C1-C3-C4	0.0	327.66	H3 -C4 -C2 -C3	0.0	426.48
H2-C3-N-C2	0.0	358.98	C5 -C5 -C1 -N	0.0	298.42

Table C.7: Intramolecular parameters for flexible torsions: number of cosines n, γ (deg) and force constants k^d in kJ/mol.

dihedral	N_{cos}	n	k^d (kJ/mol)	γ (deg)
O_W-C5-C5-C1	6	0	12.359	0.00
		1	47.082	0.00
		2	-14.155	0.00
		3	10.820	0.00
		4	1.432	0.00
		5	1.075	0.00
N-C5-C5-N	6	0	12.359	0.00
		1	47.082	0.00
		2	-14.155	0.00
		3	10.820	0.00
		4	1.432	0.00
		5	1.075	0.00

Table C.8: FF parameters for non-bonded intramolecular interactions: σ^{intra} (Å) and ϵ^{intra} (kJ/mol).

Pair (ij)	σ_{ij}^{intra}	ϵ_{ij}^{intra}
O-H1	1.650	0.1000
O-H1	1.650	0.1000
H1-H1	1.780	0.0100
O-O	2.850	0.2000

C.4 MD simulations, calculation of absorption spectra

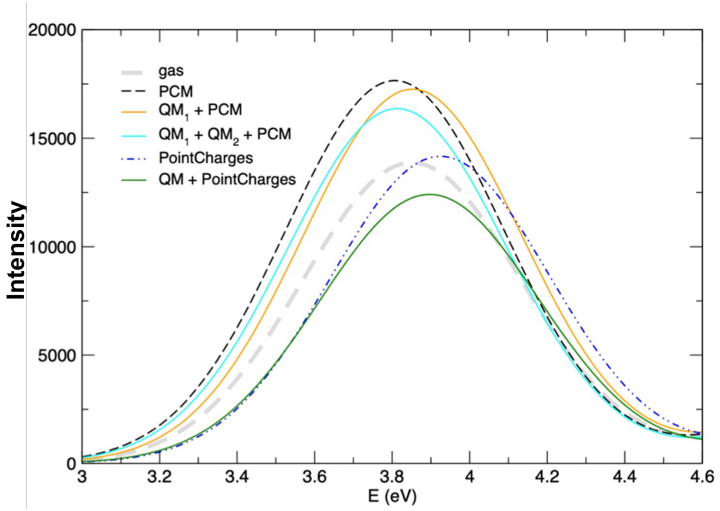


Figure C.2: Diol tautomer: absorption spectrum computed using different solvation schemes (courtesy of Dr. Giacomo Prampolini). PCM: mechanical embedding. QM1+PCM: explicit water molecules, hydrogen bonded to the solute, are included in the QM part and the rest of the solvent is represented using CPCM. QM1+QM2+PCM: explicit water molecules, hydrogen bonded to the solute, are included in the QM part together with water molecules from the second solvation shell. The rest of the solvent is represented using CPCM. Point charges: the QM solute is surrounded by point charges representing the solvent. QM+point charges: the solute and explicit solvent molecules are included in the QM part, whereas the rest of the solvent is represented as point charges.

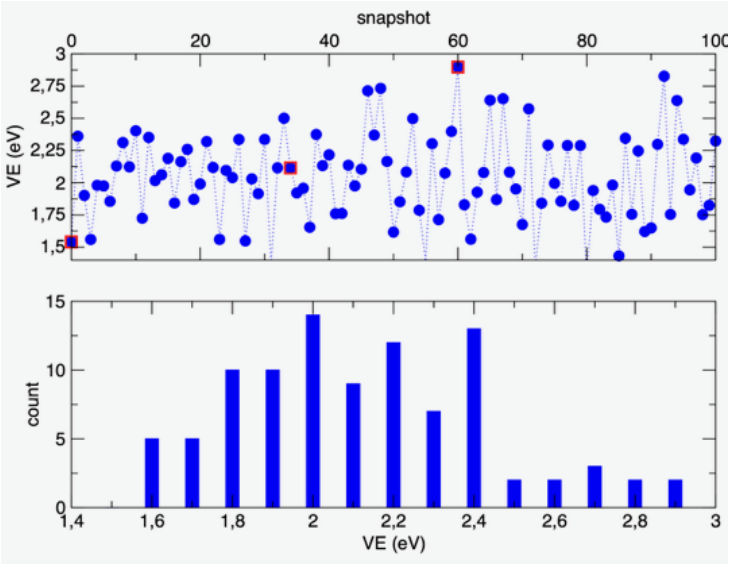


Figure C.3: Diketo tautomer: distribution of vertical excitation energies.

List of Figures

1 Structure of the aspartate anions (which will be named ASP^{-} and ASP^{2-}, depending on the protonation state) compared with the structure of the succinate anion (SUC^{2-}). 3

2 The structure of [2,2'-bipyridyl]-3,3'-diol 8

1.1 Flow chart of the Joyce protocol. 46

2.1 Top panel: structure of the three conformers of succinate and definition of the dihedral used in the analysis. SUC_1^{2-}, SUC_2^{2-}, SUC_3^{2-} correspond to the S1, S2, S3 conformers in Reference [4], respectively. Bottom panel: structure of the optimized clusters formed by SUC_1^{2-} and SUC_2^{2-} and two water molecules next to the charged groups. Green dotted lines are shown as a guide for the eye along the succinate-water hydrogen bond directions. 64

2.2 Infrared spectra obtained using different implicit solvation models (top panel) and incorporating two interacting water molecules (bottom panel). A comparison is provided with an experimental FTIR spectrum of succinate in water solution[37]. Intensities are expressed in arbitrary units. . . . 67

2.3 Top panel: structure of the two conformers of deprotonated aspartate and definition of the relevant dihedrals used in the analysis. Bottom panel: structure of the optimized clusters formed by ASP_1^- and ASP_2^- and three water molecules next to the charged groups. Relevant H-bond directions are displayed as a guide for the eye. Green dotted lines: intermolecular H-bonds, water is the donor. Blue dotted lines: intermolecular H-bond, water is the acceptor. Yellow dotted line: intramolecular H-bond. 71

2.4 Infrared spectra for the two minima ASP_1^- and ASP_2^-. Top panel: results for the anions in a continuum solvent. Bottom panel: results for the anions surrounded by three water molecules next to the polar groups in PCM. Experimental frequencies estimated in Reference[30] for different modes are displayed as black lines. Intensities are expressed in arbitrary units. . . 73

2.5 Top panel: structure of the two conformers of aspartate having a deprotonated amino group (ASP_1^{2-}, ASP_2^{2-}) and definition of the relevant dihedrals used in the analysis. Bottom panel: structure of the optimized clusters formed by ASP_1^{2-} and ASP_2^{2-} and three water molecules sitting next to the charged groups. Green dotted lines: intermolecular H-bonds, water is the donor. Yellow dotted line: intramolecular H-bond. 76

2.6 Infrared spectra for the two minima ASP_1^{2-} and ASP_2^{2-}. Top panel: results for the anions in a continuum solvent. Bottom panel: results for the anions surrounded by three water molecules next to the polar groups in PCM. Experimental frequencies estimated in Reference [30] for different modes are displayed as black lines. Intensities are expressed in arbitrary units. . . 78

2.7 Potential of mean force obtained by umbrella sampling along the χ dihedral for the three molecules considered in this work: succinate (SUC^{2-}, blue curve), deprotonated aspartate (ASP^{-}, black curve), deprotonated aspartate in a basic environment (ASP^{2-}, red curve). 80

2.8 Relevant solute-water pair distribution functions. Top panel: interaction between oxygen atoms belonging to carboxylates in succinate (black, red, green, blue curves). Middle panel, from left to right: interaction between oxygen atoms belonging to carboxylates in ASP^{-} (black, red, green, blue curves); interaction with hydrogen atoms belonging to the amino group; interaction with the nitrogen atom belonging to the amino group. Bottom panel: same as for the middle panel, in the case of ASP^{2-}. For each panel, images illustrating the structure of the solvation shell created by extracting snapshots along the simulated trajectories are displayed. The software used to create the images is VMD.[38] 82

2.9 Infrared absorption spectra of the three anions considered. Magenta line: result based on the maximum entropy method. Blue line: result based on the Wiener-Khinchin theorem. Intensities are expressed in arbitrary units. 85

3.1 The structure of $[Mg_2Al_4(OH)_{12}]_n$ 98

3.2 OH...OOC interaction between the layer and the organic intercalates . . . 105

3.3 OH...HN interaction between the layer and the organic intercalates 106

3.4 OH...OH2 interaction between the layer and water molecules 107

3.5 [HOH...HNH interaction between water molecules and the organic intercalates 108

3.6 H2O...HN interaction between water molecules and the organic intercalates 108

3.7 HOH...COO interaction between water molecules and the organic intercalates 109

3.8 H2O...HOH pair correlation functions for the interaction among intercalated water molecules 110

3.9 COO...HN interaction among organic intercalates 112

3.10 The schematic of the orientation of organic intercalates 113

3.11 Snapshot representing conformations of ASP^{2-} intercalated in LDH at two different hydration states. 115

3.12 Snapshot representing conformations of GLU^{2-} intercalated in LDH at two different hydration states. 116

3.13 Snapshot representing conformations of SUC^{2-} intercalated in LDH at two different hydration states. 116

3.14 Binding modes of carboxylates coordinated with a metal M and analogy with binding modes of the confined anions with the surface. 118

3.15 The graphs of number density for ASP 119

3.16 The graph of number density for GLU 120

3.17 The graph of number density for SUC 121

3.18 Experimental and simulated carboxyl stretch band of ASP^{2-} 123

3.19 Experimental and simulated carboxyl stretch band of GLU^{2-} 124

3.20 Experimental and simulated Carboxylate stretch band of SUC^{2-} 125

4.1 Tautomers of BPOH2 137

4.2 The experimental absorption spectra of BPOH2, Figure extracted from Ref.[12]. 142

4.3 Relaxed scan of δ_{Ph} and δ_{OH} at a QM level for diol. 143

4.4 Relaxed scan of δ_{Ph} in the diketo tautomer at the QM level. 144

4.5 The absorption spectra of diol, keto-enol, and diketo were obtained through QM calculations. 145

4.6 Tautomers of BPOH2 146

4.7 diol: relaxed scans of δ_{OH} and δ_{Ph} from QM and from the parametrized FF. Results are also shown for the Frozen Internal Rotation Approximation described in Chapter 1. 147

4.8 diketo: relaxed scans of δ_{OH} and δ_{Ph} from QM and from the parametrized FF. Results are also shown for the Frozen Internal Rotation Approximation described in Chapter 1. 147

4.9 Comparison of the frequencies of normal modes of two tautomers obtained from QM and MM. 148

4.10 Distribution of flexible dihedrals in diol. 149

4.11 Diketo tautomer: the distribution of N5-C6-C17-C12 dihedral. 150

4.12 Diol tautomer: distribution of N...H and O-H distances. 150

4.13 Diol tautomer: radial distribution function describing the HO...HOH interactions. 151

4.14 Diketo tautomer: distribution of N-H and O...H distances. 151

4.15 Diketo tautomer: radial distribution function describing the O...HOH interactions. 152

4.16 Absorption spectra of two tautomers obtained at the QM level by using mechanical embedding (the colored lines are the spectra computed for the snapshots and their convolution is shown as the black line). 153

A.1 Normalized distribution of two relevant dihedrals characterizing the conformations of the three anions considered in water. 169

A.2 Normalized distribution of the $C_\gamma OO^-$...HNH_2^+ distances correspond to the formation of an intramolecular hydrogen bond in aspartate (ASP^-). 170

A.3 Orientational correlations based on the calculation of the $G_1(r)$) normalized function. The different colors refer to the four different O atoms of the carboxylate groups. 170

B.1 Adsorption isotherms of LDH intercalated with organic intercalates 172

B.2 Experimental IR of intercalates 173

B.3 Simulated systems of ASP 179

B.4 Simulated systems of GLU 180

B.5 Simulated systems of SUC 181

B.6 Distribution of the C_α-C_β-C_γ dihedral in the system intercalating aspartate with increasing water content. 182

B.7 Distribution of the C_α-C_β-C_γ dihedral in the system intercalating glutamate with increasing water content. 182

B.8 Distribution of the C-C-C-C dihedral in the system intercalating succinate with increasing water content. 183

C.1 Conformers of BPOH2. 187

C.2 Diol tautomer: absorption spectrum computed using different solvation schemes (courtesy of Dr. Giacomo Prampolini). PCM: mechanical embedding. QM1+PCM: explicit water molecules, hydrogen bonded to the solute, are included in the QM part and the rest of the solvent is represented using CPCM. QM1+QM2+PCM: explicit water molecules, hydrogen bonded to the solute, are included in the QM part together with water molecules from the second solvation shell. The rest of the solvent is represented using CPCM. Point charges: the QM solute is surrounded by point charges representing the solvent. QM+point charges: the solute and explicit solvent molecules are included in the QM part, whereas the rest of the solvent is represented as point charges. 194

C.3 Diketo tautomer: distribution of vertical excitation energies. 194