



**HAL**  
open science

# Description topologique des phénomènes d'hydratation et développement méthodologique de fonctionnelles doubles hybrides à séparation de portée

Cairedine Kalai

## ► To cite this version:

Cairedine Kalai. Description topologique des phénomènes d'hydratation et développement méthodologique de fonctionnelles doubles hybrides à séparation de portée. Chimie théorique et/ou physique. Sorbonne Université, 2018. Français. NNT : 2018SORUS191 . tel-02448370

**HAL Id: tel-02448370**

**<https://theses.hal.science/tel-02448370>**

Submitted on 22 Jan 2020

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

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



SORBONNE UNIVERSITÉ  
ECOLE DOCTORALE : ED388  
LABORATOIRES : MONARIS & LCT

DESCRIPTION TOPOLOGIQUE DES PHÉNOMÈNES  
D'HYDRATATION ET DÉVELOPPEMENT MÉTHODOLOGIQUE  
DE FONCTIONNELLES DOUBLES HYBRIDES  
À SÉPARATION DE PORTÉE

PAR CAIRELINE KALAI  
THÈSE DE DOCTORAT EN CHIMIE THÉORIQUE  
DIRIGÉE PAR ESMAÏL ALIKHANI & JULIEN TOULOUSE

PRÉSENTÉE ET SOUTENUE PUBLIQUEMENT LE 4 JUILLET 2018  
DEVANT UN JURY COMPOSÉ DE :

Virginie MARRY	Professeur	Sorbonne Université	(présidente)
Thierry LEININGER	Professeur	Université Paul Sabatier	(rapporteur)
Marie-Christine BACCHUS	Chargée de recherche (HDR)	Université Lyon 1	(rapporteuse)
Laurent JOUBERT	Professeur	Université de Rouen	(examinateur)
Ilaria CIOFINI	Directeur de recherche	Chimie Paristech	(examinatrice)
Valérie BRENNER	Chercheur CEA (HDR)	CEA	(examinatrice)
Julien TOULOUSE	MCF (HDR)	Sorbonne Université	(co-directeur)
Esmail ALIKHANI	Professeur	Sorbonne Université	(directeur)



# Table des matières

<b>I</b>	<b>Rappels théoriques</b>	<b>15</b>
<b>1</b>	<b>Les méthodes de fonction d'onde</b>	<b>16</b>
1.1	Hamiltonien électronique non relativiste dans l'approximation de Born-Oppenheimer . . . . .	17
1.2	Spin-orbitales et déterminant de Slater . . . . .	19
1.3	Le théorème variationnel . . . . .	21
1.4	L'approximation Hartree-Fock . . . . .	21
1.4.1	Observables calculées avec un déterminant de Slater . . . . .	21
1.4.2	Les équations Hartree-Fock . . . . .	22
1.4.3	Les formalismes "restricted" et "unrestricted" . . . . .	24
1.4.4	Les équations de Roothaan . . . . .	26
1.5	Au delà de la méthode Hartree-Fock . . . . .	28
1.5.1	La méthode d'interaction de Configuration . . . . .	28
1.5.2	La théorie des perturbations . . . . .	29
<b>2</b>	<b>La théorie de la fonctionnelle de la densité</b>	<b>34</b>
2.1	Théorèmes fondateurs d'Hohenberg-Kohn . . . . .	35
2.2	Formulation de Levy-Lieb . . . . .	36
2.3	L'approche Kohn-Sham . . . . .	37
2.3.1	L'approche LDA . . . . .	38
2.3.2	Tenir compte des inhomogénéités, les approches semi-locales et hybrides . . . . .	39
2.4	Au delà de l'approche Kohn-Sham . . . . .	41
2.4.1	État de l'art sur les méthodes doubles hybrides . . . . .	41
2.4.2	Extension multi-déterminantale basée sur une connexion adiabatique linéaire . . . . .	47

2.4.3	Extension multi-déterminantale basée sur une connexion adiabatique avec séparation de portée . . . . .	50
<b>3</b>	<b>Les méthodes interprétatives permettant une caractérisation des interactions mises en jeu au sein des complexes micro-hydratés</b>	<b>55</b>
3.1	Utilisation du MEP pour la prédiction des complexes interagissant par liaison hydrogène . . . . .	56
3.2	L'analyse de la densité électronique par la théorie QTAIM . . . . .	58
<b>II</b>	<b>Développement Méthodologique</b>	<b>62</b>
<b>4</b>	<b>A general range-separated double-hybrid density-functional theory</b>	<b>63</b>
4.1	Abstract . . . . .	64
4.2	Introduction . . . . .	64
4.3	Exact theory . . . . .	67
4.4	Single-determinant approximation . . . . .	70
4.5	Second-order Møller–Plesset perturbation theory . . . . .	71
4.6	Complement short-range exchange-correlation density functional . . . . .	74
4.6.1	Decomposition into exchange and correlation . . . . .	74
4.6.2	Expression of $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$ in the Coulomb limit $\mu \rightarrow 0$ and in the high-density limit . . . . .	75
4.6.3	Expression of $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$ in the short-range limit $\mu \rightarrow \infty$ and the low-density limit . . . . .	76
4.6.4	Assessment of the approximations for $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$ on the uniform-electron gas . . . . .	79
4.6.5	Computational details . . . . .	81
4.7	Results and discussion . . . . .	83
4.7.1	Optimization of the parameters on the AE6 and BH6 sets . . . . .	83
4.8	Assessment on the AE49 and DBH24/08 sets of atomization energies and reaction barrier heights . . . . .	84
4.8.1	Assessment of the basis convergence . . . . .	85
4.8.2	Assessment on the S22 set of intermolecular interactions . . . . .	86
4.9	Conclusion . . . . .	87
4.10	Uniform coordinate scaling relation and Coulomb/high-density and short-range/low-density limits of $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$ . . . . .	88
4.10.1	Scaling relation for $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$ . . . . .	88

4.10.2	Coulomb limit and high-density limit of $\bar{E}_c^{sr,\mu,\lambda}[n]$ . . . . .	90
<b>III</b>	<b>Applications</b>	<b>96</b>
<b>5</b>	<b>Etude des complexes <math>\beta</math>-propiolactone micro-hydratés.</b>	<b>97</b>
5.1	De l'importance de l'étude des complexes micro-hydratés. . . . .	98
5.2	Caractérisation des régions nucléophiles et électrophiles de la BPL - étape 1	101
5.3	Etude du complexe BPL(H <sub>2</sub> O). . . . .	103
5.3.1	Identification des structures initiales - étape 2 . . . . .	103
5.3.2	Etude de la PES pour les complexes BPL(H <sub>2</sub> O) - étape 3 . . . . .	105
5.3.3	Caractérisation QTAIM des conformères BPL :H <sub>2</sub> O - étape 4 . . . . .	110
5.4	Etude du complexe BPL(H <sub>2</sub> O) <sub>2</sub> . . . . .	114
5.4.1	Caractérisation des régions nucléophiles et électrophiles des complexes BPL(H <sub>2</sub> O) et du dimère d'eau - étape 1 . . . . .	114
5.4.2	Proposition de structures initiales pour le complexe di-hydraté - étape 2 . . . . .	119
5.4.3	Etude de la PES pour les complexes di-hydratés - étape 3 . . . . .	121
5.4.4	Caractérisation QTAIM des conformères BPL(H <sub>2</sub> O) <sub>2</sub> - étape 4 . . . . .	126
5.5	Etude du complexe BPL(H <sub>2</sub> O) <sub>3</sub> . . . . .	128
5.5.1	Caractérisation des régions nucléophiles et électrophiles des complexes BPL(H <sub>2</sub> O) <sub>2</sub> et du trimère d'eau - étape 1 . . . . .	128
5.5.2	Proposition de structures initiales pour le complexe tri-hydraté - étape 2 . . . . .	131
5.5.3	Etude de la PES pour les complexes tri-hydratés - étape 3 . . . . .	133
5.6	Caractérisation QTAIM des conformères BPL(H <sub>2</sub> O) <sub>3</sub> - étape 4 . . . . .	136
<b>6</b>	<b>A theoretical investigation of watersolute interactions : from facial parallel to guesthost structures</b>	<b>138</b>
6.1	Abstract . . . . .	139
6.2	Introduction . . . . .	139
6.3	Method calibration . . . . .	142
6.4	BPL :(H <sub>2</sub> O) <sub>4</sub> and BPL :(H <sub>2</sub> O) <sub>5</sub> complexes : A MESP picture to rationalize the structures experimentally deduced . . . . .	145
6.5	BPL :(H <sub>2</sub> O) <sub>6</sub> complex . . . . .	148
6.6	BPL :(H <sub>2</sub> O) <sub>10</sub> complex . . . . .	148
6.7	BPL in or out of water cluster : BPL :(H <sub>2</sub> O) <sub>20</sub> and BPL :(H <sub>2</sub> O) <sub>24</sub> complexes	152

6.8 Conclusion . . . . .	154
6.9 Supporting Information . . . . .	156
References . . . . .	193

## Remerciements

Je voudrais remercier les différents acteurs qui ont rendu cette thèse possible : les laboratoires MONARIS et LCT dirigés respectivement par Christophe PETIT et Olivier PARISEL, pour m'avoir accueilli pendant ces trois années, l'école doctorale ED388 et le Labex Michem pour avoir financée ma thèse.

Ensuite, je voudrais remercier mon directeur de thèse, Mohammad Esmail Alikhani, mon co-directeur de thèse Julien TOULOUSE et Emilie-Laure ZINS. Je les remercie pour leurs gentillesse, leurs patiences, leurs disponibilités dont ils ont fait preuve avec constance.

Je voudrais également remercier les membres de mon jury de thèse, Mme Virginie MARRY, M Thierry LEININGER, Mme Marie-Christine BACCHUS, M Laurent JOUBERT, Mme Ilaria CIOFINI et Mme Valérie BRENNER d'avoir accepté de juger ce travail et de participer au jury de cette thèse.

Concernant l'aspect informatique de cette thèse, je tiens à remercier Peter REINHARDT, pour son aide en programmation, Bastien MUSSARD pour son aide concernant le logiciel MOLPRO. Je voudrais également remercier Bruno MADEBENE, maitre de conférence et également ingénieur système du laboratoire MONARIS et Marie-France COURET, administrateur systèmes et réseau du laboratoire LCT.

Un grand merci à tous les membres des laboratoires MONARIS et LCT pour les discussions scientifiques ou non et pour tous les moments que nous avons partagés. Je remercie particulièrement les doctorants Sendres, Marlène, Mindaugas, Salvatore, Alejandro, Yacine, Jing wen, Mindaugas, Mohamed, Olivier, Maya, Julien, Frédéric et Andi.

Enfin, un grand merci à ma famille (en particulier à mes parents) et mes amis de m'avoir entouré et supporté pendant ces trois années. Il est délicat d'écrire des remerciements et de n'oublier personne. Je m'excuse auprès de ceux qui ne seront pas mentionnés, qu'ils veuillent bien me pardonner.

Paris, le 28 mai 2018

# Introduction générale

L'existence de la matière condensée est une attestation de la présence de forces attractives ou d'interactions non-covalentes entre les molécules. On peut classer les interactions non-covalentes selon leur portée : les interactions à courte portée pour lesquelles l'énergie décroît de manière exponentielle avec la distance et les interactions à longue portée qui présente une décroissance en  $1/R^n$  ( $R$  étant la distance intermoléculaire et  $n$  pouvant prendre des valeurs particulières entières qui seront précisées ci-dessous). L'énergie d'interaction  $U$  entre deux atomes est une fonction de la distance intermoléculaire  $R$  et prend la forme indiquée dans la figure 1 :

Il existe une distance d'interaction  $R_m$  pour laquelle l'énergie d'interaction atteint un minimum (noté  $-\Omega$ ). A la distance d'interaction  $\sigma$ , l'énergie d'interaction est nulle.

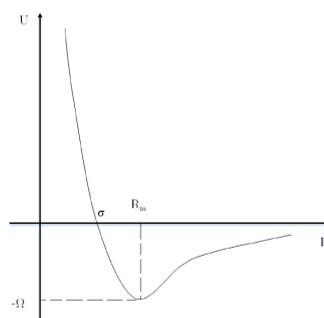


FIGURE 1 – Illustration d'une courbe d'énergie potentielle pour une paire d'atomes.

Dans le cas d'une paire d'atomes, la courbe d'énergie potentielle ne fait intervenir qu'un seul paramètre géométrique. Dans les autres cas, la courbe d'énergie potentielle est remplacée par une surface d'énergie potentielle qui est une fonction dépendant de plusieurs variables géométriques. Dans le cadre de ma thèse, je me suis essentiellement intéressé aux interactions à longue portée qui sont responsables de l'attraction entre des molécules. Ces interactions à longue portée peuvent être classées en 4 catégories, charge-charge, dipôle-permanent - dipôle-permanent, induction et dispersion [1] :

- Les interactions entre molécules chargées sont inversement proportionnelles à la distance intermoléculaire ( $1/R$ ). Ces interactions ont la particularité de ne pas présenter d'effets à plusieurs corps.
- Les interactions entre dipôles permanents sont également appelées "force de Keesom". Elles existent uniquement entre molécules polaires et décroissent en  $1/R^6$ . D'autres contributions en  $1/R^5$  et  $1/R^4$  sont à prendre en compte, elles traduisent respectivement les interactions entre moments quadripolaires et les interactions dipôle-quadripôle.
- Les interactions entre un dipôle induit et un dipôle permanent appelées également "forces de Debye" résultent du fait que la distribution de charge d'une molécule va être perturbée en raison de la présence de molécules voisines et réciproquement.

Ces interactions ont un comportement asymptotique en  $1/R^6$ .

- Les interactions de nature dispersive sont liées à des effets de corrélation électronique à longue portée. Ces interactions ne requièrent ni une polarité des deux fragments ni un recouvrement orbitalaire des fonctions d'onde impliquées. Ces interactions sont toujours attractives et ont un comportement asymptotique en  $1/R^6$ .

Les interactions à longue portée charge-charge, dipôle permanent - dipôle permanent, induction et dispersion sont appelées ainsi car elles sont les seules à subsister lorsque la distance intermoléculaire tend vers l'infini. Il faut noter que ces interactions sont également présentes à courte distance. Comme il a été énoncé précédemment, les interactions à longue portée sont usuellement décrites en une série en  $1/R^n$ . Une telle série ne peut que diverger lorsque la distance intermoléculaire tend vers zéro. D'autres contributions sont à prendre en compte lorsque  $R \rightarrow 0$ . Aux courtes distances intermoléculaires, le phénomène d'échange intermoléculaire est présent. Il s'agit d'un terme répulsif qui est dû à l'antisymétrisation de la fonction d'onde. En effet, la fonction d'onde du système ne peut être considérée comme le produit des fonctions d'onde des monomères.

La décomposition de l'énergie d'interaction en termes d'interaction charge-charge, dipôles-permanents - dipôle -permanent, induction et dispersion peut être considérée comme une vision du physicien. Les méthodes SAPT (de l'anglais pour "Symmetry-adapted perturbation theory") décomposent l'énergie d'interaction en ces diverses contributions [2]. Il existe d'autres façons de décomposer l'énergie d'interaction :

- l'approche IQA (de l'anglais pour "Interacting Quantum Atoms") [3] repose sur la partition de l'espace réel selon la théorie "*Atoms In Molecules*" (AIM) de Richard Bader [4]. Cette approche considère l'énergie d'interaction comme une somme d'interaction intra-atomique et inter-atomique. Chaque atome est défini par un bassin suite à l'application de la théorie AIM.
- Nous pouvons également citer la décomposition énergétique EDA (de l'anglais pour "Energy Decomposition Analysis") [5]. Dans cette approche, l'énergie d'interaction se décompose selon cinq composantes, électrostatique (comprenant les interactions entre charges et multipôles permanents), polarisation (comprenant les interactions entre multipôles permanents et multipôles induits), transfert de charge, échange-répulsion et un terme dit *MIX* défini comme étant la différence entre l'énergie d'interaction et les quatre premiers termes.

Cette thèse s'intéresse aux phénomènes d'hydratation de composés organiques à l'échelle moléculaire et place la liaison hydrogène au cœur du travail. La liaison hydrogène est

simplement un cas particulier d'interaction qui pourrait être vu comme une somme d'interactions issues d'une décomposition énergétique (SAPT, IQA ou EDA), mais en raison de son importance dans les milieux biologiques et atmosphériques entre autre [6] [7], elle est souvent considérée comme une interaction à part entière. Une définition précise de la liaison hydrogène a été donnée par les recommandations de l'IUPAC [8] (de l'anglais "International Union of Pure and Applied Chemistry") : *la liaison hydrogène est une interaction attractive entre un atome d'hydrogène d'une molécule X-H - dans laquelle le groupement X est plus électronégatif que l'atome H - et un atome ou un fragment d'une molécule appartenant soit à la même molécule ou à une molécule différente, dans lequel il existe une évidence d'une formation d'une interaction. Une liaison hydrogène typique peut être labellisé ainsi X-H ... Y-Z où les trois points représentent la liaison hydrogène. X-H représente le donneur d'hydrogène et Y l'accepteur. Dans tous les cas, l'accepteur possède une région riche en électrons, comme par exemple un doublet non-liant ou des électrons  $\pi$ . L'évidence de la formation de la liaison hydrogène peut être expérimentale ou théorique, ou idéalement une combinaison des deux.* Nous nous intéressons uniquement dans le cadre de cette thèse aux liaisons hydrogènes non-covalentes [9] [10].

D'un point de vue expérimental, différentes techniques permettent d'identifier ces interactions. En particulier, la spectroscopie infrarouge met en évidence la liaison hydrogène en mesurant le décalage de la fréquence d'élongation du groupement X-H. La RMN du proton mesure la différence du déplacement chimique du proton impliqué dans la liaison hydrogène. Une autre manière de mettre en évidence expérimentalement une formation de liaison hydrogène est de recourir à la spectroscopie micro-ondes (généralement couplée à une expansion supersonique) dans laquelle on effectue des substitutions isotopiques [11]. Une variation du moment d'inertie est alors mesurée permettant la détermination des positions atomiques.

D'un point de vue théorique la liaison hydrogène peut être mise en évidence par l'étude de la topologie de la liaison chimique. L'étude de la topologie de la densité électronique est un moyen rigoureux de caractériser des interactions inter- ou intra-moléculaires telles que les liaisons hydrogènes. Une étape significative vers une définition rigoureuse de la liaison hydrogène a été accomplie par Koch et Popelier [12], basée sur la théorie "AIM de Richard Bader [4]. La théorie AIM est un moyen de mettre en évidence entre autres des liaisons hydrogènes selon l'IUPAC [8]. Il existe d'autres méthodes topologiques comme NCI ("Non-Covalent Interaction") qui utilisent le gradient réduit de la densité pour révéler les interactions non-covalentes [13]. La topologie de la fonction ELF peut aussi être appliquée à l'étude des complexes à liaison hydrogène [14].

L'approche NBO (de l'anglais pour "Natural Bond Orbitals") qui repose sur la connais-

sance de la matrice densité d'ordre un permet de décrire la structure électronique de la molécule dans un schéma de type Lewis [15]. Les méthodes énoncées précédemment permettent de caractériser *a posteriori* la liaison hydrogène. Il nous faut également une méthode capable de prédire la formation de ces interactions non-covalentes. L'étude du potentiel électrostatique moléculaire (MEP) sur une surface d'iso-densité permet de déterminer les régions nucléophiles et les régions électrophiles [16].

Lorsqu'il existe une chaîne de liaisons hydrogènes X-H ... Y-H ...Z, les deux liaisons hydrogènes peuvent se renforcer mutuellement, on parle d'effet coopératif [17]. Les interactions à plusieurs corps (trois corps, quatre corps ...) sont les principales sources des effets coopératifs. Les interactions de nature inductive sont souvent reliées aux effets coopératifs puisqu'elles présentent également un caractère non-additif [1]. La contribution des interactions de nature dispersive à plusieurs corps à l'énergie d'interaction est relativement faible [1]. Cependant, l'importance de la dispersion croît avec la taille du système [18].

La caractérisation théorique d'une interaction nécessite au préalable la connaissance de la fonction d'onde du système. Il nous faut dans un premier temps une méthode fiable permettant de traiter ces interactions. L'équation de Schrödinger n'étant soluble analytiquement que pour quelques systèmes simples, l'objectif premier de la chimie quantique a été de trouver des résolutions approchées de cette équation pour les systèmes poly-électroniques. Une première façon consiste à approximer directement la fonction d'onde du système : on parle de méthodes WFT (de l'anglais pour "Wave Function Theory"). La description quantitative des propriétés chimiques ne peut être réalisée qu'en prenant en compte l'énergie de corrélation. De nombreuses méthodes reposant sur la construction explicite d'une approximation multi-déterminantale pour la fonction d'onde de l'état fondamental ont été développées, permettant d'approcher l'énergie de l'état fondamental avec précision [19].

Cependant, ces méthodes sont limitées par le coût de calcul qui croît de manière exponentielle avec la taille du système étudié. La méthode MP2 est la méthode post-Hartree-Fock la moins coûteuse en temps calcul (le coût de calcul étant en  $N_{\text{basis}}^5$ , mais ne permet pas un traitement rigoureux des interactions de nature dispersive. En effet, cette méthode est connue pour surestimer les interactions de nature dispersive [20]. Des méthodes de calcul plus sophistiquées telles que la méthode CCSD(T) permettent de tenir compte des effets de dispersion à plusieurs corps manquant dans la théorie MP2. En revanche, la méthode CCSD(T) possède un temps de calcul qui croît en  $N_{\text{basis}}^7$ . A cela, il faut ajouter une convergence lente de l'énergie de corrélation [21] et donc des temps de calcul qui deviennent rapidement prohibitifs.

Une seconde approche pour résoudre l'équation de Schrödinger repose sur la théorie de

la fonctionnelle de la densité (DFT). Actuellement, la DFT est appliquée dans le formalisme de Kohn-Sham [22]. La caractéristique principale réside dans l'emploi d'un système fictif d'électrons n'interagissant pas entre eux et dont la densité est la même que celle du système physique. La difficulté majeure dans cette famille de méthode réside dans l'approximation de la fameuse fonctionnelle d'échange-corrélation regroupant les interactions entre électrons. La fonctionnelle d'échange-corrélation est la seule partie inconnue et par conséquent approximée dans le formalisme de Kohn-Sham.

Les approximations habituelle dans la fonctionnelle d'échange-corrélation sont qualifiées de semi-locales. La fonctionnelle d'échange-corrélation dépend de la valeur de la densité au point  $r$ , mais fait également intervenir des dérivées premières, voir les dérivées d'ordre supérieur. De ce fait, les approximations semi-locales sont incapables de décrire les interactions de nature dispersives.

Un moyen simple de rendre compte de l'énergie de dispersion manquante est d'ajuster un certain nombre de paramètres à des énergies d'interaction [23] [24]. L'avantage de cette méthodologie est que les effets de dispersion se ressentent sur la densité, l'inconvénient majeur est que le comportement asymptotique en  $1/R^6$  est absent. Un autre moyen simple consiste à considérer que les approximations qui sont faites dans la fonctionnelle d'échange-corrélation tiennent compte des autres interactions et qu'il suffit d'ajouter un terme *a posteriori* pour tenir compte de l'énergie dispersion. L'avantage est la possibilité d'imposer le comportement asymptotique en  $1/R^6$ , l'inconvénient est que les effets de la dispersion sur la densité ne sont pas pris en compte. Ces approches sont connues sous le nom d'approches C6 [25] [26].

D'autres alternatives pour la prise en compte des effets de dispersion reposent sur l'emploi de méthodes hybrides fonction d'onde/DFT. Ceci peut s'effectuer en introduisant une séparation de portée dans le traitement des interactions électroniques. Les méthodes doubles hybrides à séparation portée combinent le meilleur des deux mondes [27]. L'emploi d'une séparation de portée permet une convergence exponentielle de l'énergie de corrélation [28, 29], le problème lié au *cusp* de la fonction d'onde électronique n'existe plus, les interactions électroniques de courte portée sont prises en compte au niveau DFT. Quant aux interactions de longue portée comprenant entre autres les interactions de nature dispersive, elles sont prises en compte par des méthodes post Hartree-Fock [30]. Avec les méthodes doubles hybrides à séparation de portée, il est également possible de prendre en compte les effets dispersion à plusieurs corps [31]. Le comportement asymptotique en  $1/R^6$  est respecté d'une part et les effets des interactions de nature dispersive sur la densité sont incluses d'autre part. Les approches doubles hybrides à séparation de portée sont plus coûteuses en temps de calcul que les approches C6, mais elles reposent sur un for-

malisme exact. Cependant, les méthodes doubles hybrides à séparation de portée utilisées actuellement peuvent être moins efficaces pour traiter les interactions non-covalentes que des méthodes recourant à une correction empirique telles les approches DFT-D telle que LC- $\omega$ PBE-D3 ou B3LYP-D3 [32].

L'un des objectifs de cette thèse consiste à améliorer les méthodes doubles hybrides à séparation de portée dans le traitement des interactions non-covalente et en particulier la liaison hydrogène. Pour cela, nous envisageons d'inclure une fraction d'échange exact et de corrélation calculée au niveau MP2 dans la région de courte portée. En effet, des études statistiques sur des systèmes liés par liaison hydrogène ont montré que l'inclusion d'une fraction d'échange exact dans la région de courte portée permet de réduire les erreurs sur les énergies d'interaction [33]. L'inclusion d'une fraction d'échange exacte et de corrélation non-local nécessite de nouvelles approximations au niveau de la fonctionnelle d'échange corrélation de courte portée. L'autre objectif de cette thèse consiste à caractériser de petits clusters  $M(H_2O)_n$  qui pourraient constituer les premières étapes de formation d'un clathrate hydraté autour d'une petite molécule organique M. Autrement dit, nous nous intéressons au mécanisme de la formation de la première couche d'hydratation d'un soluté organique. Pour cela, nous nous intéressons à des assemblages cubiques, hexagonaux et pentagonaux, bidimensionnels et tridimensionnels pour une valeur croissante de n, jusqu'à la formation d'une cage autour de la molécule M. Le soluté choisi dans cette thèse est la  $\beta$ -propiolactone (BPL). La raison principale de ce choix est qu'il existe des données expérimentales de haute précision [34].

Le manuscrit est divisé en trois parties :

- la première partie est consacrée aux rappels théoriques nécessaires à la compréhension de la thèse, elle comprend d'une part une introduction aux méthodes de structure électronique, un état de l'art des méthodes doubles hybrides et d'autre part une introduction aux méthodes interprétatives AIM et MEP.
- La deuxième partie est dédiée au développement méthodologique de méthodes doubles hybrides à séparation de portée incluant une fraction d'échange exact et de corrélation non-locale dans la région de courte portée.
- La troisième et dernière partie consiste en l'étude de la micro-hydratation de la  $\beta$ -propiolactone par les méthodes interprétatives AIM et MEP.

Première partie

**Rappels théoriques**

# Chapitre 1

## Les méthodes de fonction d'onde

## 1.1 Hamiltonien électronique non relativiste dans l'approximation de Born-Oppenheimer

Ce chapitre présente de manière brève les notions élémentaires de la chimie quantique et les notations utilisées dans cette thèse. Les systèmes étudiés dans cette thèse sont constitués d'un nombre  $N$  d'électrons non relativiste et d'une collection de noyaux caractérisés par leurs nombres atomiques  $Z_A$  et leurs masses  $m_A$ . Les électrons sont caractérisés par leur masse  $m_e = 9.11 \times 10^{-31}$  kg et par leur charge  $q_e = 1.60 \times 10^{-19}$  C. Les unités du système international ne sont pas appropriées pour étudier des systèmes à l'échelle moléculaire, par conséquent, on utilisera dans cette section les unités atomiques qui sont définies de la manière suivante :

$$\hbar = m_e = e^2/(4\pi\epsilon_0) = 1. \quad (1.1)$$

En ayant recours aux unités atomiques, l'hamiltonien électronique  $\hat{H}_{el}$  non-relativiste s'écrit :

$$\hat{H}_{el} = \hat{T} + \hat{V}_{Ne} + \hat{W}_{ee}. \quad (1.2)$$

L'opérateur  $\hat{T}$  réfère à l'énergie cinétique des électrons :  $\hat{T} = -\frac{1}{2} \sum_i^N \nabla_i^2$ . Le terme  $\nabla_i^2$  correspond au laplacien par rapport à la coordonnée électronique  $\mathbf{r}_i$ . Le potentiel  $\hat{V}_{Ne} = -\sum_i^N \sum_A^M \frac{Z_A}{r_{iA}}$  comprend les interactions attractives entre  $N$  électrons et  $M$  noyaux séparé par la distance  $r_{iA}$ . Enfin, le potentiel  $\hat{W}_{ee} = \sum_i^N \sum_{j>i}^N \frac{1}{r_{ij}}$  traduit la répulsion inter-électronique, où  $r_{ij}$  est la distance inter-électronique entre l'électron  $i$  et l'électron  $j$ .

L'approximation de Born-Oppenheimer conduit à un découplage des mouvements électroniques et nucléaires. Ceci a pour conséquence que l'énergie électronique  $E_{el}$  (valeur propre de l'opérateur  $\hat{H}_{el}$ ) dépend paramétriquement des coordonnées des noyaux. L'énergie totale d'un système quelconque est la somme de l'énergie électronique  $E_{el}$  et la répulsion nucléaire qui est un terme constant dans l'approximation de Born-Oppenheimer :

$$E_{total} = E_{el} + \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}. \quad (1.3)$$

Dans l'équation (1.3) le terme  $R_{AB}$  réfère à la distance entre deux noyaux. Jusqu'à présent, les opérateurs  $\hat{T}$ ,  $\hat{V}_{Ne}$  et  $\hat{W}_{ee}$  ont été exprimés dans le formalisme de première quantification. Le formalisme de seconde quantification permet une étude plus pratique du problème électronique. Ce formalisme s'appuie sur les opérateurs champs de création  $\hat{\Psi}_\sigma^\dagger(\mathbf{r})$  et d'annihilation  $\hat{\Psi}_\sigma(\mathbf{r})$ . Ces opérateurs font intervenir les coordonnées spatiales et

le spin de l'électron notées respectivement par  $\mathbf{r}$  et  $\sigma$ . Ces opérateurs satisfont les règles d'anti-commutation fermionique. Les opérateurs  $\hat{\Psi}_\sigma^\dagger(\mathbf{r})$  et  $\hat{\Psi}_\sigma(\mathbf{r})$  permettent de définir l'opérateur densité  $\hat{n}(\mathbf{r})$ , l'opérateur matrice densité à une particule  $\hat{n}_1(\mathbf{r}, \mathbf{r}')$  et l'opérateur matrice densité diagonal à deux particules  $\hat{n}_2(\mathbf{r}_1, \mathbf{r}_2)$  :

$$\hat{n}(\mathbf{r}) = \sum_{\sigma} \hat{\Psi}_\sigma^\dagger(\mathbf{r}) \hat{\Psi}_\sigma(\mathbf{r}), \quad (1.4)$$

$$\hat{n}_1(\mathbf{r}, \mathbf{r}') = \sum_{\sigma} \hat{\Psi}_\sigma^\dagger(\mathbf{r}') \hat{\Psi}_\sigma(\mathbf{r}), \quad (1.5)$$

$$\hat{n}_2(\mathbf{r}_1, \mathbf{r}_2) = \sum_{\sigma_1} \sum_{\sigma_2} \hat{\Psi}_{\sigma_1}^\dagger(\mathbf{r}_1) \hat{\Psi}_{\sigma_2}^\dagger(\mathbf{r}_2) \hat{\Psi}_{\sigma_2}(\mathbf{r}_2) \hat{\Psi}_{\sigma_1}(\mathbf{r}_1). \quad (1.6)$$

Nous pouvons maintenant réécrire le potentiel d'attraction noyaux-électrons  $\hat{V}_{Ne}$  et le potentiel de répulsion électrons-électrons  $\hat{W}_{ee}$  de la manière suivante :

$$\hat{V}_{Ne} = \int v_{ne}(\mathbf{r}) \hat{n}(\mathbf{r}) d\mathbf{r}, \quad (1.7)$$

avec :

$$v_{ne}(\mathbf{r}) = - \sum_A^M \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|} \quad (1.8)$$

$$\hat{W}_{ee} = \frac{1}{2} \iint \hat{n}_2(\mathbf{r}_1, \mathbf{r}_2) w_{ee}(r_{12}) d\mathbf{r}_1 d\mathbf{r}_2, \quad (1.9)$$

avec :

$$w_{ee} = 1/r_{12} \quad (1.10)$$

On peut noter que l'hamiltonien électronique défini dans l'équation (1.2) dépend seulement des coordonnées spatiales électroniques. On peut néanmoins introduire une dépendance vis-a-vis du spin électronique en tenant compte du fait que la fonction d'onde d'un système doit être antisymétrique par rapport à l'échange de coordonnées entre deux électrons.

## 1.2 Spin-orbitales et déterminant de Slater

Une orbitale spatiale  $\varphi_i(\mathbf{r})$  est une fonction d'onde mono-électronique décrivant la distribution spatiale d'un électron de telle sorte que la probabilité de trouver un électron dans un élément de volume  $d\mathbf{r}$  vaut  $|\varphi_i(\mathbf{r})|^2 d\mathbf{r}$ . La probabilité de trouver l'électron "i" dans tout l'espace étant égale à un, nous avons donc  $\int |\varphi_i(\mathbf{r})|^2 d\mathbf{r} = 1$ . Deux orbitales spatiales différentes sont orthogonales entre elles si la condition suivante est vérifiée :  $\int \varphi_i^*(\mathbf{r})\varphi_j(\mathbf{r})d\mathbf{r} = 0$ . Un ensemble d'orbitales qui possèdent simultanément ces deux propriétés est dit orthonormé et selon la notation de Dirac, un ensemble d'orbitales orthonormées vérifie :

$$\langle \varphi_i | \varphi_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j. \end{cases} \quad (1.11)$$

Les électrons sont caractérisés par leurs coordonnées spatiales et par leur coordonnée de spin  $\sigma$ . Le spin électronique est caractérisé par deux valeurs possibles  $+1/2$  ou  $1/2$  qui correspondent respectivement aux deux fonctions de spin  $\alpha(\sigma)$  ou  $\beta(\sigma)$ . Chaque électron dépend de ses coordonnées spatiales et du spin que l'on regroupera sous une seule variable  $\mathbf{x} = (\mathbf{r}, \sigma)$ . En ayant recourt à cette variable, on peut définir une spin-orbitale  $\varphi(\mathbf{x})$  qui est une fonction d'onde mono-électronique incluant le spin de l'électron. Si l'on omet l'interaction spin-orbite, les coordonnées d'espace et de spin deviennent indépendantes et par conséquent, une spin-orbitale pourra s'écrire comme étant le produit d'une orbital spatiale et d'une fonction de spin ( $\alpha(\sigma)$  ou  $\beta(\sigma)$ ) :

$$\varphi_i^\alpha(\mathbf{x}) = \varphi_i(\mathbf{r})\alpha(\sigma), \quad (1.12)$$

$$\varphi_i^\beta(\mathbf{x}) = \varphi_i(\mathbf{r})\beta(\sigma). \quad (1.13)$$

La difficulté majeure dans la résolution de l'équation de Schrödinger réside dans l'interaction entre électrons, terme décrit par l'opérateur  $\widehat{W}_{ee}$  dans l'équation (1.2). On pourrait en première approximation négliger ces interactions de telle sorte que les électrons se déplacent indépendamment dans le potentiel nucléaire. En ayant recourt à cette approximation, l'hamiltonien électronique peut être écrit comme une somme d'hamiltoniens

mono-électroniques :

$$\hat{H} = \sum_i^N \hat{h}_i, \quad (1.14)$$

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_A^M \frac{Z_A}{r_{iA}}, \quad (1.15)$$

et la fonction d'onde s'écrit comme un produit de spin-orbitales :

$$\Psi_{\text{HP}}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \varphi_1(\mathbf{x}_1)\varphi_2(\mathbf{x}_2)\dots\varphi_N(\mathbf{x}_N). \quad (1.16)$$

Ce type de fonction d'onde dénommée produit de Hartree, ne prend pas en compte l'indiscernabilité des électrons. En effet, si les coordonnées de deux électrons  $i$  et  $j$  sont inter-changées, les propriétés du système étudié doivent restées identiques. Pour qu'une fonction d'onde électronique soit acceptable, il faut imposer que celle-ci soit une fonction propre de l'opérateur permutation 'P<sub>ij</sub>' (la valeur propre étant égale à -1). Il s'agit du principe d'anti-symétrie qui stipule que la fonction d'onde doit être anti-symétrique avec la permutation des coordonnées de deux électrons. Le produit de Hartree ne satisfait pas ce principe mais nous permet d'obtenir une fonction d'onde acceptable à partir de celui-ci. S'il on considère le cas de deux électrons, le produit de Hartree s'écrit  $\Psi_{\text{HP}}(\mathbf{x}_1, \mathbf{x}_2) = \varphi_1(\mathbf{x}_1)\varphi_2(\mathbf{x}_2)$ . Si maintenant nous inter-changeons l'électron 1 avec l'électron 2, nous obtenons  $\Psi_{\text{HP}}(\mathbf{x}_2, \mathbf{x}_1) = \varphi_1(\mathbf{x}_2)\varphi_2(\mathbf{x}_1)$ . On peut construire une fonction d'onde approché qui respecte le principe d'anti-symétrie  $\Phi(\mathbf{x}_1, \mathbf{x}_2) = -\Phi(\mathbf{x}_2, \mathbf{x}_1)$  comme étant la somme des deux produits de Hartree :

$$\Phi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}}[\varphi_1(\mathbf{x}_1)\varphi_2(\mathbf{x}_2) - \varphi_1(\mathbf{x}_2)\varphi_2(\mathbf{x}_1)] = \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_1(\mathbf{x}_1) & \varphi_2(\mathbf{x}_1) \\ \varphi_1(\mathbf{x}_2) & \varphi_2(\mathbf{x}_2) \end{vmatrix}. \quad (1.17)$$

Pour un système à  $N$  électron on a :

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(\mathbf{x}_1) & \varphi_2(\mathbf{x}_1) & \dots & \varphi_N(\mathbf{x}_1) \\ \varphi_1(\mathbf{x}_2) & \varphi_2(\mathbf{x}_2) & \dots & \varphi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(\mathbf{x}_N) & \varphi_2(\mathbf{x}_N) & \dots & \varphi_N(\mathbf{x}_N) \end{vmatrix}. \quad (1.18)$$

La forme de la fonction d'onde employée ci-dessus (dans les équations (1.17)et(1.18)) est appelée déterminant de Slater. En écrivant la fonction d'onde sous cette forme, on impose automatiquement le fait que la fonction d'onde soit anti-symétrique avec la per-

mutation des coordonnées de deux électrons.

### 1.3 Le théorème variationnel

La première étape dans l'étude d'un système quelconque est l'énergie de son état fondamental, soit l'état de plus basse énergie. Il est possible d'approcher l'énergie de l'état fondamental  $E_0$  en appliquant le théorème variationnel. Ainsi pour n'importe quelle fonction d'onde normée  $\Psi_{\text{essai}}$ , nous avons l'inégalité suivante :

$$E[\Psi_{\text{essai}}] = \langle \Psi_{\text{essai}} | \hat{H} | \Psi_{\text{essai}} \rangle \geq \langle \Psi_0 | \hat{H} | \Psi_0 \rangle = E_0. \quad (1.19)$$

Le théorème variationnel fournit donc une stratégie pour déterminer l'énergie et la fonction d'onde de l'état fondamental en minimisant la fonctionnelle  $E[\Psi_{\text{essai}}]$ . Cette minimisation peut se faire en introduisant un multiplicateur de Lagrange  $E$  :

$$\delta[\langle \Psi | \hat{H} | \Psi \rangle - E \langle \Psi | \Psi \rangle] = 0. \quad (1.20)$$

La valeur optimale du multiplicateur de Lagrange correspond à l'énergie de l'état fondamental associée à la fonction d'onde  $|\Psi\rangle$ . En pratique, il est impossible d'effectuer la minimisation sur toutes les fonctions d'onde possibles, seul un sous-ensemble est considéré. Par conséquent, l'énergie approchée de l'état fondamental obtenue en ayant recourt à cette minimisation sera toujours supérieure à l'énergie exacte de l'état fondamental.

### 1.4 L'approximation Hartree-Fock

La méthode Hartree-Fock s'appuie sur le théorème variationnel. Dans cette méthode, le sous-ensemble considéré est constitué uniquement de fonctions d'onde restreintes à un seul déterminant de Slater.

#### 1.4.1 Observables calculées avec un déterminant de Slater

L'énergie électronique  $E^{\text{HF}}$  obtenue par un déterminant de Slater est la somme des éléments de matrice suivants :

$$E^{\text{HF}} = \langle \Phi | \hat{H} | \Phi \rangle = \sum_i^N h_{ii} + \sum_i^N \sum_{j>i}^N (J_{ij} - K_{ij}). \quad (1.21)$$

L'expression mathématique et les interprétations physiques des éléments de matrices

$h_{ii}$  et  $(J_{ij} - K_{ij})$  vont être données ci-dessous :

- L'élément de matrice  $h_{ii}$  correspond à l'énergie cinétique d'un électron dans la spin-orbitale  $\varphi_i$  et l'énergie potentielle provenant de l'interaction électrostatique entre l'électron dans la spin-orbitale  $\varphi_i$  et les noyaux. Cet élément de matrice est donné par l'intégrale mono-électronique suivante :

$$h_{ii} = \langle i | \hat{h}_1 | i \rangle = \int \varphi_i^*(\mathbf{x}_1) \left( -\frac{1}{2} \nabla_1^2 - \sum_A^M \frac{Z_A}{r_{1A}} \right) \varphi_i(\mathbf{x}_1) d\mathbf{x}_1. \quad (1.22)$$

- L'élément de matrice  $J_{ij}$  communément appelée intégrale de Coulomb, représente l'énergie potentielle électrostatique répulsive entre deux électrons et est donnée par l'intégrale bi-électronique suivante :

$$J_{ij} = \langle ij | ij \rangle = \int \varphi_i^*(\mathbf{x}_1) \varphi_j^*(\mathbf{x}_2) \frac{1}{r_{12}} \varphi_i(\mathbf{x}_1) \varphi_j(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2. \quad (1.23)$$

- L'élément de matrice  $K_{ij}$  communément appelée intégrale d'échange, est donnée par l'intégrale suivante :

$$K_{ij} = \langle ij | ji \rangle = \int \varphi_i^*(\mathbf{x}_1) \varphi_j^*(\mathbf{x}_2) \frac{1}{r_{12}} \varphi_j(\mathbf{x}_1) \varphi_i(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2, \quad (1.24)$$

Cette intégrale est différente de zéro uniquement si les spin-orbitales "i" et "j" possèdent la même fonction de spin.

## 1.4.2 Les équations Hartree-Fock

Les orbitales qui définissent le déterminant de Slater doivent être déterminées. Nous pouvons appliquer le théorème variationnel et supposer que les orbitales conduisant à la valeur la plus basse en énergie sont les meilleures. Une variation infinitésimale des orbitales  $\varphi_i \rightarrow \varphi_i + \delta\varphi_i$  résulte en une variation de la fonction d'onde totale et en une variation de l'énergie. Pour appliquer le théorème variationnel à l'équation (1.21), nous introduisons les opérateurs de Coulomb et d'échange par leur action sur une orbitale mono-électronique arbitraire  $\varphi_j$  :

$$\hat{J}_i \varphi_j(\mathbf{x}_2) = \int \left( \varphi_i^*(\mathbf{x}_1) \frac{1}{r_{12}} \varphi_i(\mathbf{x}_1) d\mathbf{x}_1 \right) \varphi_j(\mathbf{x}_2), \quad (1.25)$$

$$\hat{K}_i \varphi_j(\mathbf{x}_2) = \int \left( \varphi_i^*(\mathbf{x}_1) \frac{1}{r_{12}} \varphi_j(\mathbf{x}_1) d\mathbf{x}_1 \right) \varphi_i(\mathbf{x}_2). \quad (1.26)$$

L'opérateur de Coulomb est dit local puisque son action ne dépend que de l'orbitale  $\varphi_j$  en  $\mathbf{x}_2$ . En revanche, l'opérateur d'échange est dit non-local puisque son action est défini par l' orbitale  $\varphi_j$  en tout point de  $\mathbf{x}_1$ . La réécriture de l'équation (1.21) avec les opérateurs de Coulomb et d'échange peut se faire en effectuant une double sommation :

$$E^{\text{HF}} = \sum_i^N \langle \varphi_i | \hat{h}_1 | \varphi_i \rangle + \frac{1}{2} \sum_i^N \sum_j^N \langle \varphi_i | J_j | \varphi_i \rangle - \langle \varphi_i | K_j | \varphi_i \rangle. \quad (1.27)$$

On note que le dernier terme s'annule si les fonctions de spin des spin-orbitales  $\varphi_i$  et  $\varphi_j$  diffèrent. De plus, dans le cas où  $i = j$ , le terme bi-électronique s'annule ; il n'existe pas de self-interaction. La minimisation de l'énergie  $E^{\text{HF}}$  dans l'équation (1.27) doit s'effectuer de telle sorte que les spin-orbitales doivent être orthonormées. Pour cela nous introduisons le lagrangien suivant :

$$L^{\text{HF}} = E^{\text{HF}} - \sum_i^N \sum_j^N \varepsilon_{ij} (\langle \varphi_i | \varphi_j \rangle - \delta_{ij}), \quad (1.28)$$

où  $\varepsilon_{ij}$  sont les multiplicateurs de Lagrange. Minimiser l'énergie revient à résoudre l'équation  $\delta L^{\text{HF}} = 0$ . Pour une variation infinitésimale des spin-orbitales ( $\varphi_i \rightarrow \varphi_i + \delta\varphi_i$ ),  $\delta L^{\text{HF}}$  se calcule de la façon suivante :

$$\begin{aligned} \delta L^{\text{HF}} &= \sum_i^N \langle \delta\varphi_i | \hat{h}_1 | \varphi_i \rangle + \sum_i^N \sum_j^N \langle \delta\varphi_i | \hat{J}_j | \varphi_i \rangle - \langle \delta\varphi_i | \hat{K}_j | \varphi_i \rangle - \sum_i^N \sum_j^N \varepsilon_{ij} (\langle \delta\varphi_i | \varphi_j \rangle - \delta_{ij}) \\ &= \sum_i^N \langle \delta\varphi_i | \left[ \left( \hat{h} + \sum_j^N \hat{J}_j - \hat{K}_j \right) | \varphi_i \rangle - \sum_j^N \varepsilon_{ij} | \varphi_j \rangle \right]. \end{aligned} \quad (1.29)$$

Résoudre  $\delta L^{\text{HF}} = 0$  est équivalent à :

$$\left( \hat{h} + \sum_j^N \hat{J}_j - \hat{K}_j \right) | \varphi_i \rangle = \sum_j^N \varepsilon_{ij} | \varphi_j \rangle. \quad (1.31)$$

A partir de l'équation précédente, on peut définir l'opérateur de Fock :

$$\hat{F} = \hat{h}_1 + \sum_j^N \hat{J}_j - \hat{K}_j \quad (1.32)$$

Il s'agit d'un opérateur mono-électronique hermitique qui satisfait donc l'équation (1.31). Cette équation peut être simplifiée en ayant recourt à une transformation unitaire qui ne modifie pas la fonction d'onde du système et qui diagonalise la matrice des multiplicateurs de Lagrange (e.g  $\varepsilon_{ij} = 0$  et  $\varepsilon_{ii} = \varepsilon_i$ ). Ce qui nous mène à réécrire l'équation (1.31) de la manière suivante :

$$\hat{F}|\varphi_i\rangle = \varepsilon_i|\varphi_i\rangle \quad (1.33)$$

Les spin-orbitales obtenues par ce procédé sont dénommées canoniques, la valeur propre  $\varepsilon_i$  correspond à l'énergie de la spin-orbitale  $|\varphi_i\rangle$ . L'énergie associée à une fonction d'onde mono-déterminantale est liée à l'énergie des spin-orbitales. La somme des énergies des spin-orbitales occupées ne mène pas à l'énergie  $E^{\text{HF}}$ . En effet, si on somme les énergies des spin-orbitales occupées, les interactions coulombiennes et d'échange vont être comptées deux fois. L'énergie  $E^{\text{HF}}$  vaut donc :

$$E^{\text{HF}} = \sum_i^N \varepsilon_i - \frac{1}{2} \sum_i^N \sum_j^N J_{ij} - K_{ij}. \quad (1.34)$$

Dans l'approximation Hartree-Fock, on considère que le "mouvement" des électrons est indépendant et peut être réduit à la "dynamique" d'une particule évoluant dans le champ des noyaux et un champs moyen généré par les autres électrons.

Le formalisme que nous avons appliqué jusqu'ici est général et peut s'appliquer pour un nombre pair ou impair d'électrons. Dans ce cas de figure, le formalisme utilisé est appelée "Unrestricted Hartree-Fock"(UHF).

Pour un nombre pair d'électrons, on peut contraindre une orbitale moléculaire à avoir la même fonction spatiale, la fonction de spin pouvant être de type  $\alpha(\sigma)$  ou  $\beta(\sigma)$ . Dans ce cas nous sommes dans une description restreinte du système, on est donc dans un formalisme de type "Restricted Hartree-Fock"(RHF) qui ne peut être employé uniquement si le nombre d'électron présent dans le système est paire.

### 1.4.3 Les formalismes "restricted" et "unrestricted"

Puisque l'hamiltonien d'un système poly-électronique commute avec les opérateurs de spin totaux  $\hat{S}^2$  et  $\hat{S}_z$ , une fonction propre de l'hamiltonien doit être aussi une fonction

propre des opérateurs de spin. En d'autres termes la fonction d'onde exacte d'un système poly-électronique, qui est une fonction propre de l'hamiltonien, doit décrire un état pur de spin et doit obéir aux relations suivantes :

$$\hat{S}^2|\Psi\rangle = S(S + 1)|\Psi\rangle, \quad (1.35)$$

$$\hat{S}_z|\Psi\rangle = M_s|\Psi\rangle. \quad (1.36)$$

Un déterminant de Slater construit à partir de  $N_\alpha$  spin-orbital et de  $N_\beta$  spin-orbitales sera toujours une fonction propre de l'opérateur  $S_z$ , la valeur propre étant égale à  $M_s = 1/2(N_\alpha - N_\beta)$  : l'équation (1.36) est systématiquement respectée même si l'on a recourt à un seul déterminant de Slater pour approcher notre fonction d'onde poly-électronique. Cependant, ce même déterminant de Slater n'est pas forcément une fonction propre de l'opérateur  $\hat{S}^2$  et l'équation (1.35) peut ne pas être respectée. Dans le cas général, où les électrons possèdent des orbitales spatiales différentes, on aura recourt au formalisme "unrestricted Hartree-Fock" (UHF). Cela nous amène à introduire les équations (10) et (11) dans l'expression de l'opérateur de Fock et celle de l'énergie :

$$\hat{F}^\alpha \varphi_i^\alpha(\mathbf{r}) = \varepsilon_i^\alpha \varphi_i^\alpha(\mathbf{r}). \quad (1.37)$$

L'opérateur de Fock  $\hat{F}^\alpha$  étant égal à :

$$\hat{F}^\alpha = \hat{h} + \sum_j^N \hat{J}_j - \sum_j^{N_\alpha} \hat{K}_j. \quad (1.38)$$

En ce qui concerne l'opérateur de Fock, il faut distinguer le cas où il opère sur une spin-orbital dont la fonction de spin est  $\alpha(\sigma)$  ou sur une spin-orbitale de type  $\beta(\sigma)$ . En effet, l'opérateur  $\hat{K}_j^\alpha$  n'agit que sur les électrons de spin  $\alpha$ . Des équations similaires existent pour les électrons de spin  $\beta$ . Une situation plus simple peut avoir lieu lorsque l'état électronique considéré est un singulet ( $S = M_s = 0$ ), ce qui nécessite nécessairement un nombre pair d'électrons. Dans ce cas précis, on restreint chaque paire électronique à avoir la même orbitale spatiale et nous avons recourt au formalisme "Restricted Hartree-Fock"(RHF) :

$$\varphi_i^\alpha = \varphi_i(\mathbf{r})\alpha(\sigma), \quad (1.39)$$

$$\varphi_i^\beta = \varphi_i(\mathbf{r})\beta(\sigma) \quad (1.40)$$

L'opérateur de Fock dans le formalisme RHF se réécrit :

$$\hat{F} = \hat{h} + \sum_j^{N/2} (2\hat{J}_j - \hat{K}_j^\alpha). \quad (1.41)$$

L'intérêt du formalisme RHF est que le déterminant de Slater représentant la fonction d'onde est une fonction propre des opérateurs  $\hat{S}^2$  et  $\hat{S}_z$  des équations (1.35) et (1.36). Il en ressort que pour un état singulet, les propriétés relatives au spin sont correctes. Le fait que chaque paire d'électrons a la même orbitale spatiale entraîne que l'opérateur de Fock doit être identique. Il existe une méthode pour appliquer les équations (RHF) aux systèmes possédant un nombre impair d'électrons. Cette méthode est communément appelée ROHF (pour *Restricted Open shell Hartree Fock*). Dans ce formalisme, on restreint les électrons appariés à être décrits par la même orbitale spatiale et les électrons non appariés à être décrits par des orbitales spatiales différentes. Un déterminant issu d'une procédure ROHF est fonction propre des opérateurs  $\hat{S}^2$  et  $\hat{S}_z$ .

#### 1.4.4 Les équations de Roothaan

Dans la méthode de Roothaan, les équations intégrales de Hartree-Fock sont converties en un ensemble d'équations algébriques pouvant être résolues par des méthodes matricielles. Dans l'équation (1.33), les inconnues sont les coefficients permettant de développer les orbitales moléculaires sur une base d'orbitales atomiques. Pour déterminer ces coefficients on a recourt à la théorie LCAO (Linear Combination of Atomic Orbitals) c'est à dire qu'on exprime une orbitale moléculaire comme une combinaison linéaire d'orbitales atomiques :

$$\varphi_i(\mathbf{r}) = \sum_{\nu}^K C_{\nu i} \chi_{\nu}(\mathbf{r}). \quad (1.42)$$

Le développement exact devrait être fait sur une base infinie d'orbitales atomiques. Cela n'étant pas possible pour des questions évidentes de praticité on se limite à un nombre  $K$  de fonctions de base. En ayant recourt à la théorie LCAO, la minimisation de l'énergie s'effectue en fonction des coefficients  $C_{\nu i}$  caractérisant les orbitales moléculaires. L'introduction de l'équation (1.42) dans l'équation (1.33) et en multipliant par  $\langle \chi_{\mu} |$  mène à :

$$\sum_{\nu}^K C_{\nu i} \langle \chi_{\mu} | \hat{F} | \chi_{\nu} \rangle = \varepsilon_i \sum_{\nu}^K C_{\nu i} \langle \chi_{\mu} | \chi_{\nu} \rangle. \quad (1.43)$$

Dans l'équation précédente, nous avons fait apparaître la matrice de recouvrement  $\mathbf{S}$  et la matrice de Fock  $\mathbf{F}$  :

$$S_{\mu\nu} = \langle \chi_\mu | \chi_\nu \rangle, \quad (1.44)$$

$$F_{\mu\nu} = \langle \chi_\mu | \hat{F} | \chi_\nu \rangle. \quad (1.45)$$

L'équation dite de Roothaan n'est autre que l'équation (1.43) écrite sous une forme matricielle :  $\mathbf{FC} = \mathbf{SCE}$ . La matrice  $\mathbf{E}$  est une matrice diagonale contenant les énergies des orbitales,  $E_{ij} = \delta_{ij}\varepsilon_i$  et la matrice  $\mathbf{C}$  est la matrice des coefficients LCAO. La détermination des coefficients des orbitales moléculaires s'effectue à l'aide des éléments de la matrice de Fock  $F_{\mu\nu}$ . Dans un formalisme de type RHF, les éléments de la matrice de Fock  $F_{\mu\nu}$  s'écrivent :

$$F_{\mu\nu} = \hat{H}_{\mu\nu}^{\text{core}} + \sum_{\mu\lambda} P_{\mu\lambda} \left[ \langle \mu\lambda | \nu\sigma \rangle - \frac{1}{2} \langle \mu\lambda | \sigma\nu \rangle \right]. \quad (1.46)$$

Dans l'équation précédente, nous avons introduit le terme  $\hat{H}_{\mu\nu}^{\text{core}}$ , il s'agit de la matrice qui implique l'opérateur mono-électronique défini précédemment dans l'équation (20). L'équation précédente fait également intervenir la matrice densité qui dans un formalisme RHF s'écrit :

$$P_{\mu\lambda} = 2 \sum_j^{N/2} C_{\mu j} C_{\nu j}. \quad (1.47)$$

Nous avons également introduit la notation des physiciens pour les intégrales bi-électroniques déjà introduite dans l'équation 1.23 :

$$\langle \mu\lambda | \nu\sigma \rangle = \int \chi_\mu^*(\mathbf{r}_1) \chi_\lambda^*(\mathbf{r}_2) \frac{1}{r_{12}} \chi_\nu(\mathbf{r}_1) \chi_\sigma(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (1.48)$$

La méthode du champ auto-cohérent est une méthode permettant de résoudre les équations de Roothaan qui est une équation matricielle aux valeurs propres. Il peut être pratique de réécrire cette équation sous une forme plus "conventionnelle", c'est à dire sans la matrice  $\mathbf{S}$ . Cela peut se faire en exprimant les orbitales dans une base orthonormale. Nous introduisons ensuite la matrice unitaire  $\mathbf{U}$  de tel sorte que les orbitales atomiques soient orthonormales entre elles via la transformation suivante :  $\mathbf{U}^\dagger \mathbf{S} \mathbf{U} = \mathbf{I}$ , la matrice  $\mathbf{U}^\dagger$  étant l'adjointe à  $\mathbf{U}$ . Dans la méthode du champ auto-cohérent, la matrice  $\mathbf{U}$  peut être égale à  $\mathbf{S}^{1/2}$ . Nous pouvons désormais réécrire l'équation de Roothaan de la manière

suivante :

$$(\mathbf{U}^\dagger \mathbf{F} \mathbf{U})(\mathbf{U}^\dagger \mathbf{C}) = (\mathbf{U}^\dagger \mathbf{S} \mathbf{U})(\mathbf{U}^\dagger \mathbf{C}) \mathbf{E} \Leftrightarrow \mathbf{F}' \mathbf{C}' = \mathbf{C}' \mathbf{E}, \quad (1.49)$$

Avec  $\mathbf{F}' = \mathbf{U}^\dagger \mathbf{F} \mathbf{U}$  et  $\mathbf{C}' = \mathbf{U}^\dagger \mathbf{C}$ . La résolution de l'équation (1.49) s'effectue en diagonalisant la matrice  $\mathbf{F}'$ . Cependant, cette matrice dépend des éléments de la matrice  $\mathbf{C}$ . On peut obtenir une première proposition ("initial guess" en anglais) de la matrice  $\mathbf{C}$  afin de démarrer le processus. Ainsi, il nous est possible d'approcher l'opérateur de Fock. La diagonalisation de cet opérateur approché nous fournira un nouveau ensemble de spin-orbitales qui nous permettra de construire un nouvel opérateur de Fock. Ce procédé peut être réitéré jusqu'à ce que la matrice densité ou l'énergie électronique converge. Ce procédé est connu sous la dénomination de champ auto-cohérent (Self-Consistent-Field, SCF).

## 1.5 Au delà de la méthode Hartree-Fock

L'approximation Hartree-Fock consiste à remplacer le terme d'interaction inter-électronique de l'équation de Schrödinger par une interaction de champ moyen (représentée par les opérateurs de Coulomb et d'échange). L'utilisation de la méthode Hartree-Fock avec une base de fonctions suffisamment importante permet de récupérer 99 % de l'énergie électronique. Cependant, les 1% restants jouent un rôle crucial sur les différences énergétiques (réactivité chimique, énergie d'interaction intermoléculaire ...). La différence d'énergie entre l'énergie exacte non-relativiste  $E_0$  et l'énergie Hartree-Fock  $E_{\text{HF}}$  est l'énergie de corrélation électronique :  $E_c = E_0 - E_{\text{HF}}$ . Il est donc important d'aller au-delà de l'approximation Hartree-Fock et de calculer l'énergie de corrélation, ce qui est l'objectif des méthodes post-HF.

### 1.5.1 La méthode d'interaction de Configuration

La méthode des interactions de configurations consiste à optimiser une fonction d'onde  $\Psi_{\text{IC}}$  formée d'une combinaison linéaire de plusieurs déterminants de Slater :

$$|\Psi_{\text{IC}}\rangle = c_0 |\Phi_0\rangle + \sum_a^{\text{occ}} \sum_r^{\text{vir}} c_a^r |\Phi_a^r\rangle + \sum_{a<b}^{\text{occ}} \sum_{r<s}^{\text{vir}} c_{ab}^{rs} |\Phi_{ab}^{rs}\rangle + \dots \quad (1.50)$$

A partir du déterminant Hartree-Fock  $\Phi_0$ , nous pouvons générer plusieurs déterminants correspondants à des états excités, en échangeant une spin-orbitale occupée avec une spin-orbitale virtuelle (non occupée). Dans l'équation (1.50)  $|\Phi_0\rangle$  réfère au déterminant d'Hartree-Fock tandis que les déterminants  $|\Phi_a^r\rangle$  et  $|\Phi_{ab}^{rs}\rangle$  représentent respectivement des

déterminants simplement et doublement excités. Les coefficients  $c_0$ ,  $c_a^r$  et  $c_{ab}^{rs}$  correspondant à la fonction d'onde de l'état fondamental sont obtenus par une minimisation de l'énergie IC tout en imposant la contrainte de normalisation à la fonction d'onde  $|\Psi_{IC}\rangle$ . Cela nous amène à la résolution de l'équation aux valeurs propres :

$$\begin{pmatrix} \langle \Phi_0 | \hat{H} | \Phi_0 \rangle & \langle \Phi_0 | \hat{H} | \Phi_a^r \rangle & \langle \Phi_0 | \hat{H} | \Phi_{ab}^{rs} \rangle & \dots \\ \langle \Phi_{a'}^r | \hat{H} | \Phi_0 \rangle & \langle \Phi_{a'}^r | \hat{H} | \Phi_a^r \rangle & \langle \Phi_{a'}^r | \hat{H} | \Phi_{ab}^{rs} \rangle & \dots \\ \langle \Phi_{a'b'}^{r'} | \hat{H} | \Phi_0 \rangle & \langle \Phi_{a'b'}^{r'} | \hat{H} | \Phi_a^r \rangle & \langle \Phi_{a'b'}^{r'} | \hat{H} | \Phi_{ab}^{rs} \rangle & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} c_0 \\ c_a^r \\ c_{ab}^{rs} \\ \vdots \end{pmatrix} = E \begin{pmatrix} c_0 \\ c_a^r \\ c_{ab}^{rs} \\ \vdots \end{pmatrix}. \quad (1.51)$$

Le vecteur propre associé à la valeur propre la plus basse en énergie correspond à l'état fondamental du système étudié, les autres vecteurs propres décrivant des états excités. Dans le cas où tous les niveaux d'excitations sont inclus, la méthode est dénommée "interaction de configuration complète" (Full configuration-interaction (FCI)). Il en ressort que la fonction d'onde du système est exacte si le développement des orbitales moléculaires s'effectue sur une base infinie d'orbitales atomiques. D'un point de vue pratique, sauf pour quelques systèmes contenant peu d'électrons, il est impossible de prendre en compte toutes les excitations. Généralement, les méthodes utilisées sont dites IC tronquées, car on se limite à un nombre restreint d'excitations. Les méthodes IC tronquées ont l'inconvénient de pas être cohérentes en taille (size consistent) : si l'on considère deux molécules séparées par une distance suffisamment importante pour considérer que l'énergie d'interaction est nulle ; l'énergie du dimère devrait être égale à deux fois l'énergie d'un monomère. Il existe d'autres approches post-Hartree-Fock qui ont la particularité d'être cohérentes en taille. L'approche Møller -Plesset qui à recourt à la théorie des perturbations présente cette particularité.

## 1.5.2 La théorie des perturbations

Cette section fournit une introduction à la théorie des perturbations de Møller-Plesset. Un état de l'art de cette théorie est donnée dans la revue [35]. La méthode Møller-Plesset est une application de la théorie des perturbations de Rayleigh-Schrödinger à la théorie Hartree-Fock. Considérons un hamiltonien pouvant être exprimé de la façon suivante :  $\hat{H}^\varepsilon = \hat{H}^0 + \varepsilon \hat{W}$ . L'opérateur  $\hat{H}^{(0)}$  est l'hamiltonien non-perturbé (l'ordre de perturbation est égal à zéro). L'opérateur  $\hat{W}$  est l'opérateur permettant d'introduire une perturbation. Ces deux opérateurs sont choisis de telle sorte que lorsque le paramètre  $\varepsilon$  vaut un, nous retrouvons l'hamiltonien exact du système.

L'opérateur non-perturbé  $\hat{H}^{(0)}$  est choisi de telle sorte que les états propres et valeurs propres qui lui sont associés sont connus. En partant de l'équation aux valeurs propres  $\hat{H}^{(0)}|\Phi_n\rangle = E_n^{(0)}|\Phi_n\rangle$  connue, où les vecteurs propres sont orthonormaux (e.g.  $\langle\Phi_n|\Phi_m\rangle = \delta_{nm}$ ) il est possible de déterminer les vecteurs propres  $|\Psi_n^\varepsilon\rangle$  et les valeurs propres associées  $E_0^{(\varepsilon)}$  de l'hamiltonien  $\hat{H}^\varepsilon$ . L'énergie  $E_n^{(\varepsilon)}$  et la fonction d'onde de l'état fondamental  $|\Psi_0^\varepsilon\rangle$ , peuvent être écrites en utilisant un développement de Taylor :

$$E_0^{(\varepsilon)} = E_0^{(0)} + \varepsilon E_0^{(1)} + \varepsilon^2 E_0^{(2)} + \dots, \quad (1.52)$$

$$|\Psi_0^\varepsilon\rangle = |\Psi_0^{(0)}\rangle + \varepsilon|\Psi_0^{(1)}\rangle + \varepsilon^2|\Psi_0^{(2)}\rangle + \dots \quad (1.53)$$

On peut noter que la fonction d'onde non-perturbée  $\Psi_0^{(0)}$  est égale à  $\Phi_0$ . Pour des raisons de convenance, nous introduisons la condition de normalisation intermédiaire suivante :  $\langle\Phi_0|\Psi_0^\varepsilon\rangle = 1$  quelle que soit la valeur  $\varepsilon$ . De cette condition de normalisation intermédiaire, il en ressort que la correction de la fonction d'onde à chaque ordre est orthogonal à la fonction d'onde non-perturbée, autrement dit  $\langle\Phi_0|\Psi_0^{(i)}\rangle = 0$  si  $i \geq 1$ . L'insertion des équations précédentes dans l'équation de Schrödinger nous conduit à :

$$\left(\hat{H}^{(0)} + \varepsilon\hat{W}\right) \left(|\Psi_0^{(0)}\rangle + \varepsilon|\Psi_0^{(1)}\rangle + \varepsilon^2|\Psi_0^{(2)}\rangle + \dots\right) \quad (1.54)$$

$$= \left(E_0^{(0)} + \varepsilon E_0^{(1)} + \varepsilon^2 E_0^{(2)} + \dots\right) \left(|\Psi_0^{(0)}\rangle + \varepsilon|\Psi_0^{(1)}\rangle + \varepsilon^2|\Psi_0^{(2)}\rangle + \dots\right). \quad (1.55)$$

De l'équation précédente, on peut extraire de nouvelles équations et les classer selon leurs ordres de perturbations. A l'ordre zéro de la perturbation, nous retrouvons l'équation aux valeurs propres connue :  $\hat{H}^{(0)}|\Psi_0^{(0)}\rangle = E_n^{(0)}|\Phi_n\rangle$ . Au premier ordre de la perturbation, nous obtenons :

$$\hat{H}^{(0)}|\Psi_n^{(1)}\rangle + \hat{W}|\Phi_0\rangle = E_0^{(0)}|\Psi_0^{(1)}\rangle + E_0^{(1)}|\Phi_0\rangle. \quad (1.56)$$

En insérant dans l'équation précédente le bra  $\langle\Phi_0|$  et en ayant recours aux égalités suivantes :  $\langle\Phi_0|\hat{H}^{(0)} = E_0^{(0)}\langle\Phi_0|$  et  $\langle\Phi_0|\Psi_0^{(1)}\rangle = 0$ , nous pouvons exprimer la correction de l'énergie au premier ordre :

$$E_0^{(1)} = \langle\Phi_0|\hat{W}|\Phi_0\rangle. \quad (1.57)$$

En projetant l'équation (1.56) sur les états définis par  $\langle\Phi_n|$  et en ayant recours aux

égalités suivantes :  $\langle \Phi_n | \hat{H}^{(0)} = E_0^{(n)} \langle \Phi_n |$  et  $\langle \Phi_n | \Phi_0 \rangle = 0, n \neq 0$ , nous obtenons :

$$E_n^{(0)} \langle \Phi_n | \Psi_0^{(1)} \rangle + \langle \Phi_n | \hat{W} | \Phi_0 \rangle = E_0^{(0)} \langle \Phi_n | \Psi_0^{(1)} \rangle. \quad (1.58)$$

La réécriture de l'équation précédente mène à :

$$|\Psi_0^{(1)}\rangle = - \sum_{n \neq 0} \frac{\langle \Phi_n | \hat{W} | \Phi_0 \rangle}{E_n^{(0)} - E_0^{(0)}}. \quad (1.59)$$

De même, au second ordre de la perturbation, nous obtenons :

$$\hat{H}^{(0)} |\Psi_0^{(2)}\rangle + \hat{W} |\Psi_0^{(1)}\rangle = E_0^{(0)} |\Psi_0^{(2)}\rangle + E_0^{(1)} |\Psi_0^{(1)}\rangle + E_0^{(2)} |\Phi_0\rangle. \quad (1.60)$$

La correction de l'énergie au second ordre s'obtient en projetant l'équation (1.60) sur l'état représenté par le bra  $\langle \Phi_0 |$  et en ayant recourt aux relations  $\langle \Phi_0 | \hat{H}^0 = E_0^{(0)} \langle \Phi_0 |$  et  $\langle \Phi_0 | \Psi_0^2 \rangle = \langle \Phi_0 | \Psi_0^{(1)} \rangle$  :

$$E_0^{(2)} = \langle \Phi_0 | \hat{W} | \Psi_0^{(1)} \rangle, \quad (1.61)$$

ce qui nous mène à :

$$E_0^{(2)} = - \sum_{n \neq 0} \frac{|\langle \Phi_0 | \hat{W} | \Phi_n \rangle|^2}{E_n^{(0)} - E_0^{(0)}}, \quad (1.62)$$

où l'on a inséré l'équation (1.59) dans l'équation (1.61). On peut également noter que la correction de l'énergie au second ordre  $E_0^{(2)}$  diverge s'il existe un état  $\Phi_n$  ( $n \neq 0$ ) associé à une énergie  $E_n^{(0)}$  égale à  $E_0^{(0)}$ . Autrement dit, si l'hamiltonien à l'ordre zéro est associé à un état fondamental dégénéré. Dans la théorie Møller-Plesset, l'hamiltonien à l'ordre zéro est l'opérateur de Fock (e.g.  $\hat{H}^{(0)} = \hat{F}$ ). L'opérateur  $\hat{W}$  effectuant la perturbation est obtenu en effectuant la différence entre l'interaction de Coulomb et le potentiel Hartree-Fock :

$$\hat{W}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = \frac{1}{2} \sum_i^N \sum_{i \neq j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_i^N V_{\text{HF}}(\mathbf{x}_i) \quad (1.63)$$

$$\text{avec } \hat{V}_{\text{HF}}(\mathbf{x}_i) = \hat{J}_i(\mathbf{x}_i) - \hat{K}_i(\mathbf{x}_i) \quad (1.64)$$

A l'ordre zéro, la fonction d'onde décrivant l'état fondamental est la fonction d'onde mono-déterminantale issue d'un calcul Hartree-Fock ( $\Phi_0 = \Phi^{\text{HF}}$ ). L'énergie à l'ordre zéro

correspond à la somme des énergies des orbitales occupées.

$$E_0^{(0)} = \sum_a^{\text{occ}} \varepsilon_a. \quad (1.65)$$

La correction de l'énergie au premier ordre s'obtient après l'application des règles de Slater-Condon :

$$E_0^{(1)} = \langle \Phi^{\text{HF}} | \hat{W} | \Phi^{\text{HF}} \rangle = -\frac{1}{2} \sum_{a,b}^{\text{occ}} \langle ab || ab \rangle. \quad (1.66)$$

On peut remarquer que la somme des énergies  $E_0^{(0)}$  et  $E_0^{(1)}$  correspond à l'énergie Hartree-Fock. La correction de l'énergie au second ordre qui est appelée dans ce contexte, l'énergie de corrélation MP2, s'écrit :

$$E_0^{(2)} = E_{\text{corr}}^{\text{MP2}} = - \sum_{n \neq 0} \frac{|\langle \Phi^{\text{HF}} | \hat{W} | \Phi_n \rangle|^2}{E_n^{(0)} - E_0^{(0)}}. \quad (1.67)$$

A priori,  $\Phi_n$  peut représenter un déterminant simplement, doublement, triplement, ... excité. Puisque  $\hat{W}$  est un opérateur bi-électronique, les déterminants qui diffèrent par plus de deux excitations par rapport à  $\Phi^{\text{HF}}$  conduisent à des éléments de matrice  $\langle \Phi^{\text{HF}} | \hat{W} | \Phi_n \rangle$  nuls<sup>1</sup>. De plus, il s'avère que les simples excitations conduisent également à des éléments de matrice nul :

$$\langle \Phi^{\text{HF}} | \hat{W} | \Phi_a^r \rangle = \sum_b^{\text{occ}} \langle ab || rb \rangle - \langle a | \hat{V}_{\text{HF}} | r \rangle \quad (1.68)$$

$$= \sum_b^{\text{occ}} \langle ab || rb \rangle - \sum_b^{\text{occ}} \langle ab || rb \rangle = 0. \quad (1.69)$$

Seuls les déterminants doublement excités contribuent à l'énergie de corrélation ( $\langle \Phi^{\text{HF}} | \hat{W} | \Phi_{ab}^{rs} \rangle = \langle ab || rs \rangle \neq 0$ ). De plus, l'énergie à l'ordre zéro correspondant aux déterminants doublement excités  $\Phi_{ab}^{rs}$  est :

$$E_n^{(0)} = E_{ab}^{\text{rs},(0)} = E_0^{(0)} + \varepsilon_r + \varepsilon_s - \varepsilon_a - \varepsilon_b. \quad (1.70)$$

L'équation (1.67) peut désormais se réécrire de la façon suivante :

$$E_c^{\text{MP2}} = - \sum_{a < b}^{\text{occ}} \sum_{r < s}^{\text{vir}} \frac{|\langle ab || rs \rangle|^2}{\varepsilon_r + \varepsilon_s - \varepsilon_a - \varepsilon_b}. \quad (1.71)$$

---

1. Ce résultat s'obtient après l'application des règles de Slater-Condon

Dans les équations précédentes, les indices "a" et "b" représentent des spin-orbitales occupées tandis que "r" et "s" représentent des spin-orbitales virtuelles. On peut remarquer l'énergie de corrélation MP2 est toujours négative. Cette énergie diverge vers  $-\infty$  si l'un des termes du dénominateur de l'équation précédente est nul. Cela survient dans les systèmes pour lequel la théorie Hartree-Fock donne des gaps HOMO-LUMO proches ou égal à zéro. L'énergie totale MP2 est simplement la somme de l'énergie Hartree-Fock et de l'énergie de corrélation obtenue au second ordre de la perturbation. Enfin, il convient de préciser que la théorie des perturbations n'est pas une théorie variationnelle. La méthode MP2 est la méthode post-Hartree-Fock la plus utilisée en raison de son coût de calcul faible (le coût de calcul étant en  $N_{\text{basis}}^5$ ,  $N_{\text{basis}}$  étant proportionnelle à la taille du système).

## Chapitre 2

# La théorie de la fonctionnelle de la densité

Le but de cette section est de rappeler les bases théoriques nécessaires à la compréhension du chapitre 4. Dans les sous-sections suivantes, nous allons aborder les théorèmes fondateurs de la DFT, nous verrons par la suite la formulation de la méthode Kohn-Sham, l'approximation LDA, quelques contraintes exactes de la fonctionnelle universelle, la théorie RPA et les extensions de la méthode Kohn-Sham.

## 2.1 Théorèmes fondateurs d'Hohenberg-Kohn

Si l'on considère un système électronique avec un potentiel local externe arbitraire  $v(\mathbf{r})$ , alors la fonction d'onde<sup>1</sup>  $\Psi$  de l'état fondamentale peut être obtenue en résolvant l'équation de Schrödinger, à partir de cette fonction d'onde il est possible de déterminer la densité électronique associée à l'état fondamental du système. Autrement dit, le potentiel  $v(\mathbf{r})$  détermine la densité électronique  $n(\mathbf{r})$  de l'état fondamental.

En 1964, Hohenberg et Kohn [36] ont montré que cette correspondance entre le potentiel  $v(\mathbf{r})$  et la densité électronique  $n(\mathbf{r})$  de l'état fondamental pouvait être inversée. En d'autres termes, le potentiel  $v(\mathbf{r})$  est une fonctionnelle unique (à une constante additive près) de la densité électronique de l'état fondamental. La preuve de ce théorème est simple et utilise une démonstration par l'absurde. La fonction d'onde de l'état fondamental associée au potentiel  $v(\mathbf{r})$  est elle-même une fonctionnelle de la densité électronique, on la note  $\Psi[n]$ . Cette fonctionnelle a été exploitée par Hohenberg et Kohn pour définir la fonctionnelle universelle c'est-à-dire indépendante du potentiel externe  $v(\mathbf{r})$ ,

$$F[n] = \langle \Psi[n] | \hat{T} + \hat{W}_{ee} | \Psi[n] \rangle. \quad (2.1)$$

Cette fonctionnelle peut être utilisée à son tour pour définir la fonctionnelle associée à l'énergie électronique du système :

$$E[n] = F[n] + \int v_{ne}(\mathbf{r})n(\mathbf{r})d\mathbf{r}. \quad (2.2)$$

Hohenberg et Kohn ont également montré que la fonctionnelle de la densité  $E[n]$  vérifie une propriété variationnelle : l'énergie de l'état fondamental  $E_0$  du système considéré peut être obtenue en minimisant cette fonctionnelle de la densité par rapport à des densités électroniques  $v$ -représentables<sup>2</sup>

$$E_0 = \min_n \left\{ F[n] + \int v_{ne}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \right\} \quad (2.3)$$

---

1. Il peut y en avoir plusieurs si l'état fondamentale est dégénéré.

2. Une densité est  $v$ -représentable avec interaction si elle est la densité de l'état fondamental d'un hamiltonien  $\hat{T} + \hat{W}_{ee} + \int v(\mathbf{r})\hat{n}(\mathbf{r})d(\mathbf{r})$ .

## 2.2 Formulation de Levy-Lieb

En 1979, Levy [37] [38], et ensuite Lieb [39], ont proposé de redéfinir la fonction universelle  $F[n]$  en utilisant une minimisation sous contrainte :

$$F[n] = \min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{W}_{ee} | \Psi \rangle, \quad (2.4)$$

où  $\Psi \rightarrow n$  signifie que la minimisation s'effectue sur des fonctions d'ondes multi-déterminantales anti-symétriques donnant la densité  $n$  selon l'équation suivante :

$$n(\mathbf{r}) = N \int \dots \int |\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 d\sigma d\mathbf{x}_2 \dots, \mathbf{x}_N. \quad (2.5)$$

Pour une densité donnée, la fonction d'onde minimisante est notée  $\Psi[n]$  (cette fonction d'onde n'est pas unique en cas de dégénérescence). Cette définition ne requiert pas l'existence d'un potentiel local associé à la densité : la fonction d'onde  $\Psi$  est défini sur un ensemble de densités électroniques provenant d'une fonction d'onde antisymétrique (l'ensemble est désigné sous le nom de densités  $N$ -représentables). La propriété variationnelle de la fonctionnelle  $E[n]$  introduite dans l'équation (2.2) peut être démontrée en effectuant une recherche sous contrainte. On part du théorème variationnel usuel puis on effectue une minimisation en deux étapes : une minimisation de la fonction d'onde  $\Psi$  à une densité électronique fixée, suivie d'une minimisation sur la densité électronique  $n$  :

$$\begin{aligned} E_0 &= \min_{\Psi} \langle \Psi | \hat{T} + \hat{W}_{ee} + \hat{V}_{ne} | \Psi \rangle \\ &= \min_n \min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{W}_{ee} + \hat{V}_{ne} | \Psi \rangle \\ &= \min_n \left\{ \min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{W}_{ee} | \Psi \rangle + \int v_{ne}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \right\} \\ &= \min_n \left\{ F[n] + \int v_{ne}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \right\}. \end{aligned} \quad (2.6)$$

La densité électronique et l'énergie électronique de l'état fondamentale peuvent être obtenues en principe en effectuant une minimisation sur la densité électronique, une fonction dépendant de trois variables, ce qui est donc une simplification du problème polyelectronique, puisque la fonction d'onde dépend d'un nombre de variables plus conséquent. Cependant, il faut noter que l'expression de la fonctionnelle universelle  $F[n]$  en fonction de la densité n'est pas connue et les approximations directes de celle-ci ne sont pas suffisamment précises. Kohn et Sham ont remédié à ce problème en proposant une partition de la fonctionnelle  $F[n]$  et en ayant recourt à un système fictif d'électrons non-interagissant [22].

## 2.3 L'approche Kohn-Sham

Face à la difficulté d'approcher directement  $F[n]$ , Kohn et Sham (KS) [22] ont proposé de décomposer  $F[n]$  de la façon suivante :

$$F[n] = T_s[n] + E_{\text{Hxc}}[n]. \quad (2.7)$$

Dans l'équation (2.7),  $T_s[n]$  est la fonctionnelle d'énergie cinétique d'un système fictif d'électrons non-interagissant. Cette fonctionnelle s'exprime en fonction des orbitales et peut être calculée exactement :  $T_s[n] = -1/2 \sum_i^{\text{occ}} \langle \varphi_i | \nabla^2 | \varphi_i \rangle$ . La dépendance de la fonctionnelle  $T_s[n]$  vis-à-vis de la densité électronique est implicite, les orbitales étant reliées selon l'expression suivante :  $n(\mathbf{r}) = \sum_i |\varphi_i(\mathbf{r})|^2$ . La méthode KS consiste donc à utiliser l'expression exacte de  $T_s[n]$  dans la formulation de Levy-Lieb, c'est à dire que la minimisation sous-contrainte faite dans l'équation (2.6) se restreint aux fonctions d'onde mono-déterminantale :

$$\begin{aligned} E_0 &= \min_n \left\{ F[n] + \int v_{\text{ne}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \right\} \\ &= \min_n \left\{ \min_{\Phi \rightarrow n} \langle \Phi | \hat{T} | \Phi \rangle + E_{\text{Hxc}}[n] + \int v_{\text{ne}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \right\} \\ &= \min_n \min_{\Phi \rightarrow n} \left\{ \langle \Phi | \hat{T} + \hat{V}_{\text{ne}} | \Phi \rangle + E_{\text{Hxc}}[n\Phi] \right\} \\ &= \min_{\Phi} \left\{ \langle \Phi | \hat{T} + \hat{V}_{\text{ne}} | \Phi \rangle + E_{\text{Hxc}}[n\Phi] \right\}. \end{aligned} \quad (2.8)$$

Les interactions entre électrons sont prises en compte dans la fonctionnelle d'Hartree-échange-corrélation  $E_{\text{Hxc}}[n]$ . Cette fonctionnelle est généralement décomposée de la façon suivante :

$$E_{\text{Hxc}} = E_{\text{H}}[n] + E_{\text{x}}[n] + E_{\text{c}}[n]. \quad (2.9)$$

L'expression de la fonctionnelle d'Hartree en fonction de la densité électronique est connue :

$$E_{\text{H}}[n] = \frac{1}{2} \iint \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2. \quad (2.10)$$

La fonctionnelle d'Hartree traduit la répulsion électrostatique classique en électrons. Les fonctionnelles d'échange et de corrélation sont à approximer. L'expression de la fonctionnelle d'échange est :

$$E_{\text{x}}[n] = \langle \Phi[n] | \hat{W}_{\text{ee}} | \Phi[n] \rangle - E_{\text{H}}[n]. \quad (2.11)$$

$\Phi[\mathbf{n}]$  est la fonction d'onde mono-déterminantale obtenue suite à la minimisation faite dans l'équation (2.8) L'expression de la fonctionnelle de corrélation est :

$$E_c[\mathbf{n}] = \langle \Psi[\mathbf{n}] | \hat{T} + \hat{W}_{ee} | \Psi[\mathbf{n}] \rangle - \langle \Phi[\mathbf{n}] | \hat{T} + \hat{W}_{ee} | \Phi[\mathbf{n}] \rangle. \quad (2.12)$$

La résolution de l'équation (2.8) sous la contrainte d'orthonormalisation des orbitales :  $\langle \varphi_i | \varphi_j \rangle = \delta_{ij}$  nous amène à la résolution d'équations mono-électroniques, ces équations sont connues sous le nom des équations Kohn-Sham :

$$\left[ -\frac{1}{2} \nabla^2 + v^{\text{KS}}(\mathbf{r}) \right] \varphi_i(\mathbf{r}) = \varepsilon_i \varphi_i(\mathbf{r}). \quad (2.13)$$

La valeur propre  $\varepsilon_i$  correspond à l'énergie orbitalaire associée à l'orbitale  $\varphi_i$ .  $v^{\text{KS}}(\mathbf{r})$  représente le potentiel de Kohn-Sham et peut être décomposé ainsi :  $v^{\text{KS}}(\mathbf{r}) = v_{\text{ne}}(\mathbf{r}) + v_{\text{Hxc}}(\mathbf{r})$ . Le potentiel  $v_{\text{Hxc}}(\mathbf{r})$  est la dérivée fonctionnelle de  $E_{\text{Hxc}}[\mathbf{n}]$  par rapport à la densité électronique  $n(\mathbf{r})$  :  $v_{\text{Hxc}}(\mathbf{r}) = \delta E_{\text{Hxc}}[\mathbf{n}] / \delta n(\mathbf{r})$ . L'équation (2.13) est résolue de manière itérative, puisque le potentiel de Kohn-Sham,  $v^{\text{KS}}(\mathbf{r})$ , dépend des orbitales. Lorsque la convergence est atteinte, la fonction d'onde minimisant  $\Phi[\mathbf{n}]$  l'énergie donne la densité électronique du système réel. L'expression exacte de la fonctionnelle d'échange-corrélation étant inconnue, nous allons brièvement discuter des approximations usuelles pour la fonctionnelle d'échange-corrélation.

### 2.3.1 L'approche LDA

Dans l'approche LDA, introduite par Kohn et Sham [22], la fonctionnelle d'échange-corrélation est approximée de la manière suivante :

$$E_{\text{xc}}^{\text{LDA}}[\mathbf{n}] = \int n(\mathbf{r}) \varepsilon_{\text{xc}}^{\text{unif}}(n(\mathbf{r})) d\mathbf{r}, \quad (2.14)$$

Dans l'équation (2.14),  $\varepsilon_{\text{xc}}^{\text{unif}}$  est l'énergie d'échange-corrélation par particule d'un gaz d'électron homogène associé à une densité électronique fixe. L'approximation LDA consiste à assimiler l'énergie d'échange-corrélation par particule en chaque point  $\mathbf{r}$  d'un système dont la densité électronique est inhomogène à celui d'un gaz d'électron homogène ayant la même densité. La fonction  $\varepsilon_{\text{xc}}^{\text{unif}}$  peut être séparé en deux contributions :  $\varepsilon_{\text{xc}}^{\text{unif}}(n) = \varepsilon_{\text{x}}^{\text{unif}}(n) + \varepsilon_{\text{c}}^{\text{unif}}(n)$ . L'énergie d'échange par particule d'un gaz d'électron homogène peut être calculé analytiquement :

$$\varepsilon_{\text{x}}^{\text{unif}}(n) = -\frac{3}{4} \left( \frac{3n(\mathbf{r})}{\pi} \right)^{1/3}. \quad (2.15)$$

La fonctionnelle d'échange est associée aux noms de Dirac [40] et Slater [41]. L'énergie de corrélation par particule d'un gaz électron uniforme ne peut être calculée analytiquement. La quantité  $\epsilon_c^{\text{unif}}(n)$  peut être obtenue numériquement par des calculs de Monte Carlo quantiques [42] et ajustée à une fonction paramétrée de  $n$  satisfaisant les développements connus dans la limite des hautes et des basses densités.

$\epsilon_c^{\text{unif}}(n)$  peut s'exprimer en fonction du rayon de Wigner-Seitz, noté  $r_s$  ( $r_s = 3/(4\pi n)^{1/3}$ ), et dans la limite des hautes densités ( $r_s \rightarrow 0$ ) :

$$\epsilon_c^{\text{unif}}(n) = A \ln r_s + B + C r_s \ln r_s + O(r_s). \quad (2.16)$$

Dans la limite des basses densités ( $r_s \rightarrow \infty$ ),  $\epsilon_c^{\text{unif}}(n)$  s'écrit :

$$\epsilon_c^{\text{unif}}(n) = \frac{a}{r_s} + \frac{b}{r_s} r_s^{3/2} + O\left(\frac{1}{r_s}\right)^2. \quad (2.17)$$

Dans les équations (2.16) et (2.17),  $A$ ,  $B$ ,  $C$ ,  $a$  et  $b$  sont des constantes. Les deux paramétrisations les plus utilisées sont celles de Vosko Wilk, et Nusair (WVN) [43] et de Perdew et Wang (PW92) [44]. Ces paramétrisations incluent également la généralisation aux densités de spin. Pour une revue récente sur les calculs des énergies d'échange et de corrélation du gaz d'électrons homogène, nous renverrons à la référence [45].

### 2.3.2 Tenir compte des inhomogénéités, les approches semi-locales et hybrides

La façon la plus naturelle de tenir compte des inhomogénéités du milieu est de construire des fonctionnelles approximées en fonction de la densité électronique  $n$  et de son gradient  $\nabla n$  ou plutôt de la norme du gradient  $|\nabla n|$ . Puisque les fonctionnelles d'échange-corrélation doivent satisfaire une certaine dimensionnalité (l'unité de  $\epsilon_{xc}(n)$  doit être en Hartree) on fait généralement dépendre  $\epsilon_{xc}(n)$  en fonction du gradient réduit de la densité<sup>3</sup>, noté  $s$  :

$$s = \frac{|\nabla n|}{n^{4/3}}, \quad (2.18)$$

le gradient réduit étant une quantité adimensionnelle. On pourrait s'attendre à une amélioration de l'approximation LDA en mettant au point une fonctionnelle de la forme :

$$E_{xc}[n] = \int n(\mathbf{r}) \epsilon_{xc}(n) [1 + \chi(n) s^2 + \dots] d\mathbf{r}, \quad (2.19)$$

---

3. le gradient réduit de la densité a la particularité d'être invariant aux transformations d'échelle

où  $\kappa(n)$  est une fonction de la densité. Les approximations de ce type sont connues sous la dénomination de "*gradient expansion approximation*" (GEA). On pourrait s'attendre à ce que l'approximation GEA soit plus performante que l'approche LDA. Or, il s'avère que l'approche GEA est moins précise que l'approche LDA pour les atomes et molécules. Une des raisons est que l'addition du terme  $\kappa(n)s^2$  dans l'équation (2.19) est responsable du fait que l'énergie de corrélation peut être positive [46]; ce résultat étant un résultat non physique. On peut remédier à ce problème en introduisant ce qu'on l'on appelle un facteur d'amélioration, notée  $F_{xc}(n, s)$  :

$$E_{xc}^{GGA}(n) = \int n(\mathbf{r}) \epsilon_{xc}^{unif}(n) F_{xc}(n, s) d\mathbf{r}. \quad (2.20)$$

Les fonctionnelles utilisant la forme de l'équation (2.20) sont appelées GGA pour "*generalized gradient approximations*". La forme analytique de la fonction  $F_{xc}$  varie au cas par cas et est construite soit pour satisfaire des contraintes exactes, soit pour reproduire des données expérimentales ou bien les deux. Par exemple, la fonctionnelle d'échange B88 [47] a été choisi pour reproduire le bon comportement asymptotique de l'énergie d'échange par particule :

$$\epsilon_x[n](\mathbf{r}) \xrightarrow{r \rightarrow \infty} -\frac{1}{2r} \quad (2.21)$$

La fonctionnelle B88 contient également un paramètre empirique qui est ajusté à l'énergie d'échange des gaz nobles. La fonctionnelle LYP [48] qui est l'une des rares GGA a ne pas avoir été construite à partir du gaz d'électrons homogène à la particularité de donner une énergie de corrélation nulle pour les systèmes mono-électroniques. La fonctionnelle PBE a été construite entre autre pour tenir compte d'un certain nombre de contraintes exactes. On peut citer par exemple la relation dite de "spin-scaling" [49], la borne de Lieb-Oxford [50].

Il est possible d'aller plus loin en introduisant le laplacien de la densité  $\nabla^2 n(\mathbf{r})$  et/ou la densité d'énergie cinétique  $\tau(\mathbf{r})$  (d'un système où les électrons n'interagissent pas) dans la fonctionnelle d'échange-corrélation. Les approximations de ce type sont communément appelée meta-GGA. Contrairement aux GGA [51], les meta-GGA n'ont pas besoin de choisir parmi des contraintes incompatibles. Une des motivations de l'emploi de  $\tau(\mathbf{r})$  est que cette quantité permet d'annuler l'énergie de corrélation par particule  $\epsilon_c(n)$  dans les régions spatiales où la densité électronique est celle d'un système mono-électronique. Des exemples de fonctionnelles appartenant à la classe des meta-GGA construites en imposant des contraintes exactes sont les fonctionnelles TPSS [52] et SCAN [53].

Les approches GGA et meta-GGA sont qualifiées d'approches semi-locales, car ce type de fonctionnelle d'échange-corrélation dépend de la valeur de la densité au point  $r$ ,

mais font également intervenir les dérivées premières (GGA) mais aussi les dérivés d'ordre supérieur (meta-GGA). Les approches locales et semi-locales ont un défaut majeur que l'on nomme "*self-interaction*". Pour comprendre ce phénomène, on peut prendre l'exemple de l'atome d'hydrogène. Pour ce système mono-électronique, l'interaction entre électrons doit être nulle. Autrement dit, la fonctionnelle d'échange  $E_x[n]$  doit compenser la fonctionnelle d'Hartree  $E_H[n]$  et la fonctionnelle de corrélation  $E_c[n]$  doit être nulle. Hors, l'expression de la fonctionnelle d'échange-corrélation est approximée, par conséquent une erreur de "*self-interaction*" est introduite. Cette erreur est également présente pour les systèmes poly-électroniques. Une façon de réduire cette erreur consiste à introduire de l'échange exact :

$$E_x^{\text{exact}} = -\frac{1}{2} \sum_{\sigma=\alpha,\beta} \sum_{i=1}^{N_\sigma} \sum_{j=1}^{N_\sigma} \iint \frac{\varphi_{i\sigma}^*(\mathbf{r}_1)\varphi_{j\sigma}(\mathbf{r}_1)\varphi_{j\sigma}^*(\mathbf{r}_2)\varphi_{i\sigma}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \quad (2.22)$$

Les fonctionnelles introduisant de l'échange exacte sont qualifiées de méthodes hybrides. Les méthodes hybrides les plus populaires sont les méthodes B3LYP [54] et PBE0 [55, 56]. Une extension naturelle des méthodes hybrides consiste à introduire une fraction d'énergie de corrélation obtenue à un niveau post-HF. Ainsi nous obtenons des méthodes doubles hybrides. Ces méthodes vont être détaillés dans les sous-sections 2.4.1, 2.4.2 et 2.4.3.

## 2.4 Au delà de l'approche Kohn-Sham

### 2.4.1 État de l'art sur les méthodes doubles hybrides

Une extension des méthodes hybrides implique une fraction d'énergie de corrélation obtenue à un niveau post-HF, ce qui revient à employer une fonction d'onde multi-déterminantale. La fonction d'onde post-HF couramment utilisée repose sur la théorie de perturbation Møller-Plesset au deuxième ordre. La combinaison des méthodes DFT et MP2 a été proposée en 2004 par Zhao *et al.* [57] avec l'expression de l'énergie suivante :

$$E^{\text{MC3-DFT}} = c_2[E^{\text{HF/DZ}} + c_1 E_c^{\text{MP2/DZ}}] + (1 - c_2)E_{\text{xc}}^{\text{DFT/TZ}}. \quad (2.23)$$

Dans l'équation (2.23), les acronymes DZ et TZ réfèrent respectivement à une base de qualité double et triple zêta. Cette approche diffère essentiellement des autres méthodes dites doubles hybrides par le fait que l'énergie électronique est obtenue avec deux bases. De plus, l'approche nécessite le calcul de deux jeux d'orbitales, la partie de l'énergie de corrélation exprimée au niveau MP2 ayant recourt au déterminant issu d'un calcul HF.

Le terme "double hybride" a été utilisé pour la première fois dans la référence [58] et a été employé pour la fonctionnelle B2-PLYP [59]. Cette méthode peut être vue comme une simplification de la théorie de perturbation de Görling-Levy (GLPT) [60] [61]. Brièvement, cette théorie de perturbation est issue d'un déterminant Kohn-Sham comme fonction d'onde à l'ordre zéro. Par conséquent, cette théorie est basée sur une densité électronique exacte à l'ordre zéro, ce qui n'est pas le cas de la théorie MP2. De plus, dans la théorie GLPT, les orbitales sont obtenues suite à l'emploi d'un potentiel local d'échange-corrélation dans le processus auto-cohérent. Dans la théorie MP2, elles sont obtenues d'un potentiel d'échange non-local.

De ces différences, il en résulte que le théorème de Brillouin n'est pas vérifié dans le cas de la théorie GLPT. Autrement dit, la contribution des excitations simples à l'énergie de corrélation n'est pas nulle. L'approximation faite dans la méthode B2-PLYP consiste à négliger les excitations simples. L'expression de la fonctionnelle B2-PLYP d'échange-corrélation s'écrit :

$$E_{xc}^{\text{B2-PYLYP}} = a_x E_x^{\text{HF}} + (1 - a_x) E_x^{\text{B}}[n] + (1 - a_c) E_c^{\text{LYP}}[n] + a_c E_c^{\text{MP2}}. \quad (2.24)$$

Dans l'équation (2.24), le terme  $E_x^{\text{B}}[n]$  réfère à la fonctionnelle d'échange de Becke (notée B ou B88) [47] tandis que le terme  $E_c^{\text{LYP}}[n]$  [48] fait référence à fonctionnelle de corrélation LYP. Les trois premiers termes sont calculés de manière auto-cohérente tandis que le dernier est évalué à posteriori. Le succès de B2PLYP a inspiré d'autres travaux de recherche. Dans la plupart des cas, les développements ultérieurs ont été basés sur les mêmes fonctionnelles d'échange-corrélation. Les valeurs des paramètres ont été optimisées pour un certain type d'application. Des exemples de doubles hybrides sont ainsi B2K-PLYP [62], B2T-PLYP [62], B2GP-PLYP [63], ROB2-PLYP [64], B2- $\pi$ PLYP [65]. Les valeurs des paramètres  $a_x$  et  $a_c$  sont reportées dans le tableau 2.1.

Les fonctionnelles B2K-PLYP, B2T-PLYP, B2GP-PLYP et B2- $\pi$ PLYP sont des versions reparamétrisées de la méthode B2-PLYP :

- Les méthodes B2K-PLYP et B2T-PLYP ont été reparamétrisées pour traiter respectivement la thermochimie et la cinétique des réactions.
- La méthode ROB2-PLYP a été reparamétrisée dans un formalisme de type "restricted open-shell" pour le traitement des radicaux.
- La méthode B2- $\pi$ PLYP a été proposée pour le traitement des électrons  $\pi$  (systèmes aromatiques).
- La méthode B2GP-PLYP pour le traitement simultané de la thermochimie et la cinétique des réactions.

Fonctionnelle	$a_x$	$a_c$
B2-PLYP	0,53	0,27
B2K-PLYP	0,72	0,42
B2T-PLYP	0,6	0,31
B2GP-PLYP	0,65	0,36
ROB2-PLYP	0,59	0,28
B2- $\pi$ PLYP	0,602	0,273
mPW2-PLYP	0,55	0,25
mPW2K-PLYP	0,72	0,42
B2-PPW91	0,2	0,1

Tableau 2.1 – Valeurs des paramètres  $a_x$  et  $a_c$

D'autres méthodes doubles hybrides basées sur l'équation (2.24) ont été développées en ayant recourt à d'autres fonctionnelles d'échange-corrélation.

- La fonctionnelle mPW2-PLYP utilise une version modifiée de la fonctionnelle d'échange Perdew-Wang (PW) [66] au lieu de la fonctionnelle d'échange B88.
- La méthode mPW2-PLYP a été reparamétrisée pour une meilleure description de la cinétique des réactions donnant lieu à la fonctionnelle mPW2K-PLYP [62].
- Enfin, la fonctionnelle B2-PPW91 [67] utilise la fonctionnelle de corrélation Perdew-Wang au lieu de LYP et a été développée pour le calcul des propriétés électriques telle que la polarisabilité des métaux de transition de la quatrième ligne du tableau périodique.

Mise à part la fonctionnelle B2-PPW91, les fractions d'échange exacte  $a_x$  et de corrélation  $a_c$  sont compris respectivement entre 53% et 72% et entre 20% et 42%. (voir tableau 2.1). Les fractions  $a_x$  et  $a_c$  peu élevées pour la fonctionnelle B2-PPW91 pourraient s'expliquer par l'emploi d'une énergie de corrélation de type MP2 n'est pas adaptée pour les métaux de transition (systèmes multi-références pour lesquels le gap HOMO-LUMO est faible). Dans tous les cas de figures, une relation d'ordre  $a_c < a_x^2$  est observée. Une rationalisation de cette relation a été apportée par Fromager [68].

Les méthodes doubles hybrides impliquent une dépendance des orbitales virtuelles et devraient en pratique être capables de décrire les interactions de natures dispersives. Cependant, il faut garder à l'esprit que seule une fraction de l'énergie de corrélation MP2 est incluse dans ces méthodes, il en résulte une description incomplète des effets de dispersion<sup>4</sup>. Dans le but de remédier à ce problème des corrections de dispersion reposant sur une approche de type  $C_6$  ont été proposées par Grimme en 2007 donnant lieu aux fonctionnelles B2-PLYP-D et mPW2-PLYP-D [69].

4. Ceci n'est pas le cas pour les méthodes doubles hybrides à séparation de portée.

Il existe deux autres approches pour imposer le bon comportement asymptotique :

- la première approche repose sur une séparation de portée dans le traitement des interaction bi-électronique et va être vue plus en détail dans la sous-section (2.4.3).
- La seconde approche repose sur les fonctionnelles dite de van-der-Waals, plus précisément sur la variante de van-Vorhis-Vydrov (VV10) et également connue sous le nom de DFT-NL [70].

L'équation (2.25) résume le principe des méthodes dite "DFT-NL" :

$$E_{xc}^{\text{DFT-NL}} = E_{xc}^{\text{DFT}} + \int d\mathbf{r}n(\mathbf{r}) \left[ \frac{1}{32} \left[ \frac{3}{b^2} \right]^{3/4} + \frac{1}{2} \int d\mathbf{r}'n(\mathbf{r}')\varphi(\mathbf{r}, \mathbf{r}', b) \right]. \quad (2.25)$$

Dans l'équation (2.25), la double intégrale permet d'évaluer la densité électronique en deux points de l'espace ( $\mathbf{r}$  et  $\mathbf{r}'$ ). L'expression du noyau  $\varphi(\mathbf{r}, \mathbf{r}', b)$  est explicité dans la référence [70]. Le paramètre d'atténuation  $b$  permet d'éviter le double comptage d'interaction lorsque la distance entre deux électrons tend vers zéro, la fonctionnelle  $E_c[n]$  prenant en compte ces interactions. La correction DFT-NL a été appliquée aux méthodes B2-PLYP, revPBE0-DH et TPSS0-DH [71].

En 2003, la méthode "spin-component-scaled" MP2 (SCS-MP2) a été développée par Grimme [72]. Cette méthode module de manière empirique à la fois la contribution à l'énergie de corrélation due aux spins parallèles et la contribution à l'énergie de corrélation due aux spins antiparallèles. La première méthode double hybride qui a eu recours à la variante SCS-MP2 est  $\omega$ B97X-2 [73]. Cette méthode va être discutée plus loin dans cette section car elle utilise une séparation de portée dans la fonctionnelle d'échange. Kozuch *et al.* [74] ont combiné la méthode SCS-MP2 avec une correction de dispersion basée sur une approche  $C_6$  donnant lieu aux approches "Dispersion-corrected, spin-component-scaled, double hybrid" (DSD-DFT). Dans cette famille de méthodes, l'expression de l'énergie d'échange-corrélation s'écrit de la manière suivante :

$$E_{xc}^{\text{DSD-DFT}} = a_x E_x^{\text{HF}} + (1 - a_x) E_x^{\text{DFT}}[n] + c_C E_c^{\text{DFT}}[n] + c_O E_c^{\text{OS-MP2}} + c_S E_c^{\text{SS-MP2}} + E_{\text{disp}}. \quad (2.26)$$

Dans l'équation (2.26), les termes  $E_c^{\text{OS-MP2}}$  et  $E_c^{\text{SS-MP2}}$  représentent respectivement la contribution à l'énergie de corrélation MP2 due aux spins antiparallèles et parallèles, multipliés respectivement par les facteurs  $c_O$  et  $c_S$ . Le facteur  $c_C$  associé à l'énergie de corrélation DFT n'est pas forcément ajusté de telle sorte que la somme des coefficients  $c_C$ ,  $c_O$  et  $c_S$  est égale à 1. Les méthodes DSD-BLYP [75], DSD-PBEP86 [76], DSD-

PBEhB95 [74], DuT et PoT [77] sont des exemples de méthodes doubles hybrides reposant sur l'équation (2.26).

On peut également utiliser la variante "spin-opposite-scaled" MP2 (SOS-MP2). Jung *et al.* [78] ont montré que l'effort computationnel de la méthode SCS-MP2 peut être réduit en négligeant l'énergie de corrélation provenant de la contribution à l'énergie de corrélation due aux spins parallèles. Ce qui revient à dire que le coefficient  $c_S$  dans l'équation (2.26) est nul, autrement dit que la méthode SOS-MP2 inclut les diagrammes direct et ignore les diagrammes d'échange. Dès lors, il est possible d'utiliser une transformation de Laplace [79], permettant de réduire le coût computationnel à un facteur  $N_{\text{basis}}^4$  ( $N_{\text{basis}}$  représentant le nombre de fonctions employées dans la base.).

La méthode B2-OS3LYP [80] est un exemple de méthode employant la variante SOS-MP2. Cette méthode a la particularité d'employer une séparation de portée dans le traitement des interactions bi-électroniques tandis que les méthodes PTPSS et PWPB95 [81] sont des exemples de méthodes reposant sur une séparation linéaire de l'interaction bi-électronique. Les méthodes PTPSS et PWPB95 peuvent se résumer par l'équation (2.27) :

$$E_{xc}^{\text{SOS-DHDFT}} = a_x E_x^{\text{HF}} + (1-a_x) E_x^{\text{DFT}}[n] + (1-c_O) E_c^{\text{DFT}}[n] + c_O E_c^{\text{OS-MP2}} + E_{\text{disp}}. \quad (2.27)$$

Une reformulation plus rigoureuse des fonctionnelles doubles hybrides basées sur une séparation linéaire de l'interaction bi-électronique a été développée par Sharkas *et al* [82]. Brièvement, la reformulation employée dans la référence [82] repose sur une théorie des perturbations dans laquelle seules les excitations doubles contribuent à la corrélation électronique lorsque l'on s'arrête au deuxième ordre de la perturbation. Enfin, seul un paramètre suffit à contrôler le ratio DFT/fonction d'onde dans cette approche. Cette méthode sera vue plus en détail dans la sous-section 2.4.2. L'approche développée par Sharkas *et al*, nommé DS1DH (pour "density-scaled one-parameter double-hybrid") a apporté une justification théorique aux méthodes doubles hybrides et a été le point de départ à des approches plus pragmatiques telle que les méthodes LS1DH [83], PBE0-DH [84], PBE0-2 [85] et LS1DH-TPSS [86] qui peuvent être résumées par l'équation (2.28) :

$$E_{xc}^{\text{LS1DH}} = \lambda E_x^{\text{HF}} + (1-\lambda) E_x^{\text{DFT}}[n] + (1-\lambda^3) E_c^{\text{DFT}}[n] + \lambda^3 E_c^{\text{MP2}}. \quad (2.28)$$

Dans la méthode LS1DH [83] en utilisant la fonctionnelle d'échange-corrélation PBE, la valeur du paramètre  $\lambda$  a été optimisée sur un ensemble de réactions d'atomisation et de barrières énergétiques [87]. Ainsi, une valeur de  $\lambda = 0.75$  a été obtenue. Dans l'approche

PBE0-DH [84], la valeur de  $\lambda$  est fixé à 0.5 tandis que dans l’approche PBE0-2 [85] la valeur de  $\lambda^3$  est fixé à 0.5.

En 2008, Benighaus *et al.* a introduit dans la fonctionnelle B2-PLYP une séparation de portée dans l’expression de l’énergie de corrélation MP2 donnant lieu à la fonctionnelle B2-P3LYP [80]. L’objectif consistait à améliorer la description des complexes de van der Waals :

$$E_{xc}^{\text{B2-P3LYP}} = a_x E_x^{\text{HF}} + (1 - a_x) E_x^{\text{B88}}[n] + (1 - a_c) E_c^{\text{LYP}}[n] + a_c E_c^{\text{MP2}}(\mu, c_\mu). \quad (2.29)$$

Avec :

$$E_c^{\text{MP2}}(\mu, c_\mu) = - \sum_{a < b}^{\text{occ}} \sum_{r < s}^{\text{vir}} \frac{|\langle ab | \hat{g}_\mu(r_{12}) | rs \rangle - \langle ab | \hat{g}_\mu(r_{12}) | sr \rangle|^2}{\varepsilon_r + \varepsilon_s - \varepsilon_a - \varepsilon_b}, \quad (2.30)$$

$$\text{et } \hat{g}_\mu(r_{12}) = \frac{1}{r_{12}} + c_\mu \frac{\text{erf}(\mu r_{12})}{r_{12}}. \quad (2.31)$$

On rappelle que dans l’équation (2.30) les indices a et b représentent des spin-orbitales occupées tandis que r et s représentent des spin-orbitales virtuelles. L’opérateur  $\hat{g}_\mu$  est explicité dans l’équation (2.31), cet opérateur permet de prendre en compte le bon comportement asymptotique en  $1/r^6$  de l’énergie de dispersion. La méthode B2-OS3LYP [80] utilise également cette approche. La méthode LRC-XYG3 [88] utilise une approche similaire à la méthode B2-P3LYP, la différence s’effectue au niveau de l’équation (2.31) où l’opérateur  $\hat{g}_\mu(r_{12})$  se réduit à l’opérateur  $w_{ee}^{\text{lr},\mu}(r_{12}) = \text{erf}(\mu r_{12})/r_{12}$ . Ces approches n’utilisent pas de séparation de portée dans la fonctionnelle d’échange-corrélation et dans la partie d’échange exact.

La méthode  $\omega$ B97X-2 [73] utilise une séparation de portée dans la fonctionnelle d’échange et dans la partie d’échange exact. En revanche, dans cette méthode aucune séparation de portée n’est faite dans la fonctionnelle de corrélation  $E_c[n]$  et dans la corrélation SCS-MP2.

L’approche RSH+MP2 [30] est basée sur l’extension multi-déterminantale de la méthode de Kohn-Sham avec séparation de portée à la fois dans la fonctionnelle d’échange-corrélation et dans la partie fonction d’onde et sera détaillée dans la sous-section 2.4.3. Enfin, nous pouvons citer une autre approche rigoureuse dans la référence [89] qui repose sur une décomposition de type CAM et qui se distingue en autres de l’approche RSH+MP2 par l’inclusion d’une fraction de fonction d’onde dans la région de courte portée.

## 2.4.2 Extension multi-déterminantale basée sur une connexion adiabatique linéaire

L'objectif de cette sous-section consiste à rappeler les équations maîtresses qui ont apporté une justification théorique aux méthodes doubles hybrides reposant sur une séparation linéaire de l'interaction bi-électronique.

La connexion adiabatique permet de relier le système de Kohn-Sham où les électrons n'interagissent pas ( $\lambda = 0$ ) au système réel ( $\lambda = 1$ ) en variant de façon linéaire l'interaction bi-électronique. L'hamiltonien dépendant du paramètre  $\lambda$  assurant la connexion adiabatique s'écrit ainsi :

$$\hat{H}^\lambda = \hat{T} + \hat{V}_{ne} + \lambda \hat{W}_{ee} + \hat{V}_{Hxc}^\lambda[n]. \quad (2.32)$$

On rappelle que  $\hat{T}$  est l'opérateur lié à l'énergie cinétique des électrons,  $\hat{V}_{ne}$  l'opérateur lié à l'attraction noyaux-électrons,  $\hat{W}_{ee}$  est l'opérateur traduisant la répulsion électronique et  $\hat{V}_{Hxc}^\lambda[n]$  est le potentiel d'Hartree-échange-corrélation qui maintient la densité mono-électronique constante quelque soit la valeur de  $\lambda$  (comprise entre 0 et 1). En utilisant le formalisme de l'extension multi-déterminantale de la méthode de Kohn-Sham, il est possible d'exprimer quelque soit la valeur du paramètre  $\lambda$  l'énergie exacte de l'état fondamental en effectuant une minimisation sur des fonctions d'onde multi-déterminantales  $\Psi$  :

$$E_0 = \min_{\Psi} \left\{ \langle \Psi | \hat{T} + \hat{V}_{ne} + \lambda \hat{W}_{ee} | \Psi \rangle + \bar{E}_{Hxc}^\lambda[n_\Psi] \right\}. \quad (2.33)$$

Dans l'équation (2.33), la notation  $n_\Psi$  réfère à la densité électronique venant de  $\Psi$ . Le terme  $\bar{E}_{Hxc}^\lambda[n]$  correspond à la différence entre la fonctionnelle usuelle  $E_{Hxc}^\lambda[n]$  et la fonctionnelle d'Hartree-échange-corrélation  $E_{Hxc}^\lambda[n]$  dépendante du paramètre  $\lambda$  et qui est liée à l'hamiltonien défini dans l'équation (2.32). La fonctionnelle complémentaire  $\bar{E}_{Hxc}^\lambda[n_\Psi]$  génère le potentiel  $\hat{V}_{Hxc}^\lambda[n]$  dans l'équation (2.32) qui maintient la densité électronique du système constante :  $\hat{V}_{Hxc}^\lambda[n] = \int d\mathbf{r} \hat{n}(\mathbf{r}) \delta \bar{E}_{Hxc}^\lambda[n] / \delta n(\mathbf{r})$ . On rappelle que  $\hat{n}(\mathbf{r})$  est l'opérateur densité.

Les fonctionnelles d'Hartree et d'échange ont une dépendance au premier ordre vis-à-vis de l'interaction bi-électronique, il en ressort que  $\bar{E}_H^\lambda[n]$  et  $\bar{E}_x^\lambda[n]$  ont une dépendance linéaire par rapport au paramètre  $\lambda$  :

$$\bar{E}_H^\lambda[n] = (1 - \lambda)E_H[n], \quad (2.34)$$

$$\bar{E}_x^\lambda[n] = (1 - \lambda)E_x[n]. \quad (2.35)$$

On rappelle que les fonctionnelles  $E_H[n]$  et  $E_x[n]$  sont les fonctionnelles usuelles d'Hartree et d'échange. La contribution de la fonctionnelle de corrélation n'est pas linéaire. Celle-ci s'obtient en effectuant une transformation uniforme de la densité électronique. Nous renvoyons aux références [90] [91] [92] [93] pour de plus amples informations. La fonctionnelle de corrélation  $\bar{E}_c^\lambda[n]$  s'écrit :

$$\bar{E}_c^\lambda[n] = E_c[n] - \lambda^2 E_c[n_{1/\lambda}], \quad (2.36)$$

$$\text{avec : } n_{1/\lambda}(\mathbf{r}) = (1/\lambda)^3 n(\mathbf{r}/\lambda). \quad (2.37)$$

Nous rappelons que dans l'équation (2.36), le terme  $E_c[n]$  réfère à la fonctionnelle usuelle de corrélation. Nous avons également introduit la densité  $n_{1/\lambda}(\mathbf{r})$  obtenue après une transformation uniforme de la densité électronique  $n(\mathbf{r})$ . Nous pouvons définir l'approximation DS1H ("density-scaled one-parameter hybrid") qui restreint l'équation (2.33) aux fonctions d'onde mono-déterminantale  $\Phi$  :

$$E^{\text{DS1H},\lambda} = \min_{\Phi} \left\{ \langle \Phi | \hat{T} + \hat{V}_{\text{ne}} + \lambda \hat{W}_{\text{ee}} | \Phi \rangle + \bar{E}_{\text{Hxc}}^\lambda[n_\Phi] \right\}. \quad (2.38)$$

L'énergie obtenue  $E^{\text{DS1H},\lambda}$  dépend nécessairement du paramètre  $\lambda$ . La fonction d'onde représentée par un seul déterminant  $\Phi^\lambda$  est obtenue suite à la résolution de l'équation aux valeurs propres de manière auto-cohérente :

$$\left( \hat{T} + \hat{V}_{\text{ne}} + \lambda \hat{V}_{\text{Hx}}^{\text{HF}}[\Phi^\lambda] + \hat{V}_{\text{Hxc}}^\lambda[n_{\Phi^\lambda}] \right) |\Phi^\lambda\rangle = \mathcal{E}_0^\lambda |\Phi^\lambda\rangle. \quad (2.39)$$

Dans l'équation (2.39), le terme  $\hat{V}_{\text{Hx}}^{\text{HF}}[\Phi^\lambda]$  correspond au potentiel non-local Hartree-Fock évalué avec la fonction d'onde  $\Phi^\lambda$  tandis que le terme  $\hat{V}_{\text{Hxc}}^\lambda$  introduit précédemment est l'opérateur local Hartree-échange-corrélation évalué à la densité  $n_{\Phi^\lambda}$ . L'équation (2.39) est en pratique décomposée en une équation mono-électronique du type Kohn-Sham. Pour des raisons pratiques, la fonction d'onde DS1H notée  $\Phi^\lambda$  et la densité électronique notée  $n_{\Phi^\lambda}$  vont être désormais respectivement notée  $\Phi$  et  $n$  jusqu'à la fin de cette sous-section.

L'énergie DS1H s'écrit :

$$E^{\text{DS1H},\lambda} = \langle \Phi | \hat{T} + \hat{V}_{\text{ne}} | \Phi \rangle + E_{\text{H}}[\mathbf{n}] + \lambda E_{\text{x}}^{\text{HF}}[\Phi] + (1 - \lambda)E_{\text{x}}[\mathbf{n}] + E_{\text{c}}[\mathbf{n}] - \lambda^2 E_{\text{c}}[\mathbf{n}_1/\lambda]. \quad (2.40)$$

Dans l'équation (2.40), les seuls termes à approximer sont  $E_{\text{x}}[\mathbf{n}]$  et  $E_{\text{c}}[\mathbf{n}]$ . Dans l'annexe de la référence [82], les relations pour calculer la fonctionnelle  $E_{\text{c}}[\mathbf{n}_1/\lambda]$  et les potentiels associés à ce terme sont données. Il est possible d'approximer ce terme par  $E_{\text{c}}[\mathbf{n}]$ , ce qui nous amène à l'approximation 1H pour "one parameter hybrid" :

$$E^{\text{1H},\lambda} = \langle \Phi | \hat{T} + \hat{V}_{\text{ne}} | \Phi \rangle + E_{\text{H}}[\mathbf{n}] + \lambda E_{\text{x}}^{\text{HF}}[\Phi] + (1 - \lambda)E_{\text{x}}[\mathbf{n}] + (1 - \lambda^2)E_{\text{c}}[\mathbf{n}]. \quad (2.41)$$

Dans l'équation (2.38), il manque l'énergie de corrélation associé à l'opérateur  $\lambda \hat{W}_{\text{ee}}$ . Cette énergie de corrélation peut être calculée par une théorie de perturbation non-linéaire de Rayleigh-Schrödinger [94] [95], en partant de la référence DS1H, c'est-à-dire de la fonction d'onde et de la densité électronique obtenue dans l'équation (2.38). Pour cela, on introduit dans l'expression de l'énergie un paramètre de perturbation notée  $\varepsilon$  :

$$E^{\lambda,\varepsilon} = \min_{\Psi} \left\{ \langle \Psi | \hat{T} + \hat{V}_{\text{ne}} + \lambda \hat{V}_{\text{Hx}}^{\text{HF}}[\Phi] + \varepsilon \lambda \hat{\mathcal{W}} | \Psi \rangle + \bar{E}_{\text{Hxc}}^{\lambda}[\mathbf{n}_{\Psi}] \right\}, \quad (2.42)$$

où dans l'équation (2.42) le terme  $\hat{\mathcal{W}}$  est l'opérateur de perturbation et est défini comme étant la différence entre le potentiel défini par l'opérateur  $\hat{W}_{\text{ee}}$  et le potentiel Hartree-Fock  $\hat{V}_{\text{Hx}}^{\text{HF}}$ . Dans le cas, où la valeur du paramètre de perturbation  $\varepsilon$  vaut zéro, nous retrouvons l'équation (2.38). Dans le cas où la valeur du paramètre  $\varepsilon = 1$ , nous retrouvons l'énergie exacte (si l'expression des termes  $E_{\text{x}}$  et  $E_{\text{c}}$  sont connus) du système, indépendamment de la valeur de  $\lambda$ .

Par analogie avec la théorie MP2, la somme de l'énergie à l'ordre zéro et à l'ordre un, nous donne l'énergie DS1H dont l'expression est donnée dans l'équation (2.40). De même, il existe un théorème de Brillouin, c'est-à-dire que seules les excitations doubles contribuent à l'énergie de corrélation au second ordre de la perturbation. L'expression de la correction à l'énergie au second ordre s'écrit :

$$E^{\lambda,(2)} = -\lambda^2 \sum_{a < b}^{\text{occ}} \sum_{r < s}^{\text{vir}} \frac{|\langle ab || rs \rangle|^2}{\varepsilon_r + \varepsilon_s - \varepsilon_a - \varepsilon_b} = \lambda^2 E_{\text{c}}^{\text{MP2}}. \quad (2.43)$$

Nous rappelons que dans l'équation (2.43) les indices a et b représentent des spin-orbitales DS1H occupées tandis que r et s représentent des spin-orbitales DS1H vir-

tuelles et  $\langle ab||rs \rangle$  sont les intégrales bi-électroniques anti-symétriques. Enfin, l'expression de l'énergie d'échange-corrélation DS1DH s'écrit :

$$E_{xc}^{DS1DH,\lambda} = \lambda E_x^{HF} + (1 - \lambda) E_x^{DFT}[n] + E_c^{DFT} - \lambda^2 E_c^{DFT}[n_{1/\lambda}] + \lambda^2 E_c^{MP2}. \quad (2.44)$$

En négligeant le "scaling de la densité", c'est-à-dire en faisant l'approximation  $E_c[n_{1/\lambda}] \approx E_c[n]$ , il est possible d'établir une connexion avec les méthodes doubles hybrides répertoriées dans le tableau 2.1 :

$$E_{xc}^{1DH,\lambda} = \lambda E_x^{HF} + (1 - \lambda) E_x^{DFT}[n] + (1 - \lambda^2) E_c^{DFT}[n] + \lambda^2 E_c^{MP2}. \quad (2.45)$$

Ce qui revient à dire que  $\lambda = a_x$  et  $\lambda^2 = a_c$ .

### 2.4.3 Extension multi-déterminantale basée sur une connexion adiabatique avec séparation de portée

L'idée qui consiste à décomposer l'interaction bi-électronique en fonction de la distance inter-électronique a été développée dans les références [96] [27]. En règle générale, on traite les interactions électroniques de courte portée par une fonctionnelle de la densité et les interactions électroniques de longue portée par une méthode de type fonction d'onde. L'objectif étant de traiter les interactions de nature dispersive par une méthode de type fonction d'onde car elles sont manquantes dans les approximations faites dans la fonctionnelle de corrélation. De plus, en traitant la région de courte portée par une fonctionnelle de la densité, le problème lié à la convergence polynomiale de l'énergie de corrélation des méthodes post-HF n'existe plus, ce problème étant dû au "cusp" de la fonction d'onde lorsque la distance inter-électronique est égale à zéro.

Le point de départ des méthodes hybrides à séparation de portée consiste à décomposer l'interaction électronique coulombienne  $w_{ee}(r_{12}) = 1/r_{12}$  selon [96] [27] :

$$w_{ee}(r_{12}) = w_{ee}^{lr,\mu}(r_{12}) + w_{ee}^{sr,\mu}(r_{12}), \quad (2.46)$$

où  $w_{ee}^{lr,\mu}(r_{12}) = \text{erf}(\mu r)/r_{12}$  est l'interaction de longue portée<sup>5</sup> et  $w_{ee}^{sr,\mu}(r_{12}) = \text{erfc}(\mu r)/r_{12}$  est la partie complémentaire de courte portée ( $\text{erfc}(x) = 1 - \text{erf}(x)$ ). La décomposition de l'interaction est contrôlée par le paramètre  $\mu$  ( $0 \leq \mu < \infty$ ). Dans le cas où la valeur du paramètre  $\mu$  vaut zéro, l'interaction de longue portée est nulle,  $w_{ee}^{lr,\mu=0}(r) = 0$ , tandis

---

5. On a eu recours à la fonction erreur notée "erf", il existe d'autres fonctions permettant de faire cette décomposition.

que dans la limite où  $\mu \rightarrow \infty$ , l'interaction de longue portée correspond à l'interaction coulombienne, soit  $w_{ee}^{\text{lr},\mu \rightarrow \infty}(r) = 1/r$ . De manière similaire, lorsque  $\mu = 0$  l'interaction de courte portée se réduit à l'interaction coulombienne  $w_{ee}^{\text{sr},\mu=0} = 1/r$  et est nulle dans le cas où  $\mu \rightarrow \infty$ .

Une extension multi-déterminantale de la méthode de Kohn-Sham basée sur une séparation de portée consiste à décomposer la fonctionnelle universelle  $F[n]$  de la manière suivante :

$$F[n] = F^{\text{lr},\mu}[n] + \bar{F}^{\text{sr},\mu}[n], \quad (2.47)$$

$$\text{avec : } F^{\text{lr},\mu}[n] = \min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{W}_{ee}^{\text{lr},\mu} | \Psi \rangle. \quad (2.48)$$

La décomposition faite dans l'équation (2.47) est une décomposition de type longue portée pure/courte portée complémentaire. Dans l'équation (2.48),  $\Psi \rightarrow n$  représente l'ensemble des fonctions d'onde antisymétriques à  $N$  électrons donnant la densité  $n$ . Nous avons également introduit l'opérateur  $\hat{W}_{ee}^{\text{lr},\mu} = (1/2) \iint w_{ee}^{\text{lr},\mu}(r_{12}) \hat{n}_2(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2$  où  $\hat{n}_2(\mathbf{r}_1, \mathbf{r}_2) = \hat{n}(\mathbf{r}_1) \hat{n}(\mathbf{r}_2) - \delta(\mathbf{r}_1 - \mathbf{r}_2) \hat{n}(\mathbf{r}_1)$  est l'opérateur densité de paires défini dans l'équation (1.6).

La contribution complémentaire de courte portée  $\bar{F}^{\text{sr},\mu}[n]$  s'écrit comme la somme de fonctionnelles d'Hartree, d'échange et de corrélation :

$$\bar{F}^{\text{sr},\mu}[n] = E_{\text{H}}^{\text{sr},\mu}[n] + E_{\text{x}}^{\text{sr},\mu}[n] + \bar{E}_{\text{c}}^{\text{sr},\mu}[n], \quad (2.49)$$

où  $E_{\text{H}}^{\text{sr},\mu}[n]$  est la fonctionnelle de Hartree de courte portée :

$$E_{\text{H}}^{\text{sr},\mu}[n] = 1/2 \iint n(\mathbf{r}_1), n(\mathbf{r}_2) w_{ee}^{\text{sr},\mu}(r_{12}) d\mathbf{r}_1 d\mathbf{r}_2. \quad (2.50)$$

Les fonctionnelles d'échange et de corrélation de courte portée s'écrivent ainsi :

$$E_{\text{x}}^{\text{sr},\mu}[n] = E_{\text{x}}[n] - E_{\text{x}}^{\text{lr},\mu}[n] = \langle \Phi[n] | \hat{W}_{ee}^{\text{sr},\mu} | \Phi[n] \rangle - E_{\text{H}}^{\text{sr},\mu}[n], \quad (2.51)$$

$$\bar{E}_{\text{c}}^{\text{sr},\mu}[n] = E_{\text{c}}[n] - E_{\text{c}}^{\text{lr},\mu}[n] = \bar{F}^{\text{sr},\mu}[n] - \langle \Phi[n] | \hat{W}_{ee}^{\text{sr},\mu} | \Phi[n] \rangle. \quad (2.52)$$

On peut noter que dans le cas où  $\mu = 0$ , la fonctionnelle de longue portée  $F^{\text{lr},\mu}[n]$  se réduit à l'énergie cinétique du système Kohn-Sham  $T_{\text{s}}[n]$  et la fonctionnelle de courte portée  $\bar{F}^{\text{sr},\mu}[n]$  se réduit à  $E_{\text{Hxc}}[n]$ . Dans le cas où  $\mu \rightarrow \infty$ , l'opérateur  $\hat{W}_{ee}^{\text{lr},\mu}$  se réduit à l'opérateur usuel  $\hat{W}_{ee}$  et la fonctionnelle  $F^{\text{lr},\mu}[n]$  devient égale à  $F[n]$ , autrement dit la

fonctionnelle  $\bar{F}^{\text{sr},\mu}[n]$  s'annule. Il est possible d'écrire l'énergie exacte de l'état fondamental  $E_0$  d'un système à  $N$  électrons soumis à un potentiel externe  $v_{\text{ne}}(\mathbf{r})$  comme :

$$\begin{aligned}
E &= \min_{n \rightarrow N} \left\{ F[n] + \int v_{\text{ne}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \right\} \\
&= \min_{n \rightarrow N} \left\{ F^{\text{lr},\mu}[n] + \bar{E}_{\text{Hxc}}^{\text{sr},\mu}[n] + \int v_{\text{ne}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \right\} \\
&= \min_{\Psi \rightarrow N} \left\{ \langle \Psi | \hat{T} + \hat{W}_{\text{ee}}^{\text{lr},\mu} + \hat{V}_{\text{ne}} | \Psi \rangle + \bar{E}_{\text{Hxc}}^{\text{sr},\mu}[n_{\Psi}] \right\}. \tag{2.53}
\end{aligned}$$

Dans la dernière expression (2.53), la minimisation s'effectue sur toutes les fonctions d'onde multi-déterminantales, anti-symétriques normalisées à  $N$  électrons. La densité  $n_{\Psi}$  est la densité provenant de la fonction d'onde  $\Psi$  ( $\langle \Psi | \hat{n}(\mathbf{r}) | \Psi \rangle = n_{\Psi}(\mathbf{r})$ ). La fonction d'onde  $\Psi^{\mu}$  issue de la minimisation sous contrainte dans l'équation (2.53) satisfait l'équation d'Euler-Lagrange et mène à une équation de Schrödinger effective :

$$\left( \hat{T} + \hat{W}_{\text{ee}}^{\text{lr},\mu} + \hat{V}_{\text{ne}} + \hat{V}_{\text{Hxc}}^{\text{sr},\mu}[n_{\Psi^{\mu}}] \right) | \Psi^{\mu} \rangle = \mathcal{E}^{\mu} | \Psi^{\mu} \rangle. \tag{2.54}$$

Dans l'équation (2.54),  $\mathcal{E}^{\mu}$  est le multiplicateur de Lagrange associé à la contrainte de normalisation de la fonction d'onde. Le terme  $\hat{V}_{\text{Hxc}}^{\text{sr},\mu}[n] = \int v_{\text{Hxc}}^{\text{sr},\mu}(\mathbf{r})\hat{n}(\mathbf{r})d\mathbf{r}$  est le potentiel complémentaire d'Hartree-échange-corrélation de courte portée. Le potentiel Hartree-échange-corrélation de courte portée est défini comme la dérivée fonctionnelle de la fonctionnelle Hartree-échange-corrélation :  $v_{\text{Hxc}}^{\text{sr},\mu}(\mathbf{r}) = \delta \bar{E}_{\text{Hxc}}^{\text{sr},\mu}[n] / \delta n(\mathbf{r})$ . L'équation (2.54) définit un hamiltonien effectif :  $\hat{H}^{\mu} = \hat{T} + \hat{V}_{\text{ne}} + \hat{W}_{\text{ee}}^{\text{lr},\mu} + \hat{V}_{\text{Hxc}}^{\text{sr},\mu}[n_{\Psi^{\mu}}]$ . L'équation (2.54) doit être résolue itérativement pour obtenir la fonction d'onde multi-déterminantale donnant l'énergie la plus basse. Lorsque la convergence est atteinte, la fonction d'onde  $| \Psi^{\mu} \rangle$  donne la densité exacte du système indépendamment de la valeur de  $\mu$ .

Ce formalisme est exact et permet une extension de la méthode KS. Dans la pratique, des approximations doivent être faites à la fois pour la fonctionnelle d'échange-corrélation de courte portée et la fonction d'onde. Des approximations au niveau LDA [27] et GGA [97] pour la fonctionnelle d'échange-corrélation de courte portée existent. En ce qui concerne la fonction d'onde, un traitement des interactions de longue portée a été fait en interaction de configurations (IC) [98], MCSCF [99], CCSD(T) [100], RPA [101] [102], NEVPT2 [103] ou MP2 [30]. Pour pouvoir appréhender le chapitre 4 nous allons détailler la méthode RSH+MP2 de la référence [30].

Dans l'approche RSH+MP2, on restreint - de manière analogue à la méthode HF - la minimisation faite dans l'équation (2.53) aux fonctions d'onde anti-symétriques normali-

sées à N électrons composés d'un seul déterminant :

$$E = \min_{\Phi \rightarrow N} \left\{ \langle \Phi | \hat{T} + \hat{W}_{ee}^{\text{lr},\mu} + \hat{V}_{\text{ne}} | \Phi \rangle + \bar{E}_{\text{Hxc}}^{\text{sr},\mu}[\text{n}\Phi] \right\}. \quad (2.55)$$

Le déterminant associé à la minimisation de l'équation (2.55)  $\Phi^\mu$  doit satisfaire l'équation d'Euler-Lagrange :

$$\left( \hat{T} + \hat{V}_{\text{Hx}}^{\text{lr},\mu,\text{HF}}[\Phi^\mu] + \hat{V}_{\text{ne}} + \hat{V}_{\text{Hxc}}^{\text{sr},\mu}[\text{n}\Phi^\mu] \right) |\Phi^\mu\rangle = \mathcal{E}_{\text{RSH}}^\mu |\Phi^\mu\rangle. \quad (2.56)$$

Dans l'équation (2.56)  $\hat{V}_{\text{Hx}}^{\text{lr},\mu,\text{HF}}$  est l'opérateur Hartree-d'échange de longue portée et  $\mathcal{E}_{\text{RSH}}^\mu$  le multiplicateur de Lagrange associé à la contrainte de normalisation. L'équation (2.56) définit un Hamiltonien effectif qui après recombinaison du terme d'Hartree s'écrit :  $\hat{H}_{\text{RSH}}^\mu = \hat{T} + \hat{V}_{\text{ne}} + \hat{V}_{\text{H}}[\Phi^\mu] + \hat{V}_{\text{x}}^{\text{lr},\mu,\text{HF}}[\Phi^\mu] + \hat{V}_{\text{xc}}^{\text{sr},\mu}[\text{n}\Phi^\mu]$ . Le caractère hybride de cet hamiltonien vient de la combinaison d'un potentiel d'échange de longue portée non-local avec un potentiel d'échange courte portée local. Afin d'obtenir la fonction d'onde  $\Phi^\mu$ , il faut résoudre l'équation (2.56) de manière itérative. Il est important de noter que la fonction d'onde  $\Phi^\mu$  ne peut reproduire la densité exacte du système.

Un moyen de tenir compte des effets de corrélation de longue portée est d'utiliser le déterminant  $\Phi^\mu$  comme point de départ et d'appliquer la théorie des perturbations non-linéaire de Rayleigh-Schrödinger [94] [95]. Pour cela, on introduit un paramètre de perturbation que l'on note  $\varepsilon$ . Nous pouvons ainsi exprimer l'énergie du système de la façon suivante :

$$E^{\mu,\varepsilon} = \min_{\Psi \rightarrow N} \left\{ \langle \Psi | \hat{T} + \hat{V}_{\text{ne}} + \hat{V}_{\text{Hx,HF}}^{\text{lr},\mu}[\Phi^\mu] + \varepsilon \hat{\mathcal{W}}^{\text{lr},\mu} | \Psi \rangle + \bar{E}_{\text{Hxc}}^{\text{sr},\mu}[\text{n}\Psi] \right\}, \quad (2.57)$$

où la recherche s'effectue sur des fonctions d'ondes anti-symétriques multi-déterminantales  $\Psi$  normalisées à N électrons. Dans l'équation (2.57), nous avons introduit l'opérateur  $\mathcal{W}^\mu$  qui est un analogue de l'opérateur de perturbation dans la théorie Møller-Plesset :  $\mathcal{W}^{\text{lr},\mu} = \hat{W}_{ee}^{\text{lr},\mu} - \hat{V}_{\text{Hx,HF}}^{\text{lr},\mu}[\Phi^\mu]$ .

La fonction d'onde qui minimise l'équation (2.57) est obtenue grâce à l'équation d'Euler-Lagrange suivante :

$$\left( \hat{T} + \hat{V}_{\text{ne}} + \hat{V}_{\text{Hx,HF}}^{\text{lr},\mu}[\Phi^\mu] + \varepsilon \hat{\mathcal{W}}^{\text{lr},\mu} + \hat{V}_{\text{Hxc}}^{\text{sr},\mu}[\text{n}\Psi^{\mu,\varepsilon}] \right) |\Psi^{\mu,\varepsilon}\rangle = \mathcal{E}^{\mu,\varepsilon} |\Psi^{\mu,\varepsilon}\rangle. \quad (2.58)$$

Dans le cas où  $\varepsilon = 0$ , l'équation (2.58) se réduit à l'équation (2.56), c'est-à-dire que  $\Psi^{\mu,\varepsilon=0} = \Phi^\mu$  et  $\mathcal{E}^{\mu,\varepsilon=0} = \mathcal{E}_{\text{RSH}}^\mu$ . Dans le cas où  $\varepsilon = 1$ , l'équation  $\Psi^{\mu,\varepsilon=1} = \Psi^\mu$  et  $\mathcal{E}^{\mu,\varepsilon=1} = \mathcal{E}^\mu$ , nous retrouvons l'équation (2.54). En effectuant un développement en puissance d' $\varepsilon$  de l'énergie, de la fonction d'onde et du multiplicateurs de Lagrange ( $\mathcal{E}^{\mu,\lambda,\varepsilon} = \sum_{k=0}^{\infty} \varepsilon^k \mathcal{E}^{\mu,(k)}$ ,  $\Psi^{\mu,\varepsilon} = \sum_{k=0}^{\infty} \varepsilon^k \Psi^{\mu,(k)}$ , et  $E^{\mu,\varepsilon} = \sum_{k=0}^{\infty} \varepsilon^k E^{\mu,(k)}$ ), nous retrouvons après sommation des termes énergétiques d'ordre zéro et premier, l'expression de l'énergie RSH :

$$E^{\mu,\text{RSH}} = E^{\mu,(0)} + E^{\mu,(1)} = \langle \Phi^\mu | \hat{T} | \Phi^\mu \rangle + \int d\mathbf{r} v_{\text{ne}}(\mathbf{r}) n_{\Phi^\mu}(\mathbf{r}) + E_{\text{H}}[n] + E_{\text{X}}^{\text{lr},\mu,\text{HF}}[\Phi^\mu] + E_{\text{XC}}^{\text{sr},\mu}[n_{\Phi^\mu}], \quad (2.59)$$

où  $E_{\text{H}}[n]$  est l'énergie d'Hartree,  $E_{\text{X}}^{\text{lr},\mu,\text{HF}}[\Phi^\mu]$  est l'énergie d'échange exact de longue portée et  $E_{\text{XC}}^{\text{sr},\mu}[n_{\Phi^\mu}]$  représente la contribution d'échange-corrélation DFT de courte portée. L'énergie de corrélation de longue portée s'obtient à partir du deuxième ordre de la perturbation. Par analogie avec la théorie MP2, les excitations simples ne contribuent pas à l'énergie de corrélation. La contribution des termes non-linéaires en  $\varepsilon$  provenant de la partie DFT étant nulle au second ordre de la perturbation. L'expression de l'énergie de corrélation au second ordre s'écrit :

$$E^{\mu,(2)} = E_{\text{C}}^{\text{lr},\mu} = - \sum_{\text{a}<\text{b}}^{\text{occ}} \sum_{\text{r}<\text{s}}^{\text{vir}} \frac{|\langle \text{ab} | \hat{w}_{\text{ee}}^{\text{lr},\mu} | \text{rs} \rangle - \langle \text{ab} | \hat{w}_{\text{ee}}^{\text{lr},\mu} | \text{sr} \rangle|^2}{\varepsilon_{\text{r}} + \varepsilon_{\text{s}} - \varepsilon_{\text{a}} - \varepsilon_{\text{b}}}. \quad (2.60)$$

Nous obtenons l'énergie RSH+MP2 en sommant l'équation (2.60) et l'équation (2.59). On rappelle que dans l'équation précédente, les indices "a" et "b" représentent des spin-orbitales occupées tandis que "r" et "s" représentent des spin-orbitales virtuelles. Le terme  $|\langle \text{ab} | \hat{w}_{\text{ee}}^{\text{lr},\mu} | \text{rs} \rangle|$  est une intégrale bi-électronique associée à l'interaction de longue portée.

## Chapitre 3

**Les méthodes interprétatives  
permettant une caractérisation  
des interactions mises en jeu  
au sein des complexes  
micro-hydratés**

### 3.1 Utilisation du MEP pour la prédiction des complexes interagissant par liaison hydrogène

Le potentiel électrostatique  $V(\mathbf{r})$  créée dans l'espace moléculaire par l'ensemble de ses particules chargées - nommé "Molecular Electrostatic Potential" (MEP) - est une grandeur physique utilisée pour l'analyse des interactions non-covalentes. Le potentiel électrostatique est lié à la loi de Coulomb. Toute charge ponctuelle  $Q_a$  crée un potentiel électrostatique dans l'espace environnant à une distance  $R$  dont la valeur,  $V(R)$ , s'écrit en unités atomique :

$$V(R) = Q_a/R. \quad (3.1)$$

Lorsqu'il s'agit d'une molécule ou d'un complexe, les noyaux et les électrons sont traités comme des charges ponctuelles. Le potentiel électrostatique qu'ils créent au point  $\mathbf{r}$  dans l'espace environnant peut être obtenu, en principe, en effectuant une sommation de l'équation (3.1) sur l'ensemble des noyaux et des électrons composant le système. En faisant cela, nous avons recourt à l'approximation de Born-Oppenheimer. Or, cette approximation ne peut être appliquée qu'aux noyaux et donc la somme de l'équation (3.1) sur les électrons est remplacée par une intégration sur la densité électronique. En tenant compte de l'ensemble des noyaux et des électrons, l'équation (3.1) se réécrit :

$$V(\mathbf{r}) = \sum_A \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|} - \int \frac{n(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r}' - \mathbf{r}|}. \quad (3.2)$$

Nous rappelons que dans l'équation (3.2)  $Z_A$  représente la charge du noyau  $A$ , situé à la position  $\mathbf{R}_A$ ,  $|\mathbf{R}_A - \mathbf{r}|$  est la distance entre le point  $\mathbf{r}$  et la position  $\mathbf{R}_A$ . Le dénominateur  $|\mathbf{r}' - \mathbf{r}|$  correspond à la position relative de la densité électronique  $n(\mathbf{r}')$  par rapport au point de référence  $\mathbf{r}$

Le MEP se distingue des propriétés définies telles que les charges atomiques et les indices de liaison qui ne sont pas des observables physiques. Le MEP est une propriété physique observable : il peut être déterminé expérimentalement, par des méthodes de diffraction aux rayons X [104] et peut être également calculé. Le MEP est usuellement calculé sur la "surface" de la molécule. En suivant, la suggestion de Bader *et al.* [105], la "surface" de la molécule correspond la surface d'iso-densité prise à 0.001 u.a. Le recourt à une surface d'iso-densité électronique de 0.001 u.a. permet de refléter les principales caractéristiques d'une molécule, les doublets non-liants et l'anisotropie de la densité électronique.

Dans le système international, l'unité du MEP est le volt. Cependant, il est d'usage utiliser une unité d'énergie le kcal/mol. Lorsqu'il est donné en unité d'énergie,  $V(\mathbf{r})$  correspond à l'énergie d'interaction avec une charge ponctuelle positive située en  $\mathbf{r}$ . Lorsque l'on calcule le MEP sur une surface d'iso-densité (typiquement 0,001 u.a.), nous utiliserons le terme  $V_S(\mathbf{r})$ . Nous employons également les termes de  $V_{S_{\max}}$  et de  $V_{S_{\min}}$  pour désigner respectivement les régions où les valeurs des MEP sont localement maximales ou minimales. En règle générale, les  $V_{S_{\max}}$  désignent les sites électrophiles alors que les  $V_{S_{\min}}$  représentent les sites nucléophiles. D'après les travaux de Politzer et al., toute interaction inter- ou intramoléculaire non-covalente peut être parfaitement caractériser comme une interaction entre un  $V_{S_{\min}}$  d'une espèce moléculaire et un  $V_{S_{\max}}$  d'une autre espèce moléculaire (ou un autre partie de la même espèce).

La valeur du MEP projetée sur une surface d'iso-densité est calculée pour les molécules dans leurs états fondamentaux non perturbés. Cependant dès qu'une molécule est au voisinage d'une autre molécule, la distribution de charge de chacun des monomères subit des modifications, c'est le phénomène de polarisation. Ce phénomène devient de plus en plus important au fur et à mesure que les monomères sont proches. La conséquence est que les potentiels électrostatiques des monomères à l'état fondamental deviennent graduellement moins significatifs.

Cependant nous tenons à souligner que les labels  $V_{S_{\max}}$  et  $V_{S_{\min}}$  réfèrent spécifiquement aux maxima et minima locaux de  $V(\mathbf{r})$  sur une surface d'iso-densité. Ils ne sont pas des maxima et des minima globaux du potentiel électrostatique. Le potentiel électrostatique d'une molécule peut avoir un ou plusieurs minima, en revanche, les maxima sont situés sur les noyaux [106].

L'utilisation du MEP pour l'analyse des interactions non-covalentes a été proposée dans le cadre de deux types d'approches :

- une approche explicative, a posteriori, pour comprendre les structures observées, notamment dans le cas des interactions tetrel, chalcogène, pnictogène et halogène. Ces études ont conduit à l'introduction des concepts de  $\sigma$ -hole et  $\pi$ -hole. [107, 108]
- une approche prédictive, a priori, qui est basée sur la topologie du MEP. Dans leur méthode, appelée EPIC (Electrostatic Potential for Intermolecular Complexes), les auteurs proposent d'utiliser la topologie du potentiel électrostatique. [109]

Dans le cadre de cette thèse, nous utilisons essentiellement la valeur du MEP sur une surface d'iso-densité dans une approche prédictive :

- permettant de guider le chimiste dans la construction des structures initiales,

- basée sur le processus de formation des complexes par approche à longue distance du soluté et du cluster d'eau.

Le MEP a été utilisé pour l'étude de l'hydratation de polluants atmosphériques, en particulier le diacétyl (DAC). L'étude MEP du DAC conduit à l'identification de deux zones enrichies en électrons, autour de chacun des atomes d'oxygène. Une zone particulièrement appauvrie en électrons est identifiée au-dessus et en-dessous du plan de la molécule, au voisinage de la liaison C-C entre les deux fonctions carbonyles (Figure 3.1). Cette particularité peut être rapprochée des  $\pi$ -holes mis en évidence notamment dans le cas des cétones : il s'agit de régions électrophiles situées perpendiculairement à une liaison chimique.

Les fonctions méthyles constituent également des zones électrophiles, au niveau de chaque axe C-H se trouve une zone particulièrement appauvrie en électrons que l'on nomme des  $\sigma$ -hole. La molécule d'eau est caractérisée par deux zones : une zone électrophile au voisinage des deux atomes d'hydrogène, et une zone nucléophile au voisinage de l'atome d'oxygène (Figure 3.1).

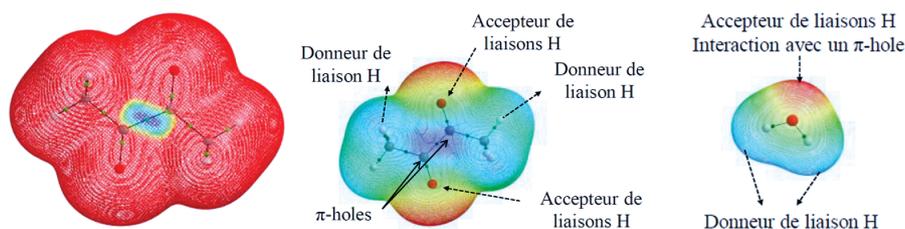


FIGURE 3.1 – Mise en évidence des régions électrophiles et nucléophiles du DAC et de la molécule d'eau

Cette analyse a permis de mettre en évidence trois directions d'approches d'une molécule d'eau par rapport au DAC. Les optimisations de géométrie montrent que ces trois directions d'approche conduisent à trois isomères stables et quasi iso-énergétiques.

## 3.2 L'analyse de la densité électronique par la théorie QTAIM

Une analyse topologique consiste à partitionner l'espace moléculaire en utilisant la théorie mathématique de systèmes dynamiques gradient appliquée à une fonction locale continue et dérivable en tout point de l'espace. Cette fonction locale présente des points critiques qui lui sont caractéristiques : des attracteurs (maxima), des minima, et des points selles. Un point critique est un point de l'espace pour lequel la dérivée

première de la fonction s'annule. L'approche QTAIM (Quantum Theory of Atoms In Molecule) [4] [110] considère la densité électronique  $n(\mathbf{r})$  comme fonction locale du système moléculaire. L'analyse topologique de cette fonction permet de scinder l'espace moléculaire en bassin atomique. Les bassins atomiques sont séparés par des *surface de flux nul* qui satisfont l'équation suivante :

$$\nabla n(\mathbf{r}) \cdot \mathbf{k}(\mathbf{r}) = 0, \quad (3.3)$$

où  $\mathbf{k}(\mathbf{r})$  représente un vecteur normal à la surface du bassin. On peut faire la distinction entre un minimum local, un maximum local et un point selle en considérant les dérivées secondes. La densité électronique possède neuf dérivées secondes qui peuvent être arrangées dans la matrice "Hessienne". Lorsque cette matrice est évaluée sur un point critique situé à  $\mathbf{r}_c$ , elle peut s'écrire de manière suivante :

$$\mathbf{A}(\mathbf{r}_c) = \begin{pmatrix} \frac{\delta^2 n}{\delta x^2} & \frac{\delta^2 n}{\delta x \delta y} & \frac{\delta^2 n}{\delta x \delta z} \\ \frac{\delta^2 n}{\delta y \delta x} & \frac{\delta^2 n}{\delta y^2} & \frac{\delta^2 n}{\delta x \delta z} \\ \frac{\delta^2 n}{\delta z \delta x} & \frac{\delta^2 n}{\delta z \delta y} & \frac{\delta^2 n}{\delta z^2} \end{pmatrix}. \quad (3.4)$$

La matrice Hessienne de l'équation (3.4) peut être diagonalisée parce qu'elle est réelle et symétrique. Nous obtenons la matrice suivante :

$$\mathbf{A}_{\text{diag}}(\mathbf{r}_c) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}. \quad (3.5)$$

La détermination des valeurs propres  $\lambda_1, \lambda_2$  et  $\lambda_3$  permet une classification des points critiques. Le rang ( $\omega$ ) de la matrice hessienne correspond au nombre de valeurs propres non nulles. La signature ( $\zeta$ ) de la matrice hessienne est l'excès (ou le défaut) de valeurs propres positives par rapport aux valeurs propres négatives. Il existe quatre types de points critiques stables ayant trois valeurs propres non nulles :

$(\omega, \zeta)$	nombre de valeur propres positive ou négative	caractéristique du point stationnaire	nom du point critique
(3,+3)	3 valeurs propres $>0$	minimum local	cage critical point (CCP)
(3,+1)	2 valeurs propres $>0$	point selle	ring critical point (RCP)
(3,-1)	2 valeurs propres $<0$	point selle	bond critical point (BCP)
(3,-3)	3 valeurs propres $<0$	maximum local	Attracteur nucléaire (NCP)

Tableau 3.1 – Points critiques de rang 3

Le nombre  $nb$  et le type des points critiques pouvant coexister dans une molécule doivent suivre la relation topologique générale dite de Poincaré-Hopf :

$$nb_{\text{NCP}} - nb_{\text{BCP}} + nb_{\text{RCP}} - nb_{\text{CCP}} = 1, \quad (3.6)$$

pour les systèmes finis. Un "BCP" connecte deux points caractérisés par (3,-3). Selon la théorie QTAIM, une telle connexion forme un "chemin de liaison" (ou "bond-path" ). La présence d'un "bond path" (ou un BCP) entre deux atomes est la condition nécessaire et suffisante pour l'existence de liaison covalente ou non-covalente entre deux atomes en question [111]. Les liaisons chimiques sont classées en deux catégories :

- d'une part, nous avons les interactions à partage d'électrons (shared interactions) : le laplacien de la densité  $\nabla^2 n(\mathbf{r})$  est négatif au point critique de liaison (3,-1). Ce type d'interaction traduit souvent une liaison covalente. Lorsque  $\nabla^2 n(\mathbf{r}) < 0$  , on observe une concentration locale de la densité électronique au BCP.
- D'autre part, nous avons les interactions dite à couche fermée (closed Shell interactions) : le laplacien de la densité est positive au point BCP :  $\nabla^2 n(\mathbf{r}) > 0$ . Ce type d'interaction correspond aux interactions "ioniques" ou "van der Waals". Lorsque  $\nabla^2 n(\mathbf{r}) > 0$  , on observe une diminution locale de la densité électronique au BCP.

Des informations complémentaires sont fournis par les valeurs propres négatives de la matrice hessienne au point critique de liaison,  $\lambda_1$  et  $\lambda_2$ . En considérant la convention  $|\lambda_1| > |\lambda_2|$ . On peut définir l'ellipticité de la liaison :

$$\varepsilon = \frac{\lambda_1}{\lambda_2} - 1 \quad (3.7)$$

L'ellipticité fournit une mesure quantitative de l'anisotropie de la densité électronique au BCP. L'ensemble des trois critères (densité électronique, le laplacien de la densité électronique et l'ellipticité) concernant un point critique décrit totalement la liaison chimique d'un point de vue topologique. L'application de cette théorie aux systèmes à liaison hydrogène a conduit à la proposition d'un certain nombre de traits caractéristiques [12] qui peuvent être résumés comme suit :

- Il existe un chemin de liaison ("bond-path") entre l'hydrogène (A-H) et l'accepteur d'hydrogène (B), autrement dit un BCP "B... H" reliant ces deux atomes. B étant l'accepteur de proton dans la liaison hydrogène.
- La densité électronique au BCP "B... H" est comprise dans l'intervalle [0.002-0.034] u.a.
- Le laplacien de la densité électronique au BCP "B... H" devrait être compris dans

l'intervalle [0.024-0.139] u.a.

- La pénétration mutuelle des atomes d'hydrogène et d'accepteur doit être observée lors de la formation d'une liaison hydrogène : la distance caractéristique de la liaison hydrogène " $r(B...H)$ " est inférieure à la somme des rayons de van der Waals.
- Il y a une perte de charge électronique de l'atome d'hydrogène.
- On peut mettre en évidence une déstabilisation énergétique de l'atome d'hydrogène car il existe une différence positive entre l'énergie totale de l'atome d'hydrogène dans le complexe et l'énergie de l'atome H non impliqué dans la liaison hydrogène.
- La polarisation dipolaire diminue lors de la complexation de l'atome d'hydrogène.
- Il y a une diminution du volume de l'atome d'hydrogène après formation d'une liaison hydrogène.

Il est intéressant de noter que la densité électronique au point critique de la liaison hydrogène peut être corrélée à l'énergie d'interaction d'une part, et à la variation de la distance A-H d'autre part. [112,113]

Deuxième partie

Développement  
Méthodologique

## Chapitre 4

# A general range-separated double-hybrid density-functional theory

Ce chapitre a été écrit sous la direction de J. Toulouse et été publié dans Journal of Chemical Physics [114]

## 4.1 Abstract

A range-separated double-hybrid (RSDH) scheme which generalizes the usual range-separated hybrids and double hybrids is developed. This scheme consistently uses a two-parameter Coulomb-attenuating-method (CAM)-like decomposition of the electron-electron interaction for both exchange and correlation in order to combine Hartree-Fock exchange and second-order Møller-Plesset (MP2) correlation with a density functional. The RSDH scheme relies on an exact theory which is presented in some detail. Several semi-local approximations are developed for the short-range exchange-correlation density functional involved in this scheme. After finding optimal values for the two parameters of the CAM-like decomposition, the RSDH scheme is shown to have a relatively small basis dependence and to provide atomization energies, reaction barrier heights, and weak intermolecular interactions globally more accurate or comparable to range-separated MP2 or standard MP2. The RSDH scheme represents a new family of double hybrids with minimal empiricism which could be useful for general chemical applications.

## 4.2 Introduction

Over the past two decades, density-functional theory (DFT) [36] within the Kohn-Sham (KS) scheme [22] has been a method of choice to study ground-state properties of electronic systems. KS DFT is formally exact, but it involves the so-called exchange-correlation energy functional whose explicit form in terms of the electron density is still unknown. Hence, families of approximations to this quantity have been developed : semi-local approximations (local-density approximation (LDA), generalized-gradient approximations (GGAs) and meta-GGAs), hybrid approximations, and approximations depending on virtual orbitals (see, e.g., Ref. [115] for a recent review). This last family of approximations includes approaches combining semi-local density-functional approximations (DFAs) with Hartree-Fock (HF) exchange and second-order Møller-Plesset (MP2) correlation, either based on a range separation of the electron-electron interaction or a linear separation. In the range-separated hybrid (RSH) variant, the Coulomb electron-electron interaction  $w_{ee}(r_{12}) = 1/r_{12}$  is decomposed as [27, 96]

$$w_{ee}(r_{12}) = w_{ee}^{\text{lr},\mu}(r_{12}) + w_{ee}^{\text{sr},\mu}(r_{12}), \quad (4.1)$$

where  $w_{ee}^{\text{lr},\mu}(r_{12}) = \text{erf}(\mu r)/r_{12}$  is a long-range interaction (written with the error function erf) and  $w_{ee}^{\text{sr},\mu}(r_{12}) = \text{erfc}(\mu r)/r_{12}$  is the complementary short-range interaction (written

with the complementary error function  $\text{erfc}$ ), the decomposition being controlled by the parameter  $\mu$  ( $0 \leq \mu < \infty$ ). HF exchange and MP2 correlation can then be used for the long-range part of the energy, while a semi-local exchange-correlation DFA is used for the complementary short-range part, resulting in a method that is denoted by RSH+MP2 [30]. Among the main advantages of such an approach are the explicit description of van der Waals dispersion interactions via the long-range MP2 part (see, e.g., Ref. [32]) and the fast (exponential) convergence of the long-range MP2 correlation energy with respect to the size of the one-electron basis set [29]. On the other hand, the short-range exchange-correlation DFAs used still exhibit significant errors, such as self-interaction errors [116], limiting the accuracy for the calculations of atomization energies or non-covalent electrostatic interactions for example. Similarly, the double-hybrid (DH) variant [59] (see Ref. [117] for a review) for combining MP2 and a semi-local DFA can be considered as corresponding to a linear separation of the Coulomb electron-electron interaction [30]

$$w_{ee}(\mathbf{r}_{12}) = \lambda w_{ee}(\mathbf{r}_{12}) + (1 - \lambda)w_{ee}(\mathbf{r}_{12}), \quad (4.2)$$

where  $\lambda$  is a parameter ( $0 \leq \lambda \leq 1$ ). If HF exchange and MP2 correlation is used for the part of the energy associated with the interaction  $\lambda w_{ee}(\mathbf{r}_{12})$  and a semi-local exchange-correlation DFA is used for the complementary part, then a one-parameter version of the DH approximations is obtained [30]. One of the main advantages of the DH approximations is their quite efficient reduction of the self-interaction error [116] thanks to their large fraction of HF exchange ( $\lambda \approx 0.5$  or more). On the other hand, they inherit (a fraction of) the slow (polynomial) basis convergence of standard MP2 [63], and they are insufficiently accurate for the description of van der Waals dispersion interactions and need the addition of semi-empirical dispersion corrections [69]. In this work, we consider range-separated double-hybrid (RSDH)<sup>1</sup> approximations which combine the two above-mentioned approaches, based on the following decomposition of the Coulomb electron-electron interaction

$$w_{ee}(\mathbf{r}_{12}) = \left[ w_{ee}^{\text{lr},\mu}(\mathbf{r}_{12}) + \lambda w_{ee}^{\text{sr},\mu}(\mathbf{r}_{12}) \right] + (1 - \lambda)w_{ee}^{\text{sr},\mu}(\mathbf{r}_{12}), \quad (4.3)$$

---

1. The term range-separated double hybrid (RSDH) has already been used in Ref. 92 to designate the RSH+MP2 method and in Refs. [89] and [118] to generically designate several kinds of range-separated MP2/DFT hybrids. In the present work, we use RSDH to designate the specific scheme corresponding to Eq. 4.28.

where, again, the energy corresponding to the first part of the interaction (in square brackets) is calculated with HF exchange and MP2 correlation, and the complementary part is treated by a semi-local exchange-correlation DFA. The expected features of such an approach are the explicit description of van der Waals dispersion interactions through the long-range part, and reduced self-interaction errors in the short-range part (and thus improved calculations of properties such as atomization energies). The decomposition of Eq. (4.3) is in fact a special case of the decomposition used in the Coulomb-attenuating method (CAM) [119]

$$w_{ee}(\mathbf{r}_{12}) = \left[ (\alpha + \beta) w_{ee}^{\text{lr},\mu}(\mathbf{r}_{12}) + \alpha w_{ee}^{\text{sr},\mu}(\mathbf{r}_{12}) \right] + \left[ (1 - \alpha - \beta) w_{ee}^{\text{lr},\mu}(\mathbf{r}_{12}) + (1 - \alpha) w_{ee}^{\text{sr},\mu}(\mathbf{r}_{12}) \right], \quad (4.4)$$

with the parameters  $\alpha + \beta = 1$  and  $\alpha = \lambda$ . The choice  $\alpha + \beta = 1$  is known to be appropriate for Rydberg and charge-transfer excitation energies [120] and for reaction-barrier heights [121]. We also expect it to be appropriate for the description of long-range van der Waals dispersion interactions. It should be noted that the CAM decomposition has been introduced in Ref. [119] at the exchange level only, i.e. for combining HF exchange with a semi-local exchange DFA without modifying the semi-local correlation DFA. Only recently, Cornaton and Fromager [89] (see also Ref. [118]) pointed out the possibility of a CAM double-hybrid approximation but remarked that the inclusion of a fraction of short-range electron-electron interaction in the MP2 part would limit the basis convergence, and preferred to develop an alternative approach which uses only the perturbation expansion of a long-range interacting wave function. Despite the expected slower basis convergence of DH approximations based on the decomposition in Eq. (4.3) [or in Eq. (4.4)] in comparison to the RSH+MP2 method based on the decomposition in Eq. (4.1), we still believe it worthwhile to explore this kind of DH approximations in light of the above-mentioned expected advantages. In fact, we will show that the basis convergence of the RSDH approximations is relatively fast, owing to the inclusion of a modest fraction of short-range MP2 correlation. The decomposition in Eq. (4.3) has been used several times at the exchange level [24]. A few DH approximations including either long-range exchange or long-range correlation terms have been proposed. The  $\omega$ B97X-2 approximation [73] adds a full-range MP2 correlation term to a hybrid approximation including long-range HF exchange. The B2-P3LYP approximation [80] and the lrc-XYG3 approximation [88] add a long-range MP2 correlation correction to a standard DH approximation including full-range HF exchange. Only in Ref. [122] the decomposition in Eq. (4.3) was consistently used at the exchange and correlation level, combining a pair coupled-cluster doubles ap-

proximation with a semi-local exchange-correlation DFA, in the goal of describing static correlation. However, the formulation of the exact theory based on the decomposition in Eq. (4.3), as well as the performance of the MP2 and semi-local DFAs in this context, have not been explored. This is what we undertake in the present work. The paper is organized as follows. In Section 4.3, the theory underlying the RSDH approximations is presented, and approximations for the corresponding short-range correlation density functional are developed. Computational details are given in Section 4.6.5. In Section 4.7, we give and discuss the results, concerning the optimization of the parameters  $\mu$  and  $\lambda$  on small sets of atomization energies (AE6 set) and reaction barrier heights (BH6 set), the study of the basis convergence, and the tests on large sets of atomization energies (AE49 set), reaction barrier heights (DBH24 set), and weak intermolecular interactions (S22 set). Section 4.9 contains conclusions and future work prospects. Finally, the Appendix contains the derivation of the uniform coordinate scaling relation and the Coulomb/high-density and short-range/low-density limits of the short-range correlation density functional involved in this work. Unless otherwise specified, Hartree atomic units are tacitly assumed throughout this work.

### 4.3 Exact theory

The derivation of the RSDH density-functional theory starts from the universal density functional [37],

$$F[n] = \min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{W}_{ee} | \Psi \rangle, \quad (4.5)$$

where  $\hat{T}$  is the kinetic-energy operator,  $\hat{W}_{ee}$  the Coulomb electron-electron repulsion operator, and the minimization is done over normalized antisymmetric multideterminant wave functions  $\Psi$  giving a fixed density  $n$ . The universal density functional is then decomposed as

$$F[n] = F^{\mu,\lambda}[n] + \bar{E}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[n], \quad (4.6)$$

where  $F^{\mu,\lambda}[n]$  is defined as

$$F^{\mu,\lambda}[n] = \min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{W}_{ee}^{\text{lr},\mu} + \lambda \hat{W}_{ee}^{\text{sr},\mu} | \Psi \rangle. \quad (4.7)$$

In Eq. (4.7),  $\hat{W}_{ee}^{\text{lr},\mu}$  is the long-range electron-electron repulsion operator and  $\lambda \hat{W}_{ee}^{\text{sr},\mu}$  is the short-range electron-electron repulsion operator scaled by the constant  $\lambda$ , with expres-

sions :

$$\hat{W}_{ee}^{\text{lr},\mu} = \frac{1}{2} \iint w_{ee}^{\text{lr},\mu}(\mathbf{r}_{12}) \hat{n}_2(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2, \quad (4.8)$$

$$\hat{W}_{ee}^{\text{sr},\mu} = \frac{1}{2} \iint w_{ee}^{\text{sr},\mu}(\mathbf{r}_{12}) \hat{n}_2(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2, \quad (4.9)$$

where  $\hat{n}_2(\mathbf{r}_1, \mathbf{r}_2) = \hat{n}(\mathbf{r}_1)\hat{n}(\mathbf{r}_2) - \delta(\mathbf{r}_1 - \mathbf{r}_2)\hat{n}(\mathbf{r}_1)$  is the pair-density operator, written with the density operator  $\hat{n}(\mathbf{r})$ . Equation (4.6) defines the complement short-range Hartree-exchange-correlation density functional  $\bar{E}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[n]$  depending on the two parameters  $\mu$  and  $\lambda$ . It can itself be decomposed as

$$\bar{E}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[n] = E_{\text{H}}^{\text{sr},\mu,\lambda}[n] + \bar{E}_{\text{xc}}^{\text{sr},\mu,\lambda}[n], \quad (4.10)$$

where  $E_{\text{H}}^{\text{sr},\mu,\lambda}[n]$  is the short-range Hartree contribution,

$$E_{\text{H}}^{\text{sr},\mu,\lambda}[n] = (1 - \lambda) \times \frac{1}{2} \iint w_{ee}^{\text{sr},\mu}(\mathbf{r}_{12}) n(\mathbf{r}_1) n(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2, \quad (4.11)$$

and  $\bar{E}_{\text{xc}}^{\text{sr},\mu,\lambda}[n]$  is the short-range exchange-correlation contribution. The exact ground-state electronic energy of a N-electron system in the external nuclei-electron potential  $v_{\text{ne}}(\mathbf{r})$  can be expressed as

$$\begin{aligned} E &= \min_{\mathbf{n} \rightarrow \text{N}} \left\{ F[\mathbf{n}] + \int v_{\text{ne}}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} \right\} \\ &= \min_{\mathbf{n} \rightarrow \text{N}} \left\{ F^{\mu,\lambda}[\mathbf{n}] + \bar{E}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[\mathbf{n}] + \int v_{\text{ne}}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} \right\} \\ &= \min_{\Psi \rightarrow \text{N}} \left\{ \langle \Psi | \hat{T} + \hat{W}_{ee}^{\text{lr},\mu} + \lambda \hat{W}_{ee}^{\text{sr},\mu} + \hat{V}_{\text{ne}} | \Psi \rangle \right. \\ &\quad \left. + \bar{E}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[n_{\Psi}] \right\}, \end{aligned} \quad (4.12)$$

where  $\mathbf{n} \rightarrow \text{N}$  refers to N-representable densities,  $\Psi \rightarrow \text{N}$  refers to N-electron normalized antisymmetric multideterminant wave functions, and  $n_{\Psi}$  denotes the density coming from  $\Psi$ , i.e.  $n_{\Psi}(\mathbf{r}) = \langle \Psi | \hat{n}(\mathbf{r}) | \Psi \rangle$ . In Eq. (4.12), the last line was obtained by using the expression of  $F^{\mu,\lambda}[\mathbf{n}]$  in Eq. (4.7), introducing the nuclei-electron potential operator  $\hat{V}_{\text{ne}} = \int v_{\text{ne}}(\mathbf{r}) \hat{n}(\mathbf{r}) d\mathbf{r}$ , and recomposing the two-step minimization into a single one, i.e.  $\min_{\mathbf{n} \rightarrow \text{N}} \min_{\Psi \rightarrow \mathbf{n}} = \min_{\Psi \rightarrow \text{N}}$ . The minimizing wave function  $\Psi^{\mu,\lambda}$  in Eq. (4.12) satisfies

the corresponding Euler-Lagrange equation, leading to the Schrödinger-like equation

$$\left( \hat{T} + \hat{W}_{ee}^{\text{lr},\mu} + \lambda \hat{W}_{ee}^{\text{sr},\mu} + \hat{V}_{\text{ne}} + \hat{V}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[\mathbf{n}_{\Psi^{\mu,\lambda}}] \right) |\Psi^{\mu,\lambda}\rangle = \mathcal{E}^{\mu,\lambda} |\Psi^{\mu,\lambda}\rangle, \quad (4.13)$$

where  $\mathcal{E}^{\mu,\lambda}$  is the Lagrange multiplier associated with the normalization constraint of the wave function. In Eq. (4.13),  $\hat{V}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[\mathbf{n}] = \int v_{\text{Hxc}}^{\text{sr},\mu,\lambda}(\mathbf{r}) \hat{n}(\mathbf{r}) d\mathbf{r}$  is the complement short-range Hartree-exchange-correlation potential operator with  $v_{\text{Hxc}}^{\text{sr},\mu,\lambda}(\mathbf{r}) = \delta \bar{E}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[\mathbf{n}] / \delta n(\mathbf{r})$ . Equation (4.13) defines an effective Hamiltonian  $\hat{H}^{\mu,\lambda} = \hat{T} + \hat{V}_{\text{ne}} + \hat{W}_{ee}^{\text{lr},\mu} + \lambda \hat{W}_{ee}^{\text{sr},\mu} + \hat{V}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[\mathbf{n}_{\Psi^{\mu,\lambda}}]$  that must be solved iteratively for its ground-state multideterminant wave function  $\Psi^{\mu,\lambda}$  which gives the exact ground-state density and the exact ground-state energy via Eq. (4.12), independently of  $\mu$  and  $\lambda$ . We have therefore defined an exact formalism combining a wave-function calculation with a density functional. This formalism encompasses several important special cases :

- $\mu = 0$  and  $\lambda = 0$ . In Eq. (4.12), the electron-electron operator vanishes,  $\hat{W}_{ee}^{\text{lr},\mu=0} + 0 \times \hat{W}_{ee}^{\text{sr},\mu=0} = 0$ , and the density functional reduces to the KS Hartree-exchange-correlation density functional,  $\bar{E}_{\text{Hxc}}^{\text{sr},\mu=0,\lambda=0}[\mathbf{n}] = E_{\text{Hxc}}[\mathbf{n}]$ , so that we recover standard KS DFT

$$E = \min_{\Phi \rightarrow \mathbf{N}} \left\{ \langle \Phi | \hat{T} + \hat{V}_{\text{ne}} | \Phi \rangle + E_{\text{Hxc}}[\mathbf{n}_{\Phi}] \right\}, \quad (4.14)$$

where  $\Phi$  is a single-determinant wave function.

- $\mu \rightarrow \infty$  or  $\lambda = 1$ . In Eq. (4.12), the electron-electron operator reduces to the Coulomb interaction  $\hat{W}_{ee}^{\text{lr},\mu \rightarrow \infty} + \lambda \hat{W}_{ee}^{\text{sr},\mu \rightarrow \infty} = \hat{W}_{ee}$  or  $\hat{W}_{ee}^{\text{lr},\mu} + 1 \times \hat{W}_{ee}^{\text{sr},\mu} = \hat{W}_{ee}$ , and the density functional vanishes,  $\bar{E}_{\text{Hxc}}^{\text{sr},\mu \rightarrow \infty,\lambda}[\mathbf{n}] = 0$  or  $\bar{E}_{\text{Hxc}}^{\text{sr},\mu,\lambda=1}[\mathbf{n}] = 0$ , so that we recover standard wave-function theory

$$E = \min_{\Psi \rightarrow \mathbf{N}} \langle \Psi | \hat{T} + \hat{W}_{ee} + \hat{V}_{\text{ne}} | \Psi \rangle. \quad (4.15)$$

- $0 < \mu < \infty$  and  $\lambda = 0$ . In Eq. (4.12), the electron-electron operator reduces to the long-range interaction  $\hat{W}_{ee}^{\text{lr},\mu} + 0 \times \hat{W}_{ee}^{\text{sr},\mu} = \hat{W}_{ee}^{\text{lr},\mu}$ , and the density functional reduces to the usual short-range density functional,  $\bar{E}_{\text{Hxc}}^{\text{sr},\mu,\lambda=0}[\mathbf{n}] = \bar{E}_{\text{Hxc}}^{\text{sr},\mu}[\mathbf{n}]$ , so that we recover range-separated DFT [27, 96]

$$E = \min_{\Psi \rightarrow \mathbf{N}} \left\{ \langle \Psi | \hat{T} + \hat{W}_{ee}^{\text{lr},\mu} + \hat{V}_{\text{ne}} | \Psi \rangle + \bar{E}_{\text{Hxc}}^{\text{sr},\mu}[\mathbf{n}_{\Psi}] \right\}. \quad (4.16)$$

•  $\mu = 0$  and  $0 < \lambda < 1$ . In Eq. (4.12), the electron-electron operator reduces to the scaled Coulomb interaction  $\hat{W}_{ee}^{\text{lr},\mu=0} + \lambda \hat{W}_{ee}^{\text{sr},\mu=0} = \lambda \hat{W}_{ee}$ , and the density functional reduces to the  $\lambda$ -complement density functional,  $\bar{E}_{\text{Hxc}}^{\text{sr},\mu=0,\lambda}[\mathbf{n}] = \bar{E}_{\text{Hxc}}^\lambda[\mathbf{n}]$ , so that we recover the multideterminant extension of KS DFT based on the linear decomposition of the electron-electron interaction [82, 123]

$$E = \min_{\Psi \rightarrow \mathbf{N}} \left\{ \langle \Psi | \hat{T} + \lambda \hat{W}_{ee} + \hat{V}_{\text{ne}} | \Psi \rangle + \bar{E}_{\text{Hxc}}^\lambda[\mathbf{n}_\Psi] \right\}. \quad (4.17)$$

## 4.4 Single-determinant approximation

As a first step, we introduce a single-determinant approximation in Eq. (4.12),

$$E_0^{\mu,\lambda} = \min_{\Phi \rightarrow \mathbf{N}} \left\{ \langle \Phi | \hat{T} + \hat{W}_{ee}^{\text{lr},\mu} + \lambda \hat{W}_{ee}^{\text{sr},\mu} + \hat{V}_{\text{ne}} | \Phi \rangle + \bar{E}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[\mathbf{n}_\Phi] \right\}, \quad (4.18)$$

where the search is over  $\mathbf{N}$ -electron normalized single-determinant wave functions. The minimizing single determinant  $\Phi^{\mu,\lambda}$  is given by the HF- or KS-like equation

$$\left( \hat{T} + \hat{V}_{\text{ne}} + \hat{V}_{\text{Hx,HF}}^{\text{lr},\mu}[\Phi^{\mu,\lambda}] + \lambda \hat{V}_{\text{Hx,HF}}^{\text{sr},\mu}[\Phi^{\mu,\lambda}] + \hat{V}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[\mathbf{n}_{\Phi^{\mu,\lambda}}] \right) |\Phi^{\mu,\lambda}\rangle = \mathcal{E}_0^{\mu,\lambda} |\Phi^{\mu,\lambda}\rangle, \quad (4.19)$$

where  $\hat{V}_{\text{Hx,HF}}^{\text{lr},\mu}[\Phi^{\mu,\lambda}]$  and  $\hat{V}_{\text{Hx,HF}}^{\text{sr},\mu}[\Phi^{\mu,\lambda}]$  are the long-range and short-range HF potential operators constructed with the single determinant  $\Phi^{\mu,\lambda}$ , and  $\mathcal{E}_0^{\mu,\lambda}$  is the Lagrange multiplier associated with the normalization constraint. Equation (4.19) must be solved self-consistently for its single-determinant ground-state wave function  $\Phi^{\mu,\lambda}$ . Note that, due to the single-determinant approximation, the density  $\mathbf{n}_{\Phi^{\mu,\lambda}}$  is not the exact ground-state density and the energy in Eq. (4.18) is not the exact ground-state energy and depends on the parameters  $\mu$  and  $\lambda$ . It can be rewritten in the form

$$E_0^{\mu,\lambda} = \langle \Phi^{\mu,\lambda} | \hat{T} + \hat{V}_{\text{ne}} | \Phi^{\mu,\lambda} \rangle + E_{\text{H}}[\mathbf{n}_{\Phi^{\mu,\lambda}}] + E_{\text{x,HF}}^{\text{lr},\mu}[\Phi^{\mu,\lambda}] + \lambda E_{\text{x,HF}}^{\text{sr},\mu}[\Phi^{\mu,\lambda}] + \bar{E}_{\text{xc}}^{\text{sr},\mu,\lambda}[\mathbf{n}_{\Phi^{\mu,\lambda}}], \quad (4.20)$$

where  $E_{\text{H}}[\mathbf{n}] = (1/2) \iint w_{ee}(\mathbf{r}_{12}) \mathbf{n}(\mathbf{r}_1) \mathbf{n}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2$  is the standard Hartree energy with the Coulomb electron-electron interaction, and  $E_{\text{x,HF}}^{\text{lr},\mu}$  and  $E_{\text{x,HF}}^{\text{sr},\mu}$  are the long-range and

short-range HF exchange energies. For  $\mu = 0$  and  $\lambda = 0$ , we recover standard KS DFT, while for  $\mu \rightarrow \infty$  or  $\lambda = 1$  we recover standard HF theory. For intermediate values of  $\mu$  and  $\lambda$ , this scheme is very similar to the approximations of Refs. [24], except that the part of correlation associated with the interaction  $w_{ee}^{\text{lr},\mu}(r_{12}) + \lambda w_{ee}^{\text{sr},\mu}(r_{12})$  is missing in Eq. (4.20). The addition of this correlation is done in a second step with MP2 perturbation theory. where  $E_{\text{H}}[n] = (1/2) \iint w_{ee}(r_{12})n(\mathbf{r}_1)n(\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2$  is the standard Hartree energy with the Coulomb electron-electron interaction, and  $E_{x,\text{HF}}^{\text{lr},\mu}$  and  $E_{x,\text{HF}}^{\text{sr},\mu}$  are the long-range and short-range HF exchange energies. For  $\mu = 0$  and  $\lambda = 0$ , we recover standard KS DFT, while for  $\mu \rightarrow \infty$  or  $\lambda = 1$  we recover standard HF theory. For intermediate values of  $\mu$  and  $\lambda$ , this scheme is very similar to the approximations of Refs. [24], except that the part of correlation associated with the interaction  $w_{ee}^{\text{lr},\mu}(r_{12}) + \lambda w_{ee}^{\text{sr},\mu}(r_{12})$  is missing in Eq. (4.20). The addition of this correlation is done in a second step with MP2 perturbation theory.

## 4.5 Second-order Møller–Plesset perturbation theory

A rigorous non-linear Rayleigh-Schrödinger perturbation theory starting from the single-determinant reference of Section 4.4 can be developed, similarly to what was done for the RSH+MP2 method in Refs [30, 94, 95] and for the one-parameter DH approximations in Ref. [82]. This is done by introducing a perturbation strength parameter  $\varepsilon$  and defining the energy expression :

$$E^{\mu,\lambda,\varepsilon} = \min_{\Psi \rightarrow \text{N}} \left\{ \langle \Psi | \hat{T} + \hat{V}_{\text{ne}} + \hat{V}_{\text{Hx,HF}}^{\text{lr},\mu}[\Phi^{\mu,\lambda}] + \lambda \hat{V}_{\text{Hx,HF}}^{\text{sr},\mu}[\Phi^{\mu,\lambda}] + \varepsilon \hat{\mathcal{W}}^{\mu,\lambda} | \Psi \rangle + \bar{E}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[n_{\Psi}] \right\}, \quad (4.21)$$

where the search is over N-electron normalized antisymmetric multideterminant wave functions, and  $\hat{\mathcal{W}}^{\mu,\lambda}$  is a Møller–Plesset-type perturbation operator

$$\hat{\mathcal{W}}^{\mu,\lambda} = \hat{W}_{ee}^{\text{lr},\mu} + \lambda \hat{W}_{ee}^{\text{sr},\mu} - \hat{V}_{\text{Hx,HF}}^{\text{lr},\mu}[\Phi^{\mu,\lambda}] - \lambda \hat{V}_{\text{Hx,HF}}^{\text{sr},\mu}[\Phi^{\mu,\lambda}]. \quad (4.22)$$

The minimizing wave function  $\Psi^{\mu,\lambda,\varepsilon}$  in Eq. (4.21) is given by the corresponding Euler-Lagrange equation :

$$\left( \hat{T} + \hat{V}_{\text{ne}} + \hat{V}_{\text{Hx,HF}}^{\text{lr},\mu}[\Phi^{\mu,\lambda}] + \lambda \hat{V}_{\text{Hx,HF}}^{\text{sr},\mu}[\Phi^{\mu,\lambda}] + \varepsilon \hat{\mathcal{W}}^{\mu,\lambda} + \hat{V}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[n_{\Psi^{\mu,\lambda,\varepsilon}}] \right) | \Psi^{\mu,\lambda,\varepsilon} \rangle = \mathcal{E}^{\mu,\lambda,\varepsilon} | \Psi^{\mu,\lambda,\varepsilon} \rangle. \quad (4.23)$$

For  $\varepsilon = 0$ , Eq. (4.23) reduces to the single-determinant reference of Eq. (4.19), i.e.  $\Psi^{\mu,\lambda,\varepsilon=0} = \Phi^{\mu,\lambda}$  and  $\mathcal{E}^{\mu,\lambda,\varepsilon=0} = \mathcal{E}_0^{\mu,\lambda}$ . For  $\varepsilon = 1$ , Eq. (4.23) reduces to Eq. (4.13), i.e.  $\Psi^{\mu,\lambda,\varepsilon=1} = \Psi^{\mu,\lambda}$  and  $\mathcal{E}^{\mu,\lambda,\varepsilon=1} = \mathcal{E}^{\mu,\lambda}$ , and Eq. (4.21) reduces to Eq. (4.12), i.e. we recover the physical energy  $E^{\mu,\lambda,\varepsilon=1} = E$ , independently of  $\mu$  and  $\lambda$ . The perturbation theory is then obtained by expanding these quantities in  $\varepsilon$  around  $\varepsilon = 0$ :  $\mathcal{E}^{\mu,\lambda,\varepsilon} = \sum_{k=0}^{\infty} \varepsilon^k \mathcal{E}^{\mu,\lambda,(k)}$ ,  $\Psi^{\mu,\lambda,\varepsilon} = \sum_{k=0}^{\infty} \varepsilon^k \Psi^{\mu,\lambda,(k)}$ , and  $E^{\mu,\lambda,\varepsilon} = \sum_{k=0}^{\infty} \varepsilon^k E^{\mu,\lambda,(k)}$ . Following the same steps as in Ref. [30], we find the zeroth-order energy,

$$\begin{aligned} E^{\mu,\lambda,(0)} &= \langle \Phi^{\mu,\lambda} | \hat{T} + \hat{V}_{\text{ne}} + \hat{V}_{\text{Hx,HF}}^{\text{lr},\mu}[\Phi^{\mu,\lambda}] \\ &\quad + \lambda \hat{V}_{\text{Hx,HF}}^{\text{sr},\mu}[\Phi^{\mu,\lambda}] | \Phi^{\mu,\lambda} \rangle + \bar{E}_{\text{Hxc}}^{\text{sr},\mu,\lambda}[\mathfrak{n}_{\Phi^{\mu,\lambda}}], \end{aligned} \quad (4.24)$$

and the first-order energy correction,

$$E^{\mu,\lambda,(1)} = \langle \Phi^{\mu,\lambda} | \hat{\mathcal{W}}^{\mu,\lambda} | \Phi^{\mu,\lambda} \rangle, \quad (4.25)$$

so that the zeroth+first order energy gives back the energy of the single-determinant reference in Eq. (4.20),

$$E^{\mu,\lambda,(0)} + E^{\mu,\lambda,(1)} = E_0^{\mu,\lambda}. \quad (4.26)$$

The second-order energy correction involves only double-excited determinants  $\Phi_{ij \rightarrow ab}^{\mu,\lambda}$  (of energy  $\mathcal{E}_{0,ij \rightarrow ab}^{\mu,\lambda}$ ) and takes the form a MP2-like correlation energy, assuming a non-degenerate ground state in Eq. (4.19),

$$\begin{aligned} E^{\mu,\lambda,(2)} = E_{\text{c,MP}}^{\mu,\lambda} &= - \sum_{i < j}^{\text{occ}} \sum_{\text{a} < \text{b}}^{\text{vir}} \frac{|\langle \Phi_{ij \rightarrow ab}^{\mu,\lambda} | \hat{\mathcal{W}}^{\mu,\lambda} | \Phi^{\mu,\lambda} \rangle|^2}{\mathcal{E}_{0,ij \rightarrow ab}^{\mu,\lambda} - \mathcal{E}_0^{\mu,\lambda}} \\ &= - \sum_{i < j}^{\text{occ}} \sum_{\text{a} < \text{b}}^{\text{vir}} \frac{|\langle ij | \hat{w}_{\text{ee}}^{\text{lr},\mu} + \lambda \hat{w}_{\text{ee}}^{\text{sr},\mu} | \text{ab} \rangle - \langle ij | \hat{w}_{\text{ee}}^{\text{lr},\mu} + \lambda \hat{w}_{\text{ee}}^{\text{sr},\mu} | \text{ba} \rangle|^2}{\varepsilon_{\text{a}} + \varepsilon_{\text{b}} - \varepsilon_{\text{i}} - \varepsilon_{\text{j}}}, \end{aligned} \quad (4.27)$$

Also, contrary to the approach of Refs. [89, 118], the second-order energy correction does not involve any contribution from the second-order correction to the density. The total RSDH energy is finally :

$$E_{\text{RSDH}}^{\mu,\lambda} = E_0^{\mu,\lambda} + E_{\text{c,MP}}^{\mu,\lambda}. \quad (4.28)$$

It is instructive to decompose the correlation energy in Eq. (4.27) as

$$E_{c,MP}^{\mu,\lambda} = E_{c,MP}^{\text{lr},\mu} + \lambda E_{c,MP}^{\text{lr-sr},\mu} + \lambda^2 E_{c,MP}^{\text{sr},\mu}, \quad (4.29)$$

with a pure long-range contribution,

$$E_{c,MP}^{\text{lr},\mu} = - \sum_{i<j}^{\text{occ}} \sum_{a<b}^{\text{vir}} \frac{|\langle ij | \hat{w}_{ee}^{\text{lr},\mu} | ab \rangle - \langle ij | \hat{w}_{ee}^{\text{lr},\mu} | ba \rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j}, \quad (4.30)$$

a pure short-range contribution,

$$E_{c,MP}^{\text{sr},\mu} = - \sum_{i<j}^{\text{occ}} \sum_{a<b}^{\text{vir}} \frac{|\langle ij | \hat{w}_{ee}^{\text{sr},\mu} | ab \rangle - \langle ij | \hat{w}_{ee}^{\text{sr},\mu} | ba \rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j}, \quad (4.31)$$

and a mixed long-range/short-range contribution,

$$E_{c,MP}^{\text{lr-sr},\mu} = - \sum_{i<j}^{\text{occ}} \sum_{a<b}^{\text{vir}} \frac{(\langle ij | \hat{w}_{ee}^{\text{lr},\mu} | ab \rangle - \langle ij | \hat{w}_{ee}^{\text{lr},\mu} | ba \rangle) (\langle ab | \hat{w}_{ee}^{\text{sr},\mu} | ij \rangle - \langle ba | \hat{w}_{ee}^{\text{sr},\mu} | ij \rangle)}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} + \text{c.c.} \quad (4.32)$$

where c.c. stands for the complex conjugate. The exchange-correlation energy in the RSDH approximation is thus :

$$E_{xc,RSDH}^{\mu,\lambda} = E_{x,HF}^{\text{lr},\mu} + \lambda E_{x,HF}^{\text{sr},\mu} + E_{c,MP}^{\text{lr},\mu} + \lambda E_{c,MP}^{\text{lr-sr},\mu} + \lambda^2 E_{c,MP}^{\text{sr},\mu} + \bar{E}_{xc}^{\text{sr},\mu,\lambda}[\mathbf{n}]. \quad (4.33)$$

It remains to develop approximations for the complement short-range exchange-correlation density functional  $\bar{E}_{xc}^{\text{sr},\mu,\lambda}[\mathbf{n}]$ , which is done in Section 4.6.

## 4.6 Complement short-range exchange-correlation density functional

### 4.6.1 Decomposition into exchange and correlation

The complement short-range exchange-correlation density functional  $\bar{E}_{xc}^{sr,\mu,\lambda}[n]$  can be decomposed into exchange and correlation contributions,

$$\bar{E}_{xc}^{sr,\mu,\lambda}[n] = E_x^{sr,\mu,\lambda}[n] + \bar{E}_c^{sr,\mu,\lambda}[n], \quad (4.34)$$

where the exchange part is defined with the KS single determinant  $\Phi[n]$  and is linear with respect to  $\lambda$ ,

$$\begin{aligned} E_x^{sr,\mu,\lambda}[n] &= \langle \Phi[n] | (1-\lambda)\hat{W}_{ee}^{sr,\mu} | \Phi[n] \rangle - E_H^{sr,\mu,\lambda}[n] \\ &= (1-\lambda)E_x^{sr,\mu}[n], \end{aligned} \quad (4.35)$$

The complement short-range exchange-correlation density functional  $\bar{E}_{xc}^{sr,\mu,\lambda}[n]$  can be decomposed into exchange and correlation contributions,

$$\bar{E}_{xc}^{sr,\mu,\lambda}[n] = E_x^{sr,\mu,\lambda}[n] + \bar{E}_c^{sr,\mu,\lambda}[n], \quad (4.36)$$

where the exchange part is defined with the KS single determinant  $\Phi[n]$  and is linear with respect to  $\lambda$ ,

$$\begin{aligned} E_x^{sr,\mu,\lambda}[n] &= \langle \Phi[n] | (1-\lambda)\hat{W}_{ee}^{sr,\mu} | \Phi[n] \rangle - E_H^{sr,\mu,\lambda}[n] \\ &= (1-\lambda)E_x^{sr,\mu}[n], \end{aligned} \quad (4.37)$$

where  $E_x^{sr,\mu}[n] = E_x^{sr,\mu,\lambda=0}[n]$  is the usual short-range exchange density functional, as already introduced, e.g., in Ref. [27]). Several (semi-)local approximations have been proposed for  $E_x^{sr,\mu}[n]$  (see, e.g., Refs. [96, 97, 100, 124–129]). By contrast, the complement short-range correlation density functional  $\bar{E}_c^{sr,\mu,\lambda}[n]$  cannot be exactly expressed in terms of the short-range correlation density functional  $\bar{E}_c^{sr,\mu}[n] = \bar{E}_c^{sr,\mu,\lambda=0}[n]$  of Ref. [27] for which several (semi-)local approximations have been proposed [27, 97, 100, 127–130]. Note that in the approach of Ref. [89] the complement density functional was defined using the pure long-range interacting wave function  $\Psi^\mu = \Psi^{\mu,\lambda=1}$  and it was possible, using uniform coordinate scaling relations, to find an exact expression for it in terms of previously studied density functionals. This is not the case in the present approach because

the complement density functional is defined using the wave function  $\Psi^{\mu,\lambda}$  obtained with both long-range and short-range interactions. As explained in the Appendix, uniform coordinate scaling relations do not allow one to obtain an exact expression for  $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$  in terms of previously studied density functionals. Therefore, the difficulty lies in developing approximations for  $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$ . For this, we first give the exact expression of  $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$  in the Coulomb limit  $\mu \rightarrow 0$  (and the related high-density limit) and in the short-range limit  $\mu \rightarrow \infty$  (and the related low-density limit).

#### 4.6.2 Expression of $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$ in the Coulomb limit $\mu \rightarrow 0$ and in the high-density limit

The complement short-range correlation density functional  $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$  can be written as

$$\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}] = E_c[\mathbf{n}] - E_c^{\mu,\lambda}[\mathbf{n}], \quad (4.38)$$

where  $E_c[\mathbf{n}]$  is the standard KS correlation density functional and  $E_c^{\mu,\lambda}[\mathbf{n}]$  is the correlation density functional associated with the interaction  $w_{ee}^{\text{lr},\mu}(r_{12}) + \lambda w_{ee}^{\text{sr},\mu}(r_{12})$ . For  $\mu = 0$ , the density functional  $E_c^{\mu=0,\lambda}[\mathbf{n}] = E_c^\lambda[\mathbf{n}]$  corresponds to the correlation functional associated with the scaled Coulomb interaction  $\lambda w_{ee}(r_{12})$ , which can be exactly expressed as  $E_c^\lambda[\mathbf{n}] = \lambda^2 E_c[\mathbf{n}_{1/\lambda}]$  [92, 131] where  $\mathbf{n}_{1/\lambda}(\mathbf{r}) = (1/\lambda^3)\mathbf{n}(\mathbf{r}/\lambda)$  is the density with coordinates uniformly scaled by  $1/\lambda$ . Therefore, for  $\mu = 0$ , the complement short-range correlation density functional is

$$\bar{E}_c^{\text{sr},\mu=0,\lambda}[\mathbf{n}] = E_c[\mathbf{n}] - \lambda^2 E_c[\mathbf{n}_{1/\lambda}], \quad (4.39)$$

which is the correlation functional used in the density-scaled one-parameter double-hybrid (DS1DH) scheme of Sharkas *et al.* [82]. For a KS system with a non-degenerate ground state, we have in the  $\lambda \rightarrow 0$  limit :  $E_c[\mathbf{n}_{1/\lambda}] = E_c^{\text{GL2}}[\mathbf{n}] + \mathcal{O}(\lambda)$  where  $E_c^{\text{GL2}}[\mathbf{n}]$  is the second-order Görling–Levy (GL2) correlation energy [60]. Therefore, in this case,  $\bar{E}_c^{\text{sr},\mu=0,\lambda}[\mathbf{n}]$  has a quadratic dependence in  $\lambda$  near  $\lambda = 0$ . In practice with GGA functionals, it has been found that the density scaling in Eq. (4.39) can sometimes be advantageously neglected, i.e.  $E_c[\mathbf{n}_{1/\lambda}] \approx E_c[\mathbf{n}]$  [82, 123], giving

$$\bar{E}_c^{\text{sr},\mu=0,\lambda}[\mathbf{n}] \approx (1 - \lambda^2)E_c[\mathbf{n}]. \quad (4.40)$$

Even if we do not plan to apply the RSDH scheme with  $\mu = 0$ , the condition in Eq. (4.39) is in fact relevant for an arbitrary value of  $\mu$  in the high-density limit, i.e.  $n_\gamma(\mathbf{r}) = \gamma^3 n(\gamma\mathbf{r})$  with  $\gamma \rightarrow \infty$ , since in this limit the short-range interaction becomes equivalent to the Coulomb interaction in the complement short-range correlation density functional :  $\lim_{\gamma \rightarrow \infty} \bar{E}_c^{\text{sr},\mu,\lambda}[n_\gamma] = \lim_{\gamma \rightarrow \infty} \bar{E}_c^{\text{sr},\mu=0,\lambda}[n_\gamma]$  (see Appendix). In fact, for a KS system with a non-degenerate ground state, the approximate condition in Eq. (4.40) is sufficient to recover the exact high-density limit for an arbitrary value of  $\mu$  which is (see Appendix)

$$\lim_{\gamma \rightarrow \infty} \bar{E}_c^{\text{sr},\mu,\lambda}[n_\gamma] = (1 - \lambda^2) E_c^{\text{GL2}}[n]. \quad (4.41)$$

### 4.6.3 Expression of $\bar{E}_c^{\text{sr},\mu,\lambda}[n]$ in the short-range limit $\mu \rightarrow \infty$ and the low-density limit

The leading term in the asymptotic expansion of  $\bar{E}_c^{\text{sr},\mu,\lambda}[n]$  as  $\mu \rightarrow \infty$  is (see Appendix)

$$\begin{aligned} \bar{E}_c^{\text{sr},\mu,\lambda}[n] &= (1 - \lambda) \frac{\pi}{2\mu^2} \int n_{2,c}[n](\mathbf{r}, \mathbf{r}) d\mathbf{r} \\ &+ \mathcal{O}\left(\frac{1}{\mu^3}\right), \end{aligned} \quad (4.42)$$

where  $n_{2,c}[n](\mathbf{r}, \mathbf{r})$  is the correlation part of the on-top pair density for the Coulomb interaction. We thus see that, for  $\mu \rightarrow \infty$ ,  $\bar{E}_c^{\text{sr},\mu,\lambda}[n]$  is linear with respect to  $\lambda$ . In fact, since the asymptotic expansion of the usual short-range correlation functional is  $\bar{E}_c^{\text{sr},\mu}[n] = \pi/(2\mu^2) \int n_{2,c}[n](\mathbf{r}, \mathbf{r}) d\mathbf{r} + \mathcal{O}(1/\mu^3)$  [27], we can write for  $\mu \rightarrow \infty$ ,

$$\bar{E}_c^{\text{sr},\mu,\lambda}[n] = (1 - \lambda) \bar{E}_c^{\text{sr},\mu}[n] + \mathcal{O}\left(\frac{1}{\mu^3}\right). \quad (4.43)$$

The low-density limit, i.e.  $n_\gamma(\mathbf{r}) = \gamma^3 n(\gamma\mathbf{r})$  with  $\gamma \rightarrow 0$ , is closely related to the limit  $\mu \rightarrow \infty$  (see Appendix)

$$\begin{aligned} \bar{E}_c^{\text{sr},\mu,\lambda}[n_\gamma] &\underset{\gamma \rightarrow 0}{\sim} \frac{\gamma^3(1 - \lambda)\pi}{2\mu^2} \int n_{2,c}^{1/\gamma}[n](\mathbf{r}, \mathbf{r}) d\mathbf{r} \\ &\underset{\gamma \rightarrow 0}{\sim} \frac{\gamma^3(1 - \lambda)\pi}{4\mu^2} \int [-n(\mathbf{r})^2 + m(\mathbf{r})^2] d\mathbf{r}, \end{aligned} \quad (4.44)$$

in which appears  $n_{2,c}^{1/\gamma}[n](\mathbf{r}, \mathbf{r})$ , the on-top pair density for the scaled Coulomb interaction  $(1/\gamma)w_{ee}(r_{12})$ , and its strong-interaction limit  $\lim_{\gamma \rightarrow 0} n_{2,c}^{1/\gamma}[n](\mathbf{r}, \mathbf{r}) = -n(\mathbf{r})^2/2 +$

$m(\mathbf{r})^2/2$  [132] where  $m(\mathbf{r})$  is the spin magnetization. Thus, in the low-density limit, contrary to the usual KS correlation functional  $E_c[n]$  which goes to zero linearly in  $\gamma$  [92], times the complicated nonlocal strictly-correlated electron functional [133], the complement short-range correlation functional  $\bar{E}_c^{\text{sr},\mu,\lambda}[n]$  goes to zero like  $\gamma^3$  and becomes a simple local functional of  $n(\mathbf{r})$  and  $m(\mathbf{r})$ . We now propose several simple approximations for  $\bar{E}_c^{\text{sr},\mu,\lambda}[n]$ . On the one hand, Eq. (4.40) suggests the approximation

$$\bar{E}_{c,\text{approx1}}^{\text{sr},\mu,\lambda}[n] = (1 - \lambda^2)\bar{E}_c^{\text{sr},\mu}[n], \quad (4.45)$$

which is correctly quadratic in  $\lambda$  at  $\mu = 0$  but is not linear in  $\lambda$  for  $\mu \rightarrow \infty$ . On the other hand, Eq. (4.43) suggests the approximation

$$\bar{E}_{c,\text{approx2}}^{\text{sr},\mu,\lambda}[n] = (1 - \lambda)\bar{E}_c^{\text{sr},\mu}[n], \quad (4.46)$$

which is correctly linear in  $\lambda$  for  $\mu \rightarrow \infty$  but not quadratic in  $\lambda$  at  $\mu = 0$ . However, it is possible to impose simultaneously the two limiting behaviors for  $\mu = 0$  and  $\mu \rightarrow \infty$  with the following approximation

$$\bar{E}_{c,\text{approx3}}^{\text{sr},\mu,\lambda}[n] = \bar{E}_c^{\text{sr},\mu}[n] - \lambda^2 \bar{E}_c^{\text{sr},\mu\sqrt{\lambda}}[n], \quad (4.47)$$

which reduces to Eq. (4.40) for  $\mu = 0$  and satisfies Eq. (4.42) for  $\mu \rightarrow \infty$ . Another possibility, proposed in Ref. [122], is

$$\bar{E}_{c,\text{approx4}}^{\text{sr},\mu,\lambda}[n] = \bar{E}_c^{\text{sr},\mu}[n] - \lambda^2 \bar{E}_c^{\text{sr},\mu/\lambda}[n_{1/\lambda}], \quad (4.48)$$

which correctly reduces to Eq. (4.39) for  $\mu = 0$ . For  $\mu \rightarrow \infty$ , its asymptotic expansion is

$$\begin{aligned} \bar{E}_{c,\text{approx4}}^{\text{sr},\mu,\lambda}[n] &= \frac{\pi}{2\mu^2} \int n_{2,c}[n](\mathbf{r}, \mathbf{r}) d\mathbf{r} \\ &\quad - \lambda^4 \frac{\pi}{2\mu^2} \int n_{2,c}[n_{1/\lambda}](\mathbf{r}, \mathbf{r}) d\mathbf{r} + \mathcal{O}\left(\frac{1}{\mu^3}\right), \end{aligned} \quad (4.49)$$

i.e. it does not satisfy Eq. (4.42). Contrary to what was suggested in Ref. [122], Eq. (4.48) is not exact but only an approximation. However, using the scaling relation on the system-averaged on-top pair density [132]

$$\int n_{2,c}[n_\gamma](\mathbf{r}, \mathbf{r}) d\mathbf{r} = \gamma^3 \int n_{2,c}^{1/\gamma}[n](\mathbf{r}, \mathbf{r}) d\mathbf{r}, \quad (4.50)$$

it can be seen that, in the low-density limit  $\gamma \rightarrow 0$ , Eq. (4.49) correctly reduces to Eq. (4.44). In Ref. [122], the authors propose to neglect the scaling of the density in Eq. (4.48) leading to

$$\bar{E}_{c,\text{approx5}}^{\text{sr},\mu,\lambda}[\mathbf{n}] = \bar{E}_c^{\text{sr},\mu}[\mathbf{n}] - \lambda^2 \bar{E}_c^{\text{sr},\mu/\lambda}[\mathbf{n}], \quad (4.51)$$

which reduces to Eq. (4.40) for  $\mu = 0$ , but which has also a wrong  $\lambda$ -dependence for large  $\mu$

$$\begin{aligned} \bar{E}_{c,\text{approx5}}^{\text{sr},\mu,\lambda}[\mathbf{n}] &= (1 - \lambda^4) \frac{\pi}{2\mu^2} \int n_{2,c}[\mathbf{n}](\mathbf{r}, \mathbf{r}) d\mathbf{r} \\ &\quad + \mathcal{O}\left(\frac{1}{\mu^3}\right), \end{aligned} \quad (4.52)$$

and does not anymore satisfy the low-density limit. Another strategy is to start from the decomposition of the MP2-like correlation energy in Eq. (4.29) which suggests the following approximation for the complement short-range correlation functional

$$\begin{aligned} \bar{E}_{c,\text{approx6}}^{\text{sr},\mu,\lambda}[\mathbf{n}] &= (1 - \lambda) E_c^{\text{lr-sr},\mu}[\mathbf{n}] \\ &\quad + (1 - \lambda^2) E_c^{\text{sr},\mu}[\mathbf{n}], \end{aligned} \quad (4.53)$$

where  $E_c^{\text{lr-sr},\mu}[\mathbf{n}] = \bar{E}_c^{\text{sr},\mu}[\mathbf{n}] - E_c^{\text{sr},\mu}[\mathbf{n}]$  is the mixed long-range/short-range correlation functional [134, 135] and  $E_c^{\text{sr},\mu}[\mathbf{n}]$  is the pure short-range correlation functional associated with the short-range interaction  $w_{ee}^{\text{sr},\mu}(r_{12})$  [134, 135]. An LDA functional has been constructed for  $E_c^{\text{sr},\mu}[\mathbf{n}]$  [136]. Since  $E_c^{\text{lr-sr},\mu=0}[\mathbf{n}] = 0$  and  $E_c^{\text{sr},\mu=0}[\mathbf{n}] = E_c[\mathbf{n}]$ , the approximation in Eq. (4.53) reduces to Eq. (4.40) for  $\mu = 0$ . For  $\mu \rightarrow \infty$ , since  $E_c^{\text{sr},\mu}[\mathbf{n}]$  decays faster than  $1/\mu^2$ , i.e.  $E_c^{\text{sr},\mu}[\mathbf{n}] = \mathcal{O}(1/\mu^3)$  [136],  $E_c^{\text{lr-sr},\mu}[\mathbf{n}]$  and  $\bar{E}_c^{\text{sr},\mu}[\mathbf{n}]$  have the same leading term in the large- $\mu$  expansion, i.e.  $E_c^{\text{lr-sr},\mu}[\mathbf{n}] = \bar{E}_c^{\text{sr},\mu}[\mathbf{n}] + \mathcal{O}(1/\mu^3)$ , and thus the approximation in Eq. (4.53) satisfies Eq. (4.42) or (4.43). One can also enforce the exact condition at  $\mu = 0$ , Eq. (4.40), by introducing a scaling of the density

$$\begin{aligned} \bar{E}_{c,\text{approx7}}^{\text{sr},\mu,\lambda}[\mathbf{n}] &= (1 - \lambda) E_c^{\text{lr-sr},\mu}[\mathbf{n}] + E_c^{\text{sr},\mu}[\mathbf{n}] \\ &\quad - \lambda^2 E_c^{\text{sr},\mu/\lambda}[\mathbf{n}_1/\lambda]. \end{aligned} \quad (4.54)$$

#### 4.6.4 Assessment of the approximations for $\bar{E}_c^{\text{sr},\mu,\lambda}[\text{n}]$ on the uniform-electron gas

We now test the approximations for the complement short-range correlation functional  $\bar{E}_c^{\text{sr},\mu,\lambda}[\text{n}]$  introduced in Sec. 4.6.3 on the spin-unpolarized uniform-electron gas. As a reference, for several values of the Wigner-Seitz radius  $r_s = (3/(4\pi\text{n}))^{1/3}$  and the parameters  $\mu$  and  $\lambda$ , we have calculated the complement short-range correlation energy per particle as

$$\bar{\varepsilon}_{c,\text{unif}}^{\text{sr},\mu,\lambda}(r_s) = \varepsilon_{c,\text{unif}}(r_s) - \varepsilon_{c,\text{unif}}^{\mu,\lambda}(r_s), \quad (4.55)$$

where  $\varepsilon_{c,\text{unif}}(r_s)$  is the correlation energy per particle of the uniform-electron gas with the Coulomb electron-electron  $w_{ee}(r_{12})$  and  $\varepsilon_{c,\text{unif}}^{\mu,\lambda}(r_s)$  is the correlation energy per particle of an uniform-electron gas with the modified electron-electron  $w_{ee}^{\text{r},\mu}(r_{12}) + \lambda w_{ee}^{\text{sr},\mu}(r_{12})$ . We used what is known today as the direct random-phase approximation + second-order screened exchange (dRPA+SOSEX) method (an approximation to coupled-cluster doubles) [31, 137] as introduced for the uniform-electron gas by Freeman [138] and extended for modified electron-electron interactions in Refs. [96, 127], and which is known to give reasonably accurate correlation energies per particle of the spin-unpolarized electron gas (error less than 1 millihartree for  $r_s < 10$ ). We note that these calculations would allow us to construct a complement short-range LDA correlation functional, but we refrain from doing that since we prefer to avoid having to do a complicated fit of  $\bar{\varepsilon}_c^{\text{sr},\mu,\lambda}(r_s)$  with respect to  $r_s$ ,  $\mu$ , and  $\lambda$ . Moreover, this would only give a spin-independent LDA functional. We thus use these uniform-electron gas calculations only to test the approximations of Sec. 4.6.3.

For several values of  $r_s$ ,  $\mu$ , and  $\lambda$ , we have calculated the complement short-range correlation energy per particle corresponding to the approximations 1 to 7 using the LDA approximation for  $\bar{E}_c^{\text{sr},\mu}[\text{n}]$  from Ref. [130] (for approximations 1 to 7), as well as the LDA approximation for  $E_c^{\text{sr},\mu}[\text{n}]$  from Ref. [136] (for approximations 6 and 7), and the errors with respect to the dRPA+SOSEX results are reported in Fig. 4.1. Note that the accuracy of the dRPA+SOSEX reference decreases as  $r_s$  increases, the error being of the order of 1 millihartree for  $r_s = 10$ , which explains why the curves on the third graph of Fig. 4.1 appear shifted with respect to zero at large  $r_s$ .

By construction, all the approximations become exact for  $\lambda = 0$  (and trivially for  $\lambda = 1$  or in the  $\mu \rightarrow \infty$  limit since the complement short-range correlation energy goes to zero in these cases). For intermediate values of  $\lambda$  and finite values of  $\mu$ , all the approximations,

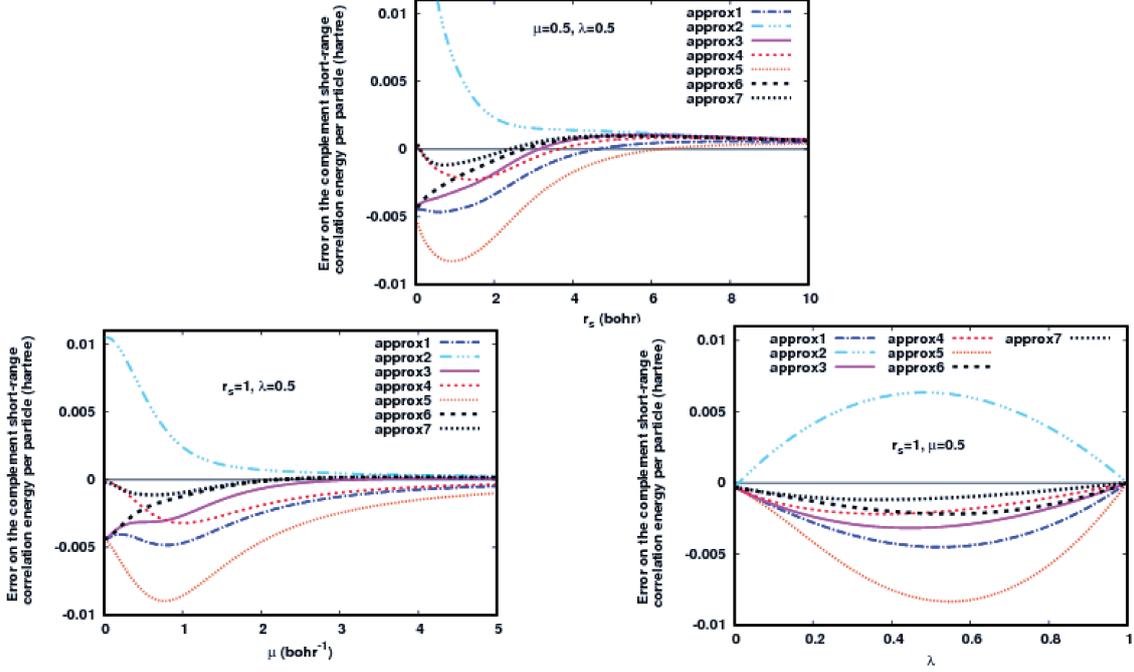


FIGURE 4.1 – Error on the complement short-range correlation energy per particle  $\bar{\varepsilon}_{c,\text{unif}}^{\text{SR},\mu,\lambda}(r_s)$  of the uniform-electron gas obtained with approximations 1 to 7 of Sec. 4.6.3 with respect to the DRPA+SOSEX results.

except approximation 2, tend to give too negative correlation energies. As it could have been expected, approximation 2, which is the only one incorrectly linear in  $\lambda$  at  $\mu = 0$ , gives quite a large error (of the order of 0.01 hartree or more) for small  $\mu$ , intermediate  $\lambda$ , and small  $r_s$  (it in fact diverges in the high-density limit  $r_s \rightarrow 0$ ), but the error goes rapidly to zero as  $\mu$  increases, reflecting the fact that this approximation has the correct leading term of the asymptotic expansion for  $\mu \rightarrow \infty$ . On the contrary, approximation 1, being quadratic in  $\lambda$ , gives a smaller error (less than 0.005 hartree) for small  $\mu$  but the error goes slower to zero as  $\mu$  increases. Approximation 3 combines the advantages of approximations 1 and 2 : it gives a small error for small  $\mu$  which goes rapidly to zero as  $\mu$  increases. Approximation 4, which contains the scaling of the density, is exact for  $\mu = 0$ , and gives a small error (at most about 0.003 hartree) for intermediate values of  $\mu$ , but the error does not go rapidly to zero as  $\mu$  increases. Again, this reflects the fact that this approximation does not give the correct leading term of the asymptotic expansion for  $\mu \rightarrow \infty$  for arbitrary  $\lambda$  and  $r_s$ . This confirms that Eq. (4.48) does not give the exact complement short-range correlation functional, contrary to what was thought in Ref. [122]. A nice feature however of approximation 4 is that it becomes exact in the high-density limit  $r_s \rightarrow 0$  of the uniform-electron gas (the scaling of the density at  $\mu = 0$  is

needed to obtain the correct high-density limit in this zero-gap system). Approximation 5, obtained from approximation 4 by neglecting the scaling of the density in the correlation functional, and used in Ref. [122], gives quite large errors for the uniform-electron gas, approaching 0.01 hartree. Approximations 6 and 7 are quite good. They both have the correct leading term of the asymptotic expansion for  $\mu \rightarrow \infty$ , but approximation 7 has the additional advantage of having also the correct  $\mu \rightarrow 0$  or  $r_s \rightarrow 0$  limit. Approximation 7 is our best approximation, with a maximal error of about 1 millihartree. Unfortunately, approximations 6 and 7 involve the pure short-range correlation functional  $E_c^{\text{sr},\mu}[\mathbf{n}]$ , for which we currently have only a spin-unpolarized LDA approximation [136]. For this reason, we do not consider these approximations in the following for molecular calculations. We will limit ourselves to approximations 1 to 5 which only involve the complement short-range correlation functional  $\bar{E}_c^{\text{sr},\mu}[\mathbf{n}]$ , for which we have spin-dependent GGAs [27, 97, 100, 128, 129].

#### 4.6.5 Computational details

The RSDH scheme has been implemented in a development version of the MOLPRO 2015 program [139]. The calculation is done in two steps : first a self-consistent-field calculation is performed according to Eqs. (4.18)-(4.20), and then the MP2-like correlation energy in Eq. (4.27) is evaluated with the previously calculated orbitals. The  $\lambda$ -dependent complement short-range exchange functional is calculated according to Eq. (4.35) and the approximations 1 to 5 [see Eqs. (4.45)-(4.51)] have been implemented for the complement short-range correlation functional, using the short-range Perdew-Becke-Ernzerhof (PBE) exchange and correlation functionals of Ref. [128] for  $E_x^{\text{sr},\mu}[\mathbf{n}]$  and  $\bar{E}_c^{\text{sr},\mu}[\mathbf{n}]$ . The RSDH scheme was applied on the AE6 and BH6 sets [140], as a first assessment of the approximations on molecular systems and in order to determine the optimal parameters  $\mu$  and  $\lambda$ . The AE6 set is a small representative benchmark of six atomization energies consisting of  $\text{SiH}_4$ ,  $\text{S}_2$ ,  $\text{SiO}$ ,  $\text{C}_3\text{H}_4$  (propyne),  $\text{C}_2\text{H}_2\text{O}_2$  (glyoxal), and  $\text{C}_4\text{H}_8$  (cyclobutane). The BH6 set is a small representative benchmark of forward and reverse hydrogen transfer barrier heights of three reactions,  $\text{OH} + \text{CH}_4 \rightarrow \text{CH}_3 + \text{H}_2\text{O}$ ,  $\text{H} + \text{OH} \rightarrow \text{O} + \text{H}_2$ , and  $\text{H} + \text{H}_2\text{S} \rightarrow \text{HS} + \text{H}_2$ . All the calculations for the AE6 and BH6 sets were performed with the Dunning cc-pVQZ basis set [141] at the geometries optimized by quadratic configuration interaction with single and double excitations with the modified Gaussian-3 basis set (QCISD/MG3)<sup>2</sup>. The reference values for the atomization energies and barrier heights are the non-relativistic FC-CCSD(T)/cc-pVQZ-F12 values of Refs. [87, 142]. For each ap-

---

2. KalTou-JJJ-XX-note1

proximation, we have first varied  $\mu$  and  $\lambda$  between 0 and 1 by steps of 0.1 to optimize the parameters on each set. We have then refined the search by steps of 0.02 to find the common optimal parameters on the two sets combined. The RSDH scheme was then tested on the AE49 set of 49 atomization energies [143] (consisting of the G2-1 set [144,145] stripped of the six molecules containing Li, Be, and Na<sup>3</sup>) and on the DBH24/08 set [147,148] of 24 forward and reverse reaction barrier heights. These calculations were performed with the cc-pVQZ basis set, with MP2(full)/6-31G\* geometries for the AE49 set, and with the aug-cc-pVQZ basis set [149] with QCISD/MG3 geometries for the DBH24/08 set. The reference values for the AE49 set are the non-relativistic FC-CCSD(T)/cc-pVQZ-F12 values of Ref. [142], and the reference values for the DBH24/08 set are the zero-point exclusive values from Ref. [148]. Finally, the RSDH scheme was tested on the S22 set of 22 weakly interacting molecular complexes [150]. These calculations were performed with the aug-cc-pVDZ basis set and the counterpoise correction, using the geometries from Ref. [150] and the complete-basis-set (CBS)-extrapolated CCSD(T) reference interaction energies from Ref. [151]. The local MP2 approach [152] is used on the largest systems in the S22 set. Core electrons are kept frozen in all our MP2 calculations. Spin-restricted calculations are performed for all the closed-shell systems, and spin-unrestricted calculations for all the open-shell systems. As statistical measures of goodness of the different methods, we compute mean absolute errors (MAEs), mean errors (MEs), root mean square deviations (RMSDs), mean absolute percentage errors (MA%E), and maximal and minimal errors.

---

3. We use the AE49 set of Fast et al. [143] to facilitate comparisons with previous related studies [82,146],<sup>93</sup> where this set was used. We do not expect any special difficulties with the six molecules that have been dropped from the G2-1 set.

## 4.7 Results and discussion

### 4.7.1 Optimization of the parameters on the AE6 and BH6 sets

We start by applying the RSDH scheme on the small AE6 and BH6 sets, and determining optimal values for the parameters  $\mu$  and  $\lambda$ . Figure 4.2 shows the MAEs for these two sets obtained with approximations 1 to 5 of Sec. 4.6.3 as a function of  $\lambda$  for  $\mu = 0.5$  and  $\mu = 0.6$ . We choose to show plots for only these two values of  $\mu$ , since they are close to the optimal value of  $\mu$  for RSH+MP2 [82,146] and also for RSDH with all the approximations except approximation 2. This last approximation is anyhow of little value for thermochemistry since it gives large MAEs on the AE6 set for intermediate values of  $\lambda$ , which must be related to the incorrect linear dependence in  $\lambda$  of this approximation in the limit  $\mu \rightarrow 0$  or the high-density limit. We thus only discuss next the other four approximations. Let

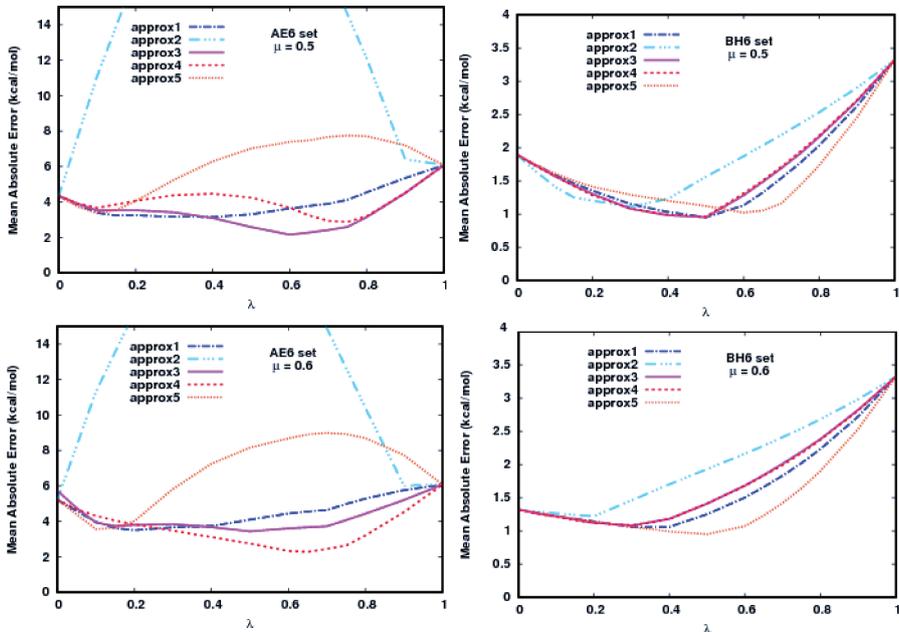


FIGURE 4.2 – MAEs for the AE6 and BH6 sets obtained with the RSDH scheme using approximations 1 to 5 of Sec. 4.6.3 (with the short-range PBE exchange-correlation functional of Ref. [128]) as a function of  $\lambda$  for  $\mu = 0.5$  and  $\mu = 0.6$ . The basis set used is cc-pVQZ.

us start by analyzing the results for the AE6 set. For the approximations 1, 3, 4, and 5, we can always find an intermediate value of  $\lambda$  giving a smaller MAE than the two limiting cases  $\lambda = 0$  (corresponding to RSH+MP2) and  $\lambda = 1$  (corresponding to standard MP2). Among these four approximations, the approximations 1 and 5 are the least effective to reduce the MAE in comparison to RSH+MP2 and MP2, which may be connected to

the fact that these two approximations are both incorrect in the low-density limit. The approximations 3 and 4, which become identical in the high- and low-density limits (for systems with non-zero gaps), are the two best approximations, giving minimal MAEs of 2.2 and 2.3 kcal/mol, respectively, at the optimal parameter values  $(\mu, \lambda) = (0.5, 0.6)$  and  $(0.6, 0.65)$ , respectively. Let us consider now the results for the BH6 set. Each MAE curve displays a marked minimum at an intermediate value of  $\lambda$ , at which the corresponding approximation is globally more accurate than both RSH+MP2 and MP2. All the approximations perform rather similarly, giving minimal MAEs of about 1 kcal/mol. In fact, for  $\mu = 0.5$  and  $\mu = 0.6$ , the approximations 3 and 4 give essentially identical MAEs for all  $\lambda$ . The optimal parameter values for these two approximations are  $(\mu, \lambda) = (0.5, 0.5)$ , i.e. relatively close to the optimal values found for the AE6 set.

For each of our two best approximations 3 and 4, we also determine optimal values of  $\mu$  and  $\lambda$  that minimize the total MAE of the combined AE6 + BH6 set, and which could be used for general chemical applications. For the approximation 3, the optimal parameter values are  $(\mu, \lambda) = (0.46, 0.58)$ , giving a total MAE of 1.68 kcal/mol. For the approximation 4, the optimal parameter values are  $(\mu, \lambda) = (0.62, 0.60)$ , giving a total MAE of 1.98 kcal/mol. In the following, we further assess the approximations 3 and 4 with these optimal parameters.

## 4.8 Assessment on the AE49 and DBH24/08 sets of atomization energies and reaction barrier heights

We assess now the RSDH scheme with the approximations 3 and 4, evaluated with the previously determined optimal parameters  $(\mu, \lambda)$ , on the larger AE49 and DBH24/08 sets of atomization energies and reaction barrier heights. The results are reported in Tables 4.1 and 4.2, and compared with other methods corresponding to limiting cases of the RSDH scheme : DS1DH [82] (with the PBE exchange-correlation functional [153]) corresponding to the  $\mu = 0$  limit of the RSDH scheme with approximation 4, RSH+MP2 [94] corresponding to the  $\lambda = 0$  limit of the RSDH scheme, and standard MP2 corresponding to the  $\mu \rightarrow \infty$  or  $\lambda = 1$  limit of the RSDH scheme. On the AE49 set, the two RSDH approximations (3 and 4) give very similar results. With a MAE of 4.3 kcal/mol and a RMSD of about 5.1 kcal/mol, they provide an overall improvement over both RSH+MP2 and standard MP2 which give MAEs larger by about 1 kcal/mol and RMSDs larger by about 2 kcal/mol. It turns out that the DS1DH approximation gives a smaller MAE of 3.2 kcal/mol than the two RSDH approximations, but a similar RMSD of 5.0 kcal/mol. On the DBH24/08 set, the two RSDH approximations give less similar but still compa-

rable results with MAEs of 1.9 and 2.7 kcal/mol for approximations 3 and 4, respectively. This is a big improvement over standard MP2 which gives a MAE of 6.2 kcal/mol, but similar to the accuracy of RSH+MP2 which gives a MAE of 2.0 kcal/mol. Again, the smallest MAE of 1.5 kcal/mol is obtained with the the DS1DH approximation. The fact that the DS1DH approximation appears to be globally more accurate than the RSDH approximations on these larger sets but not on the small AE6 and BH6 sets points to a limited representativeness of the latter small sets, and suggests that there may be room for improvement by optimizing the parameters on larger sets.

### 4.8.1 Assessment of the basis convergence

We study now the basis convergence of the RSDH scheme. Figure 4.3 shows the convergence of the total energy of He, Ne, N<sub>2</sub>, and H<sub>2</sub>O with respect to the cardinal number X for a series of Dunning basis sets cc-pVXZ (X = 2, 3, 4, 5), calculated with MP2, RSH+MP2, and RSDH with approximations 3 and 4 (with the parameters ( $\mu, \lambda$ ) optimized on the AE6+BH6 combined set).

The results for MP2 and RSH+MP2 are in agreement with what is already known. MP2 has a slow basis convergence, with the error on the total energy decreasing as a third-power law,  $\Delta E_{\text{MP}} \sim A X^{-3}$  [21, 154], due to the difficulty of describing the short-range part of the correlation hole near the electron-electron cusp. RSH+MP2 has a fast basis convergence, with the error decreasing as an exponential law,  $\Delta E_{\text{RSH+MP2}} \sim B e^{-\beta X}$  [29], since it involves only the long-range MP2 correlation energy.

Unsurprisingly, the RSDH scheme displays a basis convergence which is intermediate between that of MP2 and RSH+MP2. What should be remarked is that, for a given basis, the RSDH basis error is closer to the RSH+MP2 basis error than to the MP2 basis error. The basis dependence of RSDH is thus only moderately affected by the presence of short-range MP2 correlation. This can be understood by the fact that RSDH contains only a modest fraction  $\lambda^2 \approx 0.35$  of the pure short-range MP2 correlation energy  $E_{\text{c,MP}}^{\text{sr},\mu}$  [see Eq. (4.29)], which should have a third-power-law convergence, while the pure long-range correlation energy  $E_{\text{c,MP}}^{\text{lr},\mu}$  and the mixed long-range/short-range correlation energy  $E_{\text{c,MP}}^{\text{lr-sr},\mu}$  both should have an exponential-law convergence. We thus expect the RSDH error to decrease as  $\Delta E_{\text{RSDH}} \sim \lambda^2 A X^{-3} + B e^{-\beta X}$ , with constants A, B,  $\beta$  *a priori* different from the ones introduced for MP2 and RSH+MP2. The results of Figure 4.3 are in fact in agreement with such a basis dependence with similar constants A, B,  $\beta$  for MP2, RSH+MP2, and RSDH.

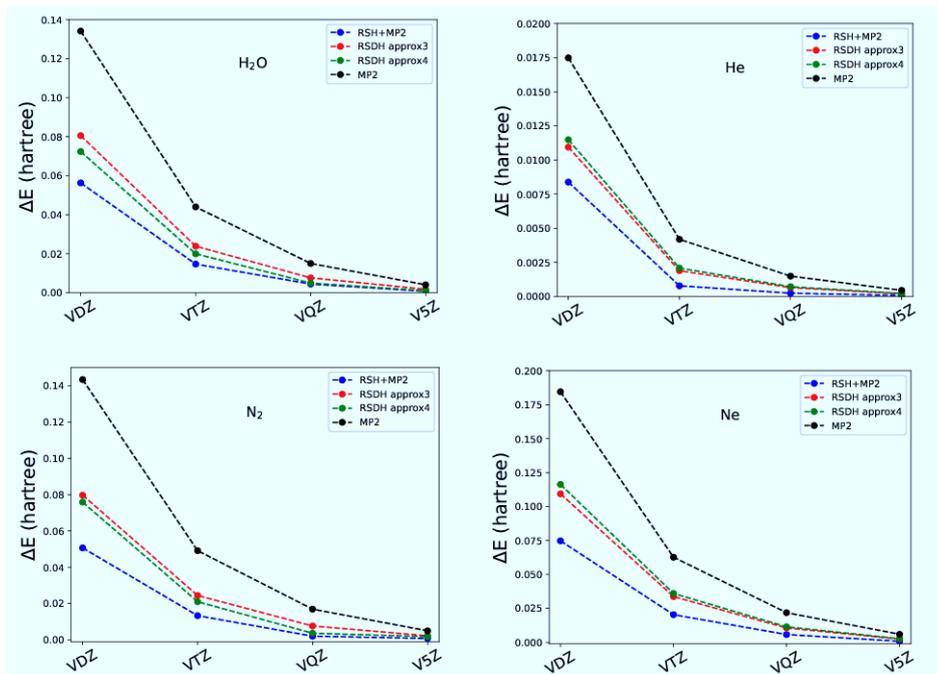


FIGURE 4.3 – Convergence of the total energy with respect to the basis set for He, Ne,  $N_2$ , and  $H_2O$ , as measured by the basis error with respect to the V6Z basis set,  $\Delta E = E_{VXZ} - E_{V6Z}$ , where VXZ stands for cc-pVXZ with  $X = 2$  (D), 3 (T), 4 (Q), 5, calculated with RSH+MP2, RSDH with approximations 3 and 4 of Sec. 4.6.3 (with the short-range PBE exchange-correlation functional of Ref. [128]), and MP2. The parameters ( $\mu, \lambda$ ) used are the ones optimized on the AE6+BH6 combined set.

## 4.8.2 Assessment on the S22 set of intermolecular interactions

We finally test the RSDH scheme on weak intermolecular interactions. Table 4.3 reports the interaction energies for the 22 molecular dimers of the S22 set calculated by RSH+MP2, RSDH (with approximations 3 and 4), and MP2, using the aug-cc-pVDZ and aug-cc-pVTZ basis sets. We also report DS1DH results, but since this method is quite inaccurate for dispersion interactions we only did calculations with the aug-cc-pVDZ basis set for a rough comparison. Again, the basis dependence of RSDH is intermediate between the small basis dependence of RSH+MP2 and the larger basis dependence of standard MP2. The basis convergence study in Section 4.8.1 suggests that the RSDH results with the aug-cc-pVTZ basis set are not too far from the CBS limit.

The two approximations (3 and 4) used in the RSDH scheme give overall similar results, which may be rationalized by the fact low-density regions primarily contribute to these intermolecular interaction energies and the approximations 3 and 4 become identical in the low-density limit. For hydrogen-bonded complexes, RSDH with the aug-cc-pVTZ basis set gives a MA%E of about 3-4%, similar to standard MP2 but in clear improvement

over RSH+MP2 which tends to give too negative interaction energies. Presumably, this is so because the explicit wave-function treatment of the short-range interaction  $\lambda w_{ee}^{sr,\mu}(r_{12})$  makes RSDH accurately describe the short-range component of the intermolecular interaction, while still correctly describes the long-range component. For complexes with a predominant dispersion contribution, RSDH with the aug-cc-pVTZ basis set gives too negative interaction energies by about 30 %, similar to both MP2 and RSH+MP2. Notably, DS1DH gives much too negative interaction energies for the largest and most polarizable systems, leading to a MA%E of more than 100 % with aug-cc-pVDZ basis set. This can be explained by the fact that the reduced amount of HF exchange at long range in DS1DH leads to smaller HOMO-LUMO gaps in these systems in comparison with RSH+MP2 and RSDH, causing a overlarge MP2 contribution. For mixed complexes, RSDH with the aug-cc-pVTZ basis set gives a MA%E of about 14-15 %, which is a bit worse than MP2 but slightly better than RSH+MP2. Again, DS1DH tends to give significantly too negative interaction energies for the largest dimers.

Overall, for weak intermolecular interactions, RSDH thus provides a big improvement over DS1DH, a small improvement over RSH+MP2, and is quite similar to standard MP2.

## 4.9 Conclusion

We have studied a wave-function/DFT hybrid approach based on a CAM-like decomposition of the electron-electron interaction in which a correlated wave-function calculation associated with the two-parameter interaction  $w_{ee}^{lr,\mu}(r_{12}) + \lambda w_{ee}^{sr,\mu}(r_{12})$  is combined with a complement short-range density functional. Specifically, we considered the case of MP2 perturbation theory for the wave-function part and obtained a scheme that we named RSDH. This RSDH scheme is a generalization of the usual one-parameter DHs (corresponding to the special case  $\mu = 0$ ) and the range-separated MP2/DFT hybrid known as RSH+MP2 (corresponding to the special case  $\lambda = 0$ ). It allows one to have both 100% HF exchange and MP2 correlation at long interelectronic distances and fractions of HF exchange and MP2 correlation at short interelectronic distances. We have also proposed a number of approximations for the complement short-range exchange-correlation density functional, based on the limits  $\mu = 0$  and  $\mu \rightarrow \infty$ , and showed their relevance on the uniform-electron gas with the corresponding electron-electron interaction, in particular in the high- and low-density limits.

The RSDH scheme with complement short-range DFAs constructed from a short-range version of the PBE functional has then been applied on small sets of atomization energies (AE6 set) and reaction barrier heights (BH6 set) in order to find optimal values for the

parameters  $\mu$  and  $\lambda$ . It turns out that the optimal values of these parameters for RSDH,  $\mu \approx 0.5$ – $0.6$  and  $\lambda \approx 0.6$ , are very similar to the usual optimal values found separately for RSH+MP2 and one-parameter DHs. With these values of the parameters, RSDH has a relatively fast convergence with respect to the size of the one-electron basis, which can be explained by the fact that it contains only a modest fraction  $\lambda^2 \approx 0.35$  of pure short-range MP2 correlation. We have tested the RSDH scheme with the two best complement short-range DFAs (referred to as approximations 3 and 4) on large sets of atomization energies (AE49 set), reaction barrier heights (DBH24 set), and weak intermolecular interactions (S22 set). The results show that the RSDH scheme is either globally more accurate or comparable to RSH+MP2 and standard MP2. If we had to recommend a computational method for general chemical applications among the methods tested in this work, it would be RSDH with approximation 3 with parameters  $(\mu, \lambda) = (0.46, 0.58)$ .

There is much room however for improvement and extension. The parameters  $\mu$  and  $\lambda$  could be optimized on larger training sets. More accurate complement short-range DFAs should be constructed. The MP2 correlation term could be replaced by random-phase approximations, which would more accurately describe dispersion interactions [31, 101], or by multireference perturbation theory [103], which would capture static correlation effects. The RSDH scheme could be extended to linear-response theory for calculating excitation energies or molecular properties, e.g. by generalizing the methods of Refs. [155–158].

## Acknowledgements

We thank Bastien Mussard for help with the MOLPRO software. We also thank Labex MiChem for providing PhD financial support for C. Kalai.

## 4.10 Uniform coordinate scaling relation and Coulomb/high-density and short-range/low-density limits of $\bar{E}_c^{\text{sr},\mu,\lambda}[\text{n}]$

### 4.10.1 Scaling relation for $\bar{E}_c^{\text{sr},\mu,\lambda}[\text{n}]$

Here, we generalize the uniform coordinate scaling relation, known for the KS correlation functional  $E_c[\text{n}]$  [90, 92, 131] and for the complement short-range correlation functional  $\bar{E}_c^{\text{sr},\mu}[\text{n}]$  [134, 159], to the  $\lambda$ -dependent complement short-range correlation functional  $\bar{E}_c^{\text{sr},\mu,\lambda}[\text{n}]$ . We first define the universal density functional, for arbitrary parameters  $\mu \geq 0$ ,

$\lambda \geq 0$ , and  $\xi \geq 0$ ,

$$F^{\mu,\lambda,\xi}[\mathbf{n}] = \min_{\Psi \rightarrow \mathbf{n}} \langle \Psi | \hat{T} + \xi \hat{W}_{ee}^{\text{lr},\mu} + \lambda \hat{W}_{ee}^{\text{sr},\mu} | \Psi \rangle, \quad (4.56)$$

which is a simple generalization of the universal functional  $F^{\mu,\lambda}[\mathbf{n}]$  in Eq. (4.7) such that  $F^{\mu,\lambda,\xi=1}[\mathbf{n}] = F^{\mu,\lambda}[\mathbf{n}]$ . The minimizing wave function in Eq. (4.56) will be denoted by  $\Psi^{\mu,\lambda,\xi}[\mathbf{n}]$ . Let us now consider the scaled wave function  $\Psi_Y^{\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}]$  defined by, for  $N$  electrons,

$$\Psi_Y^{\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}](\mathbf{r}_1, \dots, \mathbf{r}_N) = \gamma^{3N/2} \Psi^{\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}](\gamma\mathbf{r}_1, \dots, \gamma\mathbf{r}_N), \quad (4.57)$$

where  $\gamma > 0$  is a scaling factor. The wave function  $\Psi_Y^{\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}]$  yields the scaled density  $n_Y(\mathbf{r}) = \gamma^3 n(\gamma\mathbf{r})$  and minimizes  $\langle \Psi | \hat{T} + \xi \hat{W}_{ee}^{\text{lr},\mu} + \lambda \hat{W}_{ee}^{\text{sr},\mu} | \Psi \rangle$  since

$$\begin{aligned} \langle \Psi_Y^{\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}] | \hat{T} + \xi \hat{W}_{ee}^{\text{lr},\mu} + \lambda \hat{W}_{ee}^{\text{sr},\mu} | \Psi_Y^{\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}] \rangle = \\ \gamma^2 \langle \Psi^{\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}] | \hat{T} + (\xi/\gamma) \hat{W}_{ee}^{\text{lr},\mu/\gamma} + (\lambda/\gamma) \hat{W}_{ee}^{\text{sr},\mu/\gamma} | \Psi^{\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}] \rangle, \end{aligned} \quad (4.58)$$

where the right-hand side is minimal by definition of  $\Psi^{\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}]$ . Therefore, we conclude that

$$\Psi^{\mu,\lambda,\xi}[\mathbf{n}_\gamma] = \Psi_Y^{\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}], \quad (4.59)$$

and

$$F^{\mu,\lambda,\xi}[\mathbf{n}_\gamma] = \gamma^2 F^{\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}]. \quad (4.60)$$

Consequently, the corresponding correlation functional,

$$E_c^{\mu,\lambda,\xi}[\mathbf{n}] = \langle \Psi^{\mu,\lambda,\xi}[\mathbf{n}] | \hat{T} + \xi \hat{W}_{ee}^{\text{lr},\mu} + \lambda \hat{W}_{ee}^{\text{sr},\mu} | \Psi^{\mu,\lambda,\xi}[\mathbf{n}] \rangle - \langle \Phi[\mathbf{n}] | \hat{T} + \xi \hat{W}_{ee}^{\text{lr},\mu} + \lambda \hat{W}_{ee}^{\text{sr},\mu} | \Phi[\mathbf{n}] \rangle \quad (4.61)$$

with the KS single-determinant wave function  $\Phi[\mathbf{n}] = \Psi^{\mu=0,\lambda=0,\xi}[\mathbf{n}]$ , satisfies the same scaling relation

$$E_c^{\mu,\lambda,\xi}[\mathbf{n}_\gamma] = \gamma^2 E_c^{\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}]. \quad (4.62)$$

Similarly, the associated short-range complement correlation functional,

$$\bar{E}_c^{\text{sr},\mu,\lambda,\xi}[\mathbf{n}] = E_c^{\xi}[\mathbf{n}] - E_c^{\mu,\lambda,\xi}[\mathbf{n}], \quad (4.63)$$

with  $E_c^{\xi}[\mathbf{n}] = E_c^{\mu \rightarrow \infty, \lambda, \xi}[\mathbf{n}]$ , satisfies the scaling relation

$$\bar{E}_c^{\text{sr},\mu,\lambda,\xi}[\mathbf{n}_\gamma] = \gamma^2 \bar{E}_c^{\text{sr},\mu/\gamma,\lambda/\gamma,\xi/\gamma}[\mathbf{n}]. \quad (4.64)$$

Applying this relation for  $\xi = 1$  gives the scaling relation for  $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$ ,

$$\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}_\gamma] = \gamma^2 \bar{E}_c^{\text{sr},\mu/\gamma,\lambda/\gamma,1/\gamma}[\mathbf{n}], \quad (4.65)$$

from which we see that the high-density limit  $\gamma \rightarrow \infty$  is related to the Coulomb limit  $\mu \rightarrow 0$  and the low-density limit  $\gamma \rightarrow 0$  is related to the short-range limit  $\mu \rightarrow \infty$  of  $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$ . Note that by applying Eq. (4.64) with  $\lambda = 0$  and  $\gamma = \xi$  we obtain the short-range complement correlation functional associated with the interaction  $\xi w_{ee}^{\text{lr},\mu}$  in terms of the short-range complement correlation functional associated with the interaction  $w_{ee}^{\text{lr},\mu/\xi}$ , i.e.  $\bar{E}_c^{\text{sr},\mu,0,\xi}[\mathbf{n}] = \xi^2 \bar{E}_c^{\text{sr},\mu/\xi,0,1}[\mathbf{n}_{1/\xi}] = \xi^2 \bar{E}_c^{\text{sr},\mu/\xi}[\mathbf{n}_{1/\xi}]$ , as already explained in Ref. [159]. Also, by applying Eq. (4.64) with  $\xi = 1$  and  $\gamma = \lambda$  we obtain the short-range complement correlation functional associated with the interaction  $w_{ee}^{\text{lr},\mu} + \lambda w_{ee}^{\text{sr},\mu}$  in terms of the short-range complement correlation functional associated with the interaction  $(1/\lambda)w_{ee}^{\text{lr},\mu/\lambda} + w_{ee}^{\text{sr},\mu/\lambda}$ , i.e.  $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}] = \bar{E}_c^{\text{sr},\mu,\lambda,1}[\mathbf{n}] = \lambda^2 \bar{E}_c^{\text{sr},\mu/\lambda,1,1/\lambda}[\mathbf{n}_{1/\lambda}]$ . However, since the functional  $\bar{E}_c^{\text{sr},\mu,1,\xi}[\mathbf{n}]$  is equally unknown as the functional  $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$ , this relation is not useful to obtain an expression for the functional  $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$ .

#### 4.10.2 Coulomb limit and high-density limit of $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$

We first give the limit of  $\bar{E}_c^{\text{sr},\mu,\lambda,\xi}[\mathbf{n}]$  as  $\mu \rightarrow 0$ . Starting from Eq. (4.63) and noting that  $E_c^{\mu=0,\lambda,\xi}[\mathbf{n}] = E_c^\lambda[\mathbf{n}]$ , we obtain :

$$\bar{E}_c^{\text{sr},\mu=0,\lambda,\xi}[\mathbf{n}] = E_c^\xi[\mathbf{n}] - E_c^\lambda[\mathbf{n}] = \xi^2 E_c[\mathbf{n}_{1/\xi}] - \lambda^2 E_c[\mathbf{n}_{1/\lambda}], \quad (4.66)$$

where we have used the well-known relation,  $E_c^\lambda[\mathbf{n}] = \lambda^2 E_c[\mathbf{n}_{1/\lambda}]$  [82, 92, 131] [a special case of Eq. (4.62)]. In particular, for  $\xi = 1$  we obtain the limit of  $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$  as  $\mu \rightarrow 0$

$$\bar{E}_c^{\text{sr},\mu=0,\lambda}[\mathbf{n}] = E_c[\mathbf{n}] - \lambda^2 E_c[\mathbf{n}_{1/\lambda}]. \quad (4.67)$$

We can now derive the high-density limit of  $\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}]$  using the scaling relation in Eq. (4.65) and the limit  $\mu \rightarrow 0$  in Eq. (4.66)

$$\begin{aligned} \lim_{\gamma \rightarrow \infty} \bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}_\gamma] &= \lim_{\gamma \rightarrow \infty} \gamma^2 \bar{E}_c^{\text{sr},\mu/\gamma,\lambda/\gamma,1/\gamma}[\mathbf{n}] \\ &= \lim_{\gamma \rightarrow \infty} \left( E_c[\mathbf{n}_\gamma] - \lambda^2 E_c[\mathbf{n}_{\gamma/\lambda}] \right) \\ &= (1 - \lambda^2) E_c^{\text{GL2}}[\mathbf{n}], \end{aligned} \quad (4.68)$$

where we have used  $\lim_{\gamma \rightarrow \infty} E_c[n_\gamma] = E_c^{\text{GL2}}[n]$  [60] assuming a KS system with a non-degenerate ground state.

### Short-range limit and low-density limit of $\bar{E}_c^{\text{sr},\mu,\lambda}[n]$

We first derive the leading term of the asymptotic expansion of  $\bar{E}_c^{\text{sr},\mu,\lambda,\xi}[n]$  as  $\mu \rightarrow \infty$ . Taking the derivative with respect to  $\lambda$  of Eq. (4.61), and using the Hellmann-Feynman theorem which states that the derivative of  $\Psi^{\mu,\lambda,\xi}[n]$  does not contribute, we obtain :

$$\begin{aligned} \frac{\partial E_c^{\mu,\lambda,\xi}[n]}{\partial \lambda} &= \langle \Psi^{\mu,\lambda,\xi}[n] | \hat{W}_{ee}^{\text{sr},\mu} | \Psi^{\mu,\lambda,\xi}[n] \rangle - \langle \Phi[n] | \hat{W}_{ee}^{\text{sr},\mu} | \Phi[n] \rangle \\ &= \frac{1}{2} \iint n_{2,c}^{\mu,\lambda,\xi}[n](\mathbf{r}_1, \mathbf{r}_2) w_{ee}^{\text{sr},\mu}(r_{12}) d\mathbf{r}_1 d\mathbf{r}_2, \end{aligned} \quad (4.69)$$

where  $n_{2,c}^{\mu,\lambda,\xi}[n](\mathbf{r}_1, \mathbf{r}_2) = \langle \Psi^{\mu,\lambda,\xi}[n] | \hat{n}_2(\mathbf{r}_1, \mathbf{r}_2) | \Psi^{\mu,\lambda,\xi}[n] \rangle - \langle \Phi[n] | \hat{n}_2(\mathbf{r}_1, \mathbf{r}_2) | \Phi[n] \rangle$  is the correlation part of the pair density associated with the wave function  $\Psi^{\mu,\lambda,\xi}[n]$ . Noting from Eq. (4.63) that  $\partial \bar{E}_c^{\text{sr},\mu,\lambda,\xi}[n] / \partial \lambda = -\partial E_c^{\mu,\lambda,\xi}[n] / \partial \lambda$ , integrating over  $\lambda$ , and using  $\bar{E}_c^{\text{sr},\mu,\lambda=\xi,\xi}[n] = 0$ , we arrive at the exact expression :

$$\bar{E}_c^{\text{sr},\mu,\lambda,\xi}[n] = \frac{1}{2} \int_{\lambda}^{\xi} d\alpha \iint n_{2,c}^{\mu,\alpha,\xi}[n](\mathbf{r}_1, \mathbf{r}_2) w_{ee}^{\text{sr},\mu}(r_{12}) d\mathbf{r}_1 d\mathbf{r}_2. \quad (4.70)$$

Using now the asymptotic expansion of the short-range interaction [27],

$$w_{ee}^{\text{sr},\mu}(r_{12}) = \frac{\pi}{\mu^2} \delta(\mathbf{r}_{12}) + \mathcal{O}\left(\frac{1}{\mu^3}\right), \quad (4.71)$$

we obtain the leading term of the asymptotic expansion of  $\bar{E}_c^{\text{sr},\mu,\lambda,\xi}[n]$  as  $\mu \rightarrow \infty$ ,

$$\bar{E}_c^{\text{sr},\mu,\lambda,\xi}[n] = (\xi - \lambda) \frac{\pi}{2\mu^2} \int n_{2,c}^{\xi}[n](\mathbf{r}, \mathbf{r}) d\mathbf{r} + \mathcal{O}\left(\frac{1}{\mu^3}\right), \quad (4.72)$$

where  $n_{2,c}^{\xi}[n](\mathbf{r}, \mathbf{r}) = n_{2,c}^{\mu \rightarrow \infty, \alpha, \xi}[n](\mathbf{r}, \mathbf{r})$  is the correlation part of the on-top pair density associated with the scaled Coulomb interaction  $\xi w_{ee}(r_{12})$ . For the special case  $\xi = 1$ , we obtain the leading term of the asymptotic expansion of  $\bar{E}_c^{\text{sr},\mu,\lambda}[n]$

$$\bar{E}_c^{\text{sr},\mu,\lambda}[n] = (1 - \lambda) \frac{\pi}{2\mu^2} \int n_{2,c}[n](\mathbf{r}, \mathbf{r}) d\mathbf{r} + \mathcal{O}\left(\frac{1}{\mu^3}\right), \quad (4.73)$$

where  $n_{2,c}[n](\mathbf{r}, \mathbf{r})$  is the correlation part of the on-top pair density associated with the Coulomb interaction. We can now derive the low-density limit of  $\bar{E}_c^{\text{sr},\mu,\lambda}[n]$  using the scaling

relation in Eq. (4.65) and the asymptotic expansion as  $\mu \rightarrow \infty$  in Eq. (4.72)

$$\begin{aligned}
\bar{E}_c^{\text{sr},\mu,\lambda}[\mathbf{n}_\gamma] &= \gamma^2 \bar{E}_c^{\text{sr},\mu/\gamma,\lambda/\gamma,1/\gamma}[\mathbf{n}] \\
&\underset{\gamma \rightarrow 0}{\sim} \gamma^3 (1-\lambda) \frac{\pi}{2\mu^2} \int n_{2,c}^{1/\gamma}[\mathbf{n}](\mathbf{r}, \mathbf{r}) d\mathbf{r} \\
&\underset{\gamma \rightarrow 0}{\sim} \gamma^3 (1-\lambda) \frac{\pi}{4\mu^2} \int [-\mathbf{n}(\mathbf{r})^2 + \mathbf{m}(\mathbf{r})^2] d\mathbf{r}, \tag{4.74}
\end{aligned}$$

where we have used the strong-interaction limit of the on-top pair density,  $\lim_{\gamma \rightarrow 0} n_{2,c}^{1/\gamma}[\mathbf{n}](\mathbf{r}, \mathbf{r}) = -\mathbf{n}(\mathbf{r})^2/2 + \mathbf{m}(\mathbf{r})^2/2 = -2n_\uparrow(\mathbf{r})n_\downarrow(\mathbf{r})$  [132] where  $\mathbf{m}(\mathbf{r})$  is the spin magnetization and  $n_\sigma(\mathbf{r})$  are the spin densities ( $\sigma = \uparrow, \downarrow$ ).

Tableau 4.1 – Atomization energies (in kcal/mol) of the AE49 set calculated by DS1DH (with the PBE exchange-correlation functional [153]), RSH+MP2, RSDH with approximations 3 and 4 of Sec. 4.6.3 (with the short-range PBE exchange-correlation functional of Ref. [128]), and MP2. The calculations were carried out using the cc-pVQZ basis set at MP2(full)/6-31G\* geometries and with parameters  $(\mu, \lambda)$  optimized on the AE6+BH6 combined set. The reference values are the non-relativistic FC-CCSD(T)/cc-pVQZ-F12 values of Ref. [142].

Molecule	DS1DH $(\mu, \lambda) = (0, 0.70)$	RSH+MP2 $(0.58, 0)$	RSDH approx3 $(0.46, 0.58)$	RSDH approx4 $(0.62, 0.60)$	MP2	Reference
CH	81.13	78.38	79.93	79.58	79.68	83.87
CH <sub>2</sub> ( <sup>3</sup> B <sub>1</sub> )	190.68	190.19	190.42	190.32	188.70	189.74
CH <sub>2</sub> ( <sup>1</sup> A <sub>1</sub> )	175.20	170.26	173.36	173.24	174.45	180.62
CH <sub>3</sub>	305.32	302.91	304.43	304.34	303.36	306.59
CH <sub>4</sub>	415.79	410.84	414.23	414.45	414.83	418.87
NH	81.39	81.09	80.28	79.49	78.57	82.79
NH <sub>2</sub>	179.12	177.12	177.03	176.22	176.65	181.96
NH <sub>3</sub>	293.24	288.76	290.33	290.02	293.11	297.07
OH	105.26	104.49	104.02	103.81	105.78	106.96
OH <sub>2</sub>	229.78	225.48	226.96	227.21	233.83	232.56
FH	140.11	137.20	138.16	138.35	144.17	141.51
SiH <sub>2</sub> ( <sup>1</sup> A <sub>1</sub> )	146.66	143.21	146.22	146.15	145.90	153.68
SiH <sub>2</sub> ( <sup>3</sup> B <sub>1</sub> )	130.48	133.05	130.56	130.39	128.93	133.26
SiH <sub>3</sub>	222.07	220.05	222.39	222.21	220.51	228.08
SiH <sub>4</sub>	315.08	311.89	315.80	315.81	314.27	324.59
PH <sub>2</sub>	148.62	146.37	147.60	146.98	144.95	153.97
PH <sub>3</sub>	233.35	229.18	232.06	231.69	230.24	241.47
SH <sub>2</sub>	178.66	174.18	177.29	177.66	178.55	183.30
ClH	105.39	101.63	104.43	104.90	106.53	107.20
HCCH	406.75	399.05	403.84	405.17	409.58	402.76
H <sub>2</sub> CCH <sub>2</sub>	561.56	554.55	559.03	559.63	561.38	561.34
H <sub>3</sub> CCH <sub>3</sub>	708.35	701.94	706.59	706.98	707.15	710.20
CN	178.63	172.93	174.03	172.54	168.84	180.06
HCN	315.27	305.21	310.17	310.79	319.26	311.52
CO	262.93	254.60	258.48	258.64	269.29	258.88
HCO	283.19	277.00	278.81	278.50	285.79	278.28
H <sub>2</sub> CO	375.40	367.85	370.90	370.96	379.19	373.21
H <sub>3</sub> COH	510.48	505.00	507.31	507.41	513.32	511.83
N <sub>2</sub>	229.78	218.09	223.07	222.86	234.80	227.44
H <sub>2</sub> NNH <sub>2</sub>	433.52	428.92	428.93	427.80	432.46	436.70
NO	156.24	151.17	151.16	149.94	156.94	152.19
O <sub>2</sub>	126.21	119.73	120.29	119.49	128.55	120.54
HOOH	267.02	259.77	260.59	260.11	272.51	268.65
F <sub>2</sub>	38.66	31.64	32.30	31.17	42.18	38.75
CO <sub>2</sub>	400.49	390.46	393.18	393.30	409.33	388.59
Si <sub>2</sub>	71.64	67.21	69.78	70.72	70.56	73.41
P <sub>2</sub>	112.91	107.27	111.26	112.87	113.59	115.95
S <sub>2</sub>	104.29	100.90	102.71	103.56	103.67	103.11
Cl <sub>2</sub>	58.97	54.19	57.31	57.94	60.43	59.07
SiO	192.77	185.82	189.17	189.93	200.09	192.36
SC	172.35	163.07	164.01	170.43	175.16	170.98
SO	127.74	122.46	123.89	123.77	129.29	125.80
ClO	62.96	60.81	59.82	58.55	59.69	64.53
ClF	62.43	57.94	58.98	58.50	65.20	62.57
Si <sub>2</sub> H <sub>6</sub>	521.08	517.07	522.83	522.88	519.17	535.47
CH <sub>3</sub> Cl	393.93	388.23	392.44	393.14	394.57	394.52
CH <sub>3</sub> SH	470.26	463.90	468.46	469.10	469.94	473.49
HOCl	164.83	158.46	160.70	160.75	168.50	165.79
SO <sub>2</sub>	260.22	244.46	250.91	251.27	270.72	259.77
MAE	3.19	5.49	4.31	4.31	5.37	
ME	-1.18	-6.32	-3.98	-3.97	-0.24	
RMSD	4.98	7.41	5.13	5.18	6.75	
Min error	-14.39	-18.40	-12.64	-12.59	-16.30	
Max error	11.90	1.87	93 4.59	4.71	20.74	

Tableau 4.2 – Forward (F) and reverse (R) reaction barrier heights (in kcal/mol) of the DBH24/08 set calculated by DS1DH (with the PBE exchange-correlation functional [153]), RSH+MP2, RSDH with approximations 3 and 4 of Sec. 4.6.3 (with the short-range PBE exchange-correlation functional of Ref. [128]), and MP2. The calculations were carried out using the aug-cc-pVQZ basis set at QCISD/MG3 geometries and with parameters  $(\mu, \lambda)$  optimized on the AE6+BH6 combined set. The reference values are taken from Ref. [148].

Reaction	$(\mu, \lambda) =$	DS1DH (0,0.70)	RSH+MP2 (0.58,0)	RSDH approx3 (0.46,0.58)	RSDH approx4 (0.62,0.60)	MP2	Reference
		F/R	F/R	F/R	F/R	F/R	F/R
Heavy-atom transfer							
$\text{H} + \text{N}_2\text{O} \rightarrow \text{OH} + \text{N}_2$		21.64/75.80	19.34/77.14	22.76/80.39	25.01/82.80	35.94/89.26	17.13/82.47
$\text{H} + \text{ClH} \rightarrow \text{HCl} + \text{H}$		18.51/18.51	19.77/19.77	20.23/20.23	20.99/20.99	22.79/22.79	18.00/18.00
$\text{CH}_3 + \text{FCl} \rightarrow \text{CH}_3\text{F} + \text{Cl}$		7.54/60.77	8.21/63.59	9.79/64.81	11.25/66.64	19.74/74.29	6.75/60.00
Nucleophilic substitution							
$\text{Cl}^- \cdots \text{CH}_3\text{Cl} \rightarrow \text{ClCH}_3 \cdots \text{Cl}^-$		12.45/12.45	15.40/15.40	14.36/14.36	9.90/9.90	14.64/14.64	13.41/13.41
$\text{F}^- \cdots \text{CH}_3\text{Cl} \rightarrow \text{FCH}_3 \cdots \text{Cl}^-$		2.83/27.93	4.72/31.46	3.99/30.52	4.27/30.67	4.59/28.88	3.44/29.42
$\text{OH}^- + \text{CH}_3\text{F} \rightarrow \text{HOCH}_3 + \text{F}^-$		-3.11/16.76	-1.59/21.56	-1.92/19.23	-1.53/19.56	-1.75/17.86	-2.44/17.66
Unimolecular and association							
$\text{H} + \text{N}_2 \rightarrow \text{HN}_2$		16.36/10.27	14.03/13.09	17.00/11.57	18.75/11.45	27.60/8.06	14.36/10.61
$\text{H} + \text{C}_2\text{H}_4 \rightarrow \text{CH}_3\text{CH}_2$		4.15/44.13	2.70/45.76	4.34/45.49	5.40/45.89	9.32/46.54	1.72/41.75
$\text{HCN} \rightarrow \text{HNC}$		49.13/33.01	48.52/34.81	49.07/33.59	50.05/33.95	34.46/52.09	48.07/32.82
Hydrogen transfer							
$\text{OH} + \text{CH}_4 \rightarrow \text{CH}_3 + \text{H}_2\text{O}$		4.54/19.33	6.03/19.75	6.53/20.38	7.33/21.35	7.66/25.01	6.70/19.60
$\text{H} + \text{OH} \rightarrow \text{O} + \text{H}_2$		12.22/11.02	13.44/10.00	13.49/12.64	14.47/13.94	17.56/15.58	10.70/13.10
$\text{H} + \text{H}_2\text{S} \rightarrow \text{H}_2 + \text{HS}$		4.04/15.09	4.73/15.35	5.00/15.81	5.46/15.96	6.42/16.36	3.60/17.30
MAE		1.52	2.01	1.85	2.65	6.17	
ME		-0.09	1.06	1.50	1.95	4.70	
RMSD		2.09	2.36	2.30	3.26	8.56	
Min error		-6.67	-5.33	-2.08	-3.51	-13.61	
Max error		4.51	4.01	5.63	7.88	19.61	

Tableau 4.3 – Interaction energies (in kcal/mol) for the complexes of the S22 set calculated by DS1DH (with the PBE exchange-correlation functional [153]), RSH+MP2, RSDH with approximations 3 and 4 of Sec. 4.6.3 (with the short-range PBE exchange-correlation functional of Ref. [128]), and MP2. The parameters ( $\mu, \lambda$ ) used are the ones optimized on the AE6+BH6 combined set, except for the RSH+MP2 values which are taken from Ref. [101] in which  $\mu = 0.50$  was used. The basis sets used are aVDZ and aVTZ which refer to aug-cc-pVDZ and aug-cc-pVTZ, respectively, and the counterpoise correction is applied. The values in italics were obtained using the local MP2 approach, the ones with an asterisk (\*) were obtained in Ref. [160] with the density-fitting approximation, and the ones with a dagger (†) were obtained with the approximation :  $E_{\text{aVTZ}}(\text{RSDH approx4}) \approx E_{\text{aVDZ}}(\text{RSDH approx4}) + E_{\text{aVTZ}}(\text{RSDH approx3}) - E_{\text{aVDZ}}(\text{RSDH approx3})$ . The geometries of the complexes are taken from Ref. [150] and the reference interaction energies are taken as the CCSD(T)/CBS estimates of Ref. [151]. The MP2 values are also taken from Ref. [151].

Complex	$(\mu, \lambda) =$	DS1DH	RSH+MP2		RSDH approx3		RSDH approx4		MP2		Reference
		(0,0.70) aVDZ	(0.50,0) aVDZ	aVTZ	(0.46,0.58) aVDZ	aVTZ	(0.62,0.60) aVDZ	aVTZ	aVDZ	aVTZ	
Hydrogen-bonded complexes											
Ammonia dimer		-2.70	-3.13	-3.25	-3.00	-3.18	-2.94	-3.16	-2.68	-2.99	-3.17
Water dimer		-4.63	-5.34	-5.45	-5.03	-5.19	-4.93	-5.12	-4.36	-4.69	-5.02
Formic acid dimer		-17.28	-21.20	-21.57	-19.31	-20.14	-18.86	-19.80	-15.99	-17.55	-18.80
Formamide dimer		-14.63	-17.44	-17.64	-16.30	-16.81	-15.98	-16.60	-13.95	-15.03	-16.12
Uracile dimer $C_{2h}$		<i>-18.86</i>	-22.62	<i>-22.82*</i>	<i>-20.52</i>	<i>-21.77</i>	<i>-20.53</i>	<i>-21.78†</i>	-18.41	-19.60	-20.69
2-pyridoxine/2-aminopyridine		<i>-18.65</i>	-18.86	<i>-18.60*</i>	<i>-17.43</i>	<i>-17.93</i>	<i>-17.04</i>	<i>-17.55†</i>	-15.56	-16.64	-17.00
Adenine/thymine WC		<i>-17.52</i>	-18.26	<i>-18.12*</i>	<i>-16.47</i>	<i>-17.28</i>	<i>-16.23</i>	<i>-17.04†</i>	-14.71	-15.80	-16.74
MAE		1.16	1.34	1.42	0.26	0.68	0.23	0.51	1.70	0.75	
ME		0.46	-1.33	-1.42	-0.07	-0.68	0.22	-0.50	1.70	0.75	
RMSD		1.28	1.56	1.66	0.29	0.81	0.29	0.64	1.88	0.85	
MA%E		9.00	8.36	9.03	2.04	4.14	2.35	3.01	12.63	5.62	
Complexes with predominant dispersion contribution											
Methane dimer		-0.25	-0.46	-0.48	-0.42	-0.47	-0.42	-0.47	-0.39	-0.46	-0.53
Ethene dimer		-0.84	-1.45	-1.55	-1.38	-1.68	-1.33	-1.55	-1.18	-1.46	-1.50
Benzene/methane		-0.87	-1.62	-1.71	-1.56	-1.70	-1.56	-1.63	-1.47	-1.71	-1.45
Benzene dimer $C_{2h}$		-7.21	-4.08	<i>-4.24*</i>	-3.52	-3.78	-4.14	-4.40†	-4.25	-4.70	-2.62
Pyrazine dimer		-8.97	-5.97	<i>-6.04*</i>	-6.50	-6.21	-6.02	-5.73†	-6.00	-6.55	-4.20
Uracil dimer $C_2$		-13.31	-11.76	<i>-11.95*</i>	-12.70	-11.42	-10.77	-9.49†	-9.80	-10.63	-9.74
Indole/benzene		<i>-17.26</i>	-6.95	<i>-6.96*</i>	<i>-8.83</i>	<i>-6.97</i>	<i>-9.25</i>	<i>-7.39†</i>	-7.13	-7.74	-4.59
Adenine/thymine stack		<i>-20.84</i>	-15.11	<i>-14.71*</i>	<i>-14.28</i>	<i>-14.56</i>	<i>-14.25</i>	<i>-14.53†</i>	-13.24	-14.26	-11.66
MAE		4.54	1.42	1.43	1.67	1.33	1.50	1.19	1.01	1.43	
ME		-4.16	-1.39	-1.42	-1.61	-1.31	-1.43	-1.11	-0.90	-1.40	
RMSD		6.14	1.83	1.80	2.23	1.67	2.10	1.65	1.37	1.85	
MA%E		102.09	28.48	29.60	34.02	28.31	34.42	27.45	27.96	33.65	
Mixed complexes											
Ethene/ethyne		-1.28	-1.62	-1.68	-1.57	-1.68	-1.43	-1.67	-1.39	-1.58	-1.51
Benzene/water		-2.66	-3.49	-3.68	-3.33	-3.55	-3.29	-3.53	-2.98	-3.35	-3.29
Benzene/ammonia		-1.70	-2.49	-2.63	-2.39	-2.58	-2.38	-2.59	-2.21	-2.52	-2.32
Benzene/hydrogen cyanide		-3.86	-5.31	-5.38	-4.93	-5.26	-4.89	-5.26	-4.37	-4.92	-4.55
Benzene dimer $C_{2v}$		-4.57	-3.33	<i>-3.49*</i>	-3.26	-3.47	-3.26	-3.47†	-3.09	-3.46	-2.71
Indole/benzene T-shaped		<i>-11.71</i>	-6.55	<i>-6.85*</i>	<i>-7.49</i>	<i>-6.50</i>	<i>-7.91</i>	<i>-6.92†</i>	-6.10	-6.71	-5.62
Phenol dimer		<i>-8.05</i>	-8.05	<i>-8.09*</i>	<i>-6.89</i>	<i>-7.57</i>	<i>-7.15</i>	<i>-7.83†</i>	-6.79	-7.36	-7.09
MAE		1.58	0.54	0.67	0.45	0.50	0.48	0.60	0.27	0.40	
ME		-0.97	-0.54	-0.67	-0.40	-0.50	-0.46	-0.60	0.02	-0.40	
RMSD		2.31	0.76	0.86	0.75	0.57	0.90	0.71	0.50	0.69	
MA%E		38.01	12.34	17.40	10.43	13.78	11.03	15.28	7.55	10.58	
Total MAE		2.52	1.11	1.18	0.83	0.86	0.77	0.75	0.99	0.89	
Total ME		-1.67	-1.09	-1.18	-0.73	-0.85	-0.60	-0.75	0.22	-0.40	
Total RMSD		4.03	1.45	1.49	1.42	1.15	1.37	1.13	1.35	1.25	
Total MA%E		52.08	16.95	19.06	16.34	16.00	16.77	15.80	16.59	17.36	

Troisième partie

**Applications**

## Chapitre 5

# Etude des complexes $\beta$ -propiolactone micro-hydratés.

## 5.1 De l'importance de l'étude des complexes micro-hydratés.

La micro-solvatation des molécules organiques par l'eau est un phénomène majeur, notamment en chimie atmosphérique [161]. Le rôle de l'hydratation sur la réactivité des polluants atmosphériques [162] fait l'objet de récentes études [163]. La stabilité des complexes micro-hydratés est essentiellement due à la complémentarité et/ou compétitivité des interactions eau-eau et eau-soluté. L'existence d'interactions multiples peut conduire à une stabilisation accrue des agrégats, ainsi deux liaisons hydrogène peuvent mutuellement se renforcer : on parle d'effets coopératifs. Dans les clusters d'eau, les effets coopératifs se traduisent par une réduction exponentielle de la distance O - O entre deux molécules d'eau en interaction [164] et un décalage vers le rouge de la fréquence d'élongation de la liaisons covalente O-H impliquée dans la liaison hydrogène lorsque la taille du cluster croît [165]. A l'inverse, il existe également des effets anti-coopératifs, qui sont dus à un affaiblissement d'une interaction dans un complexe. Guevara-Vela et al. [166] ont montré que dans le cas de l'hexamère d'eau, les effets anti-coopératifs sont dues à l'occurrence de double donneur et double accepteur de liaison hydrogène. Les effets coopératifs sont observés lorsque chaque molécule formant le complexe est à la fois donneur et accepteur de liaison hydrogène. Ce qui revient à dire que les effets coopératifs sont caractérisés par des cycles "homodromic" et que les effets anti-coopératifs sont caractérisés par des cycles "antidromic" [167]. La figure 5.1 montre un exemple de cycle "homodromic" et "antidromic".

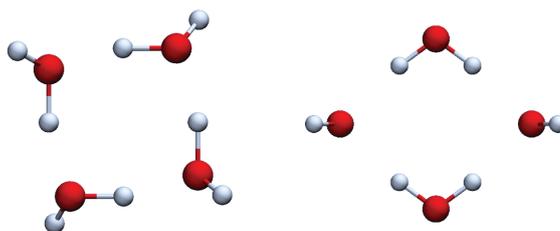


FIGURE 5.1 – Exemple de cycle dit "homodromic"(à gauche) et de cycle dit "antidromic" (à droite).

L'étude de ces effets en présence d'un soluté organique a été peu étudiée et pourrait être responsable d'un certain nombre de faits théoriques et expérimentaux. En effet, une étude [168] sur la molécule de benzène hydratée par huit molécules d'eau basée sur des calculs CCSD(T), MP2 et DFT-D a montré que la structure générale des clusters

d'eau est légèrement affectée par l'interaction avec le soluté, le réseau de liaison hydrogène n'étant pas rompu. Cette étude met également en évidence des conformères quasi-isoénergétiques. Nous pouvons également citer une étude de la micro-hydratation de la 2-aminooxazole [169] reposant sur des calculs DFT et sur un champ de force qui a montré que les interactions eau-soluté étaient plus faibles que les interactions eau-eau pour des complexes de petite taille.

Les caractérisations structurales de complexes micro-hydratés basées sur des données expérimentales de référence sont particulièrement intéressantes [170] [171]. Une manière de déterminer expérimentalement la structure des complexes repose sur des études de spectroscopie micro-ondes (généralement couplée à une expansion supersonique) dans lesquelles on effectue des substitutions isotopiques [11]. Une variation du moment d'inertie est alors mesurée permettant la détermination des positions atomiques. Il est également possible de distinguer expérimentalement deux conformères quasi-isoénergétique en déterminant leur moment dipolaire grâce à l'effet Stark [172].

Le but de ce chapitre est de comprendre l'importance des effets coopératifs dans la stabilité relative des différents conformères pour un nombre croissant de molécule d'eau. L'hydratation de la  $\beta$ -propiolactone (voir figure 5.2) a été étudiée expérimentalement par Pérez *et al.* [34].

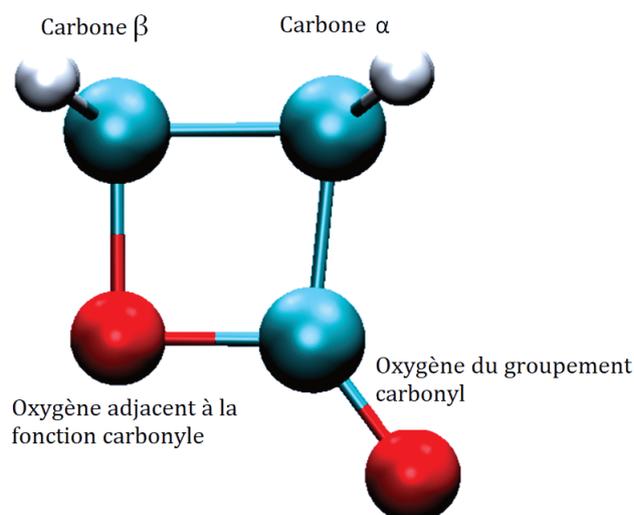


FIGURE 5.2 – Illustration de la  $\beta$ -propiolactone (BPL)

Cette molécule constitue un très bon prototype pour l'étude de différents types d'interactions :

- D'une part, elle permet de sonder des interactions "moyennes" ( $5 < E < 15$  kcal/mol) de type liaison hydrogène à forte composante électrostatique (O-H ... O=C ou O-C).
- Par ailleurs, ce soluté permet également de sonder les liaisons hydrogène plus faibles ( $E < 5$  kcal/mol) dont la composante dispersive constitue une part importante de

l'énergie d'interaction (C-H ...OH<sub>2</sub>).

Si ces dernières interactions sont plus faibles, elles peuvent jouer un rôle important dans la stabilité relative de différents conformères. Ceci a été démontré dans le cas du complexe oxirane :(H<sub>2</sub>O) [173]. Ces interactions sont cruciales dans de nombreux équilibres environnementaux, puisqu'on les retrouve dans les clathrates de méthane [174]. L'hydratation de la β-propiolactone a été caractérisée expérimentalement [34] par spectroscopie micro-ondes couplée à une expansion supersonique. Les structures des complexes BPL :(H<sub>2</sub>O)<sub>1-5</sub> ainsi que leur moment dipolaire ont été déterminés. Il s'agit de la raison majeure pour laquelle nous avons choisi ce composé comme prototype pour la mise en place d'une stratégie d'étude qui permettra à la fois de prédire et d'expliquer les structures des conformères les plus stables.

Etape 1 : localisation des zones électrophiles et nucléophiles (MEP)

Etape 2: orientation des deux partenaires de façon à faire coïncider une zone électrophile de la molécule d'eau supplémentaire avec une zone nucléophile du soluté et réciproquement.

Etape 3 : optimisation des structures construites et études des chemins réactionnels entre les minima de la surface d'énergie potentielle (PES)

Etape 4 : mise en évidence des interactions et des effets coopératifs par la théorie AIM.

FIGURE 5.3 – Stratégie proposée pour la recherche systématique des conformères hydratés les plus stables.

Une stratégie a récemment été proposée pour l'étude théorique des complexes mono-hydratés [7] les plus stables. Celle-ci repose principalement sur l'étude du potentiel électrostatique moléculaire des partenaires isolés, et sur des approches topologiques de la liaison chimique. Ici, nous proposons une stratégie similaire présentée dans la figure 5.3.

L'utilisation du potentiel électrostatique moléculaire (MEP) permet d'identifier des sites complémentaires d'interactions sur les deux partenaires A et B. Les interactions à longue distance conduiront le partenaire B à s'orienter de manière à faire coïncider au mieux ses zones nucléophiles avec les zones électrophiles du partenaire A et réciproquement. Cette étape conduit à l'identification de directions d'approches privilégiées

d'interactions, qui peuvent être assimilées aux interactions à longue distance entre les partenaires. Des structures initiales de complexes peuvent ainsi être proposées.

L'optimisation de ces structures initiales nous permet d'identifier tous les minima. Il est à noter que le choix du niveau de calcul pour l'optimisation des structures initiales est crucial pour construire une surface d'énergie potentielle la plus précise. Pour les complexes contenant de une à cinq molécules d'eau, l'approche MP2 reste encore envisageable. Nous l'avons donc sélectionnée pour l'étude des complexes BPL  $:(\text{H}_2\text{O})_{1-5}$ . Nous avons commencé par rechercher les minima locaux et globaux pour chacun des complexes. Puis, pour expliquer les structures des conformères déterminés expérimentalement, nous avons étudié une partie de la surface d'énergie potentielle et nous avons identifié les états de transition entre les minima les plus pertinents. Par ailleurs, nous avons discuté l'origine de la stabilité relative des différents conformères grâce l'analyse topologique AIM.

## 5.2 Caractérisation des régions nucléophiles et électrophiles de la BPL - étape 1

L'analyse du potentiel électrostatique moléculaire de la BPL conduit à l'identification de quatre zones électrophiles (indiquées par  $V_{\text{Smax}}$  sur la figure 5.4) :

- La première zone électrophile ( $V_{\text{Smax1}}$ ) est située sur l'axe de la liaison covalente C-H du carbone  $\alpha$ .
- Ensuite, la deuxième zone électrophiles ( $V_{\text{Smax2}}$ ) est localisée au niveau de l'axe la liaison covalente C-H, cette fois-ci au niveau du carbone  $\beta$ .
- La troisième zone électrophile ( $V_{\text{Smax3}}$ ) est situé le long de l'axe de la liaison C-O : selon la nomenclature de Politzer *et al.* il s'agit d'un  $\sigma$ -hole.
- Enfin, la quatrième et dernière zone électrophile ( $V_{\text{Smax4}}$ ) a été identifié au niveau du plan médian de la liaison C-C : ce type de zone électrophile situé perpendiculairement à l'axe d'une liaison chimique est communément appelé  $\pi$ -hole.

L'analyse du MEP de la BPL permet également d'identifier les deux zones nucléophiles situées à proximité des oxygènes (indiquées par  $V_{\text{Smin}}$  sur la figure 5.5) :

- L'oxygène du groupement carbonyle (indiqué par la flèche  $V_{\text{Smin1}}$ ) est la zone la plus nucléophile.
- L'autre zone nucléophile est située à proximité de l'oxygène adjacent à la fonction carbonyle.

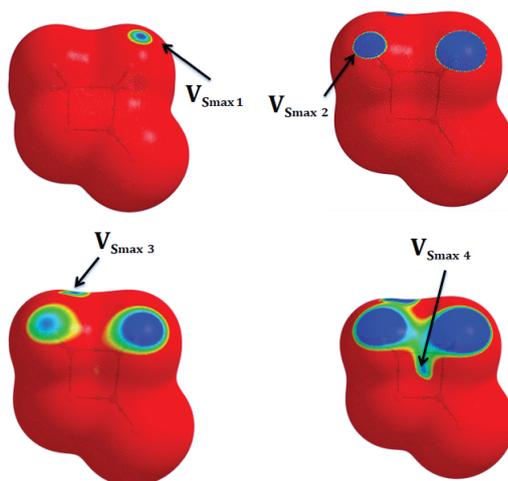


FIGURE 5.4 – Identification des zones électrophiles de la BPL sur une surface d’iso-densité électronique (0,001 u.a.).

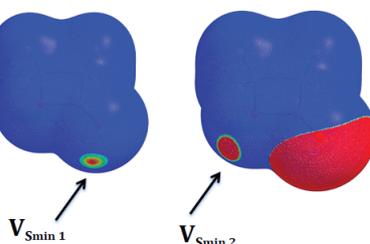


FIGURE 5.5 – Identification des zones nucléophiles de la BPL sur une surface iso-densité (0,001 u.a.)

Pour la molécule d’eau, il est évident que la zone nucléophile est située à proximité de l’oxygène tandis que les zones électrophiles sont situées à proximité des hydrogènes sur l’axe O-H. Nous avons toutefois indiqué ces zones par des flèches sur la figure 5.6.

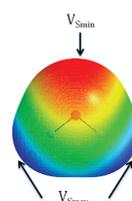


FIGURE 5.6 – Identification des maxima et du minimum du MEP de la molécule d’eau sur une surface d’iso-densité de 0,001 u.a.

Dans le tableau 5.1 nous avons rassemblé les valeurs du potentiel électrostatique sur les points remarquables identifiés dans le paragraphe précédent.

Nous observons que l'oxygène cétonique est plus nucléophile que l'autre oxygène de la BPL. Nous constatons également que l'oxygène du carbonyle a un caractère plus nucléophile que celui de la molécule d'eau. En revanche, la valeur maximale du MEP de la molécule d'eau est nettement supérieure à ceux de la BPL.

BPL		H <sub>2</sub> O	
V <sub>Smin1</sub> :	-34,51	V <sub>Smin</sub> :	-32,25
V <sub>Smin2</sub> :	-26,36		
V <sub>Smax1</sub> :	25,85	V <sub>Smax</sub> :	41,85
V <sub>Smax2</sub> :	24,79		
V <sub>Smax3</sub> :	24,77		
V <sub>Smax4</sub> :	23,22		

Tableau 5.1 – Valeurs des MEP sur une surface d'iso-densité électronique (0,001 u.a.) en kcal/mol.

Dans la section suivante, nous allons montrer comment la caractérisation des partenaires isolés peut conduire à la proposition de structures initiales, appelée également "guess structures", pour l'optimisation géométrique des conformères du complexe 1 :1 . Pour cela, nous allons recourir à la stratégie décrite en figure 5.3. Celle-ci consiste à orienter les partenaires de façon à faire coïncider une zone électrophile du soluté avec une zone nucléophile de la molécule d'eau et vice-versa.

## 5.3 Etude du complexe BPL(H<sub>2</sub>O).

### 5.3.1 Identification des structures initiales - étape 2

L'application de la stratégie employée par Politzer et Gadre [175] [176] conduit à la structure illustré sur la figure 5.7. En effet, cette structure est le résultat d'interaction entre la zone la plus nucléophile de la BPL et la zone la plus électrophile de l'eau (voir le tableau 5.1).

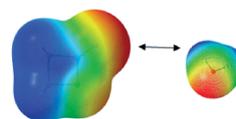


FIGURE 5.7 – Une structure initiale suggérée où la zone la plus nucléophile de la BPL avec la zone la plus électrophile de l'eau.

En revanche, notre stratégie consiste à considérer toutes les interaction plausibles entre les zones complémentaires des deux partenaires. La mise en application de cette démarche nous conduit à l'identification de neuf directions d'approche, comme l'indique la figure 5.8 :

Ces neuf approches peuvent se classer en trois catégories :

- La première catégorie est constituée par les structures guess-A, guess-B, guess-C, guess-D, guess-F, guess-G et guess-H. Ces sept approches vont conduire à des

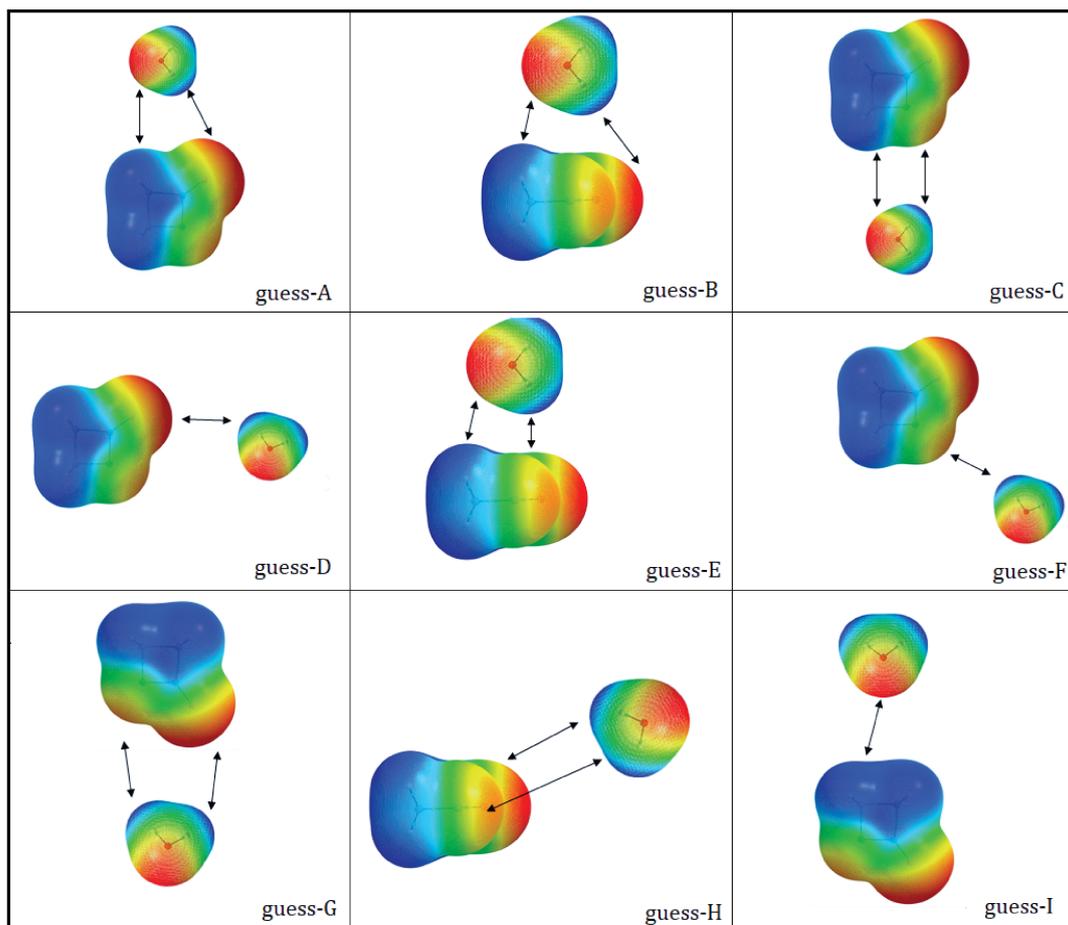


FIGURE 5.8 – Identification de neuf directions d’approches d’une molécule d’eau par rapport à la BPL à partir de l’analyse MEP.

interactions de type liaison hydrogène classique.

- La deuxième catégorie consiste à placer la molécule d’eau au niveau du  $\pi$ -hole en position "stack". La structure guess-E appartient à cette catégorie.
- Enfin, la troisième catégorie consiste en une interaction entre la molécule d’eau et le  $\sigma$ -hole de la BPL. Cette catégorie contient la structure guess-I.

La description des neuf structures "guess" est donnée ci-dessous :

- Les approches guess-A et guess-B dans lesquelles l’un des hydrogènes de la molécule d’eau est en interaction avec l’oxygène du groupement carbonyle C=O de la BPL. Dans les deux approches, l’oxygène de la molécule d’eau est en interaction avec un hydrogène porté par le carbone  $\alpha$ . L’approche guess-A est une approche dans laquelle la molécule d’eau est situé dans le même plan que celui de la  $\beta$ -propiolactone. L’approche guess-B est une approche particulière dans laquelle la molécule d’eau

est localisée dans un plan perpendiculaire au plan de symétrie défini par la BPL.

- Dans les approches guess-C et guess-F, la molécule d'eau est située dans le plan de la  $\beta$ -propiolactone tandis que dans la structure guess-E, la molécule d'eau est située dans un plan orthogonal au plan de la BPL. Dans ces trois structures (guess-C, guess-E et guess-F), l'interaction principale s'effectue entre l'oxygène adjacent à la fonction carbonyle de la BPL et un hydrogène de la molécule d'eau.
- Dans les approches guess-G et guess-H, la molécule d'eau joue le rôle de double donneur de liaison hydrogène avec les deux atomes d'oxygène de la BPL. Lors de l'approche guess-G, la molécule d'eau est située dans le plan de la BPL tandis que l'approche guess-H, la molécule d'eau est située dans un plan orthogonal au plan de la BPL.
- Quant à l'approche guess-I, l'interaction électrostatique se réalise entre le site électrophile du  $\sigma$ -hole de la BPL et l'atome d'oxygène de la molécule d'eau.

### 5.3.2 Etude de la PES pour les complexes BPL(H<sub>2</sub>O) - étape 3

A partir des structures initiales proposées dans la sous-section précédente (voir figure 5.8), nous avons effectué des optimisations de géométrie suivies de calculs de fréquences au niveau MP2/aug-cc-pVDZ. Ces calculs nous ont permis de mettre en évidence cinq minima sur la surface d'énergie potentielle. Les structures des cinq minima sont indiquées dans les figures 5.9 et 5.10. Sur ces figures, nous avons reporté les paramètres géométriques les plus importants ainsi que les écarts énergétiques par rapport au conformère le plus stable (nommé structure S1<sup>1</sup>) :

Il est intéressant de faire le lien entre les structures "guess" représentées sur la figure 5.8 et les minima représentés sur les figures 5.9 et 5.10.

- les approches guess-A et guess-B conduisent à la même structure optimisée nommée S1.
- Quant aux approches guess-E et guess-F mènent à la structure optimisée nommée S2.
- Les approches guess-G et guess-H mènent à la structure optimisée nommée S3.
- La structure S4 s'obtient à partir de la structure non-optimisée guess-D.
- Enfin, la structure initiale guess-I conduit après optimisation à la structure S5.

---

1. La lettre S renvoi au mot "single", autrement dit, il existe une seule molécule d'eau. Le chiffre indique l'ordre de stabilité.

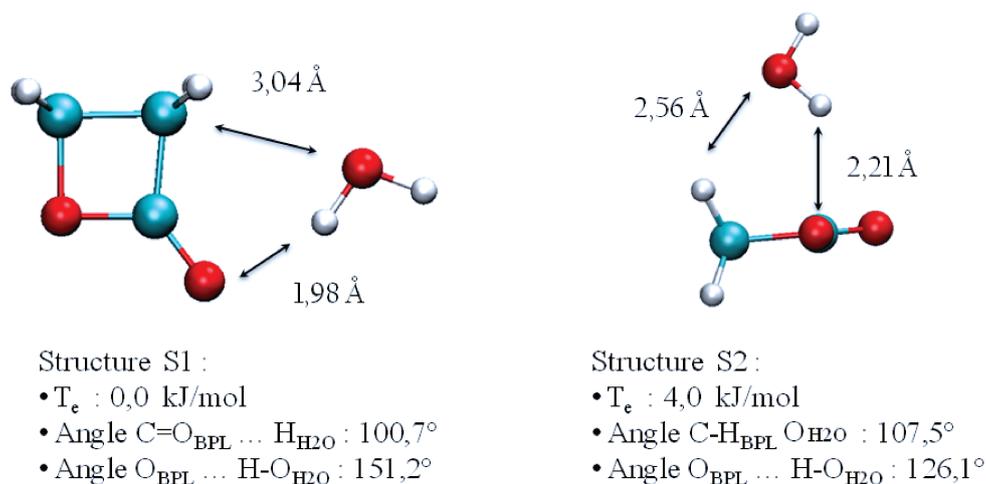


FIGURE 5.9 – Illustration des minima de la SEP pour les complexes S1 et S2.

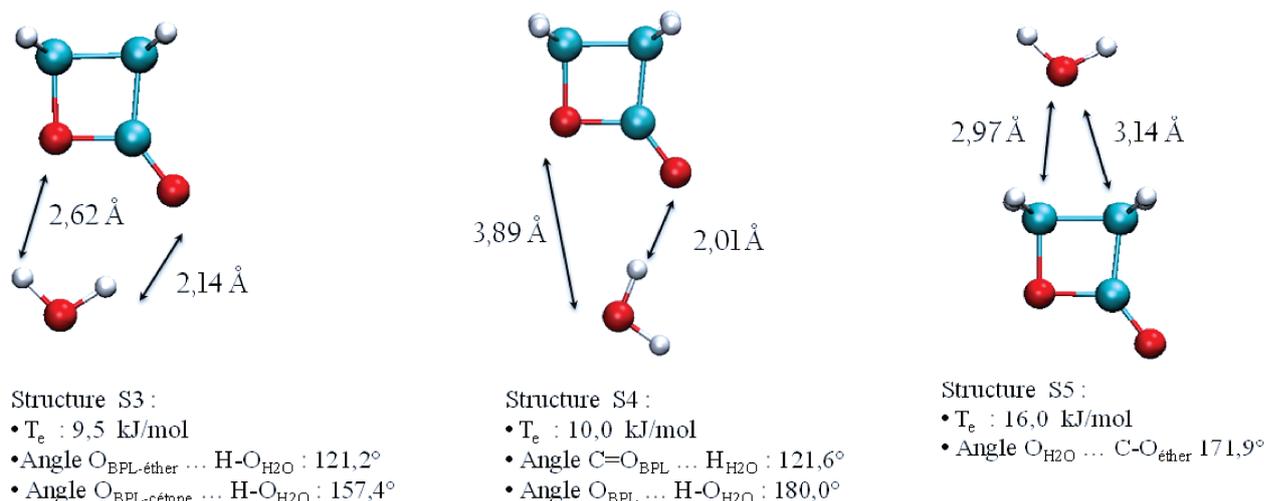


FIGURE 5.10 – Illustration des minima de la SEP pour les complexes S3, S4 et S5.

Il est à noter que dans la structure S1 (voir figure 5.9), la molécule d'eau est située dans le plan de la BPL tandis que l'étude théorique menée par Perez *et al.* [34] au niveau MP2/6-311++G(d,p) suggère que la molécule d'eau n'est pas située sur ce même plan. Ce résultat est dû à l'emploi d'une base ayant moins de fonctions diffuses que la base aug-cc-pVDZ.

Un regard attentif sur les données géométriques portées sur les figures 5.9 et 5.10 nous permet de mieux apprécier les différences structurales :

- Les structures S1, S3 et S4 sont caractérisées par une liaison hydrogène entre l'hydrogène de la molécule d'eau et l'oxygène du groupement carbonyle. Comme

l'indiquent les figures 5.9 et 5.10, les distances entre l'hydrogène de la molécule d'eau et l'oxygène du groupement carbonyle sont respectivement de 1,98 ; 2,14 et 2,01 Å. Selon la classification proposée par Jeffrey [9], elles appartiennent donc à la catégorie des liaisons hydrogènes dites "modérées" : les distances de liaison hydrogènes étant toutes comprises entre 1,5 et 2,2 Å. De même, les angles de liaison hydrogène caractéristiques sont compris dans l'intervalle 130 - 180° proposé par Jeffrey pour cette catégorie.

Cependant, le minimum S3 se distingue du minima S4 par la présence d'une autre liaison hydrogène, impliquant l'oxygène adjacent à la fonction carbonyle. Comme l'indique la figure 5.10, la distance de cette liaison hydrogène est de 2,62 Å. Cet interaction secondaire pourrait être responsable d'un allongement de la liaison hydrogène entre la molécule d'eau et l'oxygène du groupement carbonyle.

- Le minimum S2 est le seul complexe pour lequel la molécule d'eau n'est pas dans le plan de la BPL. Ce complexe est caractérisé par une liaison hydrogène non-conventionnelle. En effet, si on considérait uniquement la distance  $H_{H_2O} - O_{BPL}$  (2,21Å), on pourrait conclure qu'il s'agit d'une liaison hydrogène modéré au sens de Jeffrey. Or l'angle de liaison qui est de 126.1° n'appartient pas à l'intervalle suggéré par Jeffrey.
- Enfin, dans le complexe S5, l'oxygène de la molécule d'eau interagit avec le  $\sigma$ -hole de la BPL. La distance d'interaction entre l'oxygène de la molécule d'eau et le carbone situé à proximité du  $\sigma$ -hole est de 2,97Å. L'angle d'interaction  $O_{eau} \widehat{C}_\beta O_{ether}$  est proche de 180°.

Nous avons rassemblé les caractéristiques énergétiques des complexes étudiés dans le tableau 5.2. L'énergie relative des composés est indiquée par le symbole  $T_e$  en kJ/mol.  $T_e + ZPE$  représente cette énergie relative corrigé de l'énergie de point zéro. Les grandeurs  $D_e$  et  $C_e$  représentent respectivement l'énergie de formation et de complexation. La correction due à l'effet de superposition de base (BSSE), calculée avec la méthode de Boys et Bernardi [177], est également reportée dans le tableau 5.2 et notée  $D_E^{CP}$ .

MP2/aVDZ	$T_e$	$T_e + ZPE$	$C_e$	$C_e^{CP}$	$D_e$	$D_e^{CP}$
S1	0,0	0,0	-30,9	-24,8	-30,2	-24,1
S2	4,0	2,8	-26,2	-20,1	-26,2	-20,1
S3	9,5	7,9	-20,7	-17,1	-20,7	-17,1
S4	10,0	8,4	-20,2	-16,9	-20,2	-16,9
S5	16,2	11,4	-14,0	-10,2	-14,0	-10,2

Tableau 5.2 – Caractéristiques énergétiques des complexes 1 : 1 en kJ/mol.

Le complexe S1 est le minimum global sur la surface d'énergie potentielle alors que les quatre autres complexes (S2, S3, S4 et S5) sont des minima locaux. Afin de conforter ces résultats, nous avons effectué un calcul d'optimisation des cinq complexes au niveau MP2/aug-cc-pVTZ dont les résultats sont regroupés dans le tableau 5.3. On peut remarquer que l'utilisation d'une base plus étendue a pour effet d'accroître l'écart énergétique entre les conformères S1 et S2 et de diminuer les écarts énergétiques entre les conformères S1, S3 et S4.

MP2/aVTZ	Te	Te + ZPE	C <sub>e</sub>	C <sub>e</sub> <sup>CP</sup>	D <sub>e</sub>	D <sub>e</sub> <sup>CP</sup>
S1	0,0	0,0	-29,6	-26,9	-28,8	-26,1
S2	5,0	3,7	-24,3	-21,3	-23,9	-21,0
S3	8,5	7,5	-20,6	-18,4	-20,3	-18,1
S4	9,1	7,7	-20,1	-18,1	-19,8	-17,7
S5	16,6	12,0	-12,3	-10,9	-12,2	-10,8

Tableau 5.3 – Caractéristiques énergétiques des complexes étudiés en kJ/mol, au niveau MP2/aVTZ.

Le conformère S1 est le seul qui a été identifié pour le complexe 1 : 1 à partir des données expérimentales [34]. Or, nos données théoriques mettent en évidence cinq minima. Comment peut-on expliquer, sur la base de notre étude théorique, la mise en évidence expérimentale d'un seul conformère ? Pour répondre à cette question, nous avons étudié les états de transition permettant le passage d'un conformère à un autre.

La figure 5.11 montre que la barrière énergétique la plus élevée n'est que de 2.8 kJ/mol (barrière permettant le passage de la structure S2 à la structure S1). Or, les énergies de formations des complexes (répertoriées dans le tableau 5.2) sont nettement supérieures à cette barrière énergétique. Nous pouvons donc supposer que différents conformères pourraient se former, puis se réarranger pour conduire au conformère le plus stable dans le jet supersonique formé en amont de la caractérisation ro-vibrationnelle [34].

La structure géométrique du complexe le plus stable énergétiquement S1, ne satisfait pas les règles empiriques de Legon et Millen [178]. En effet, Legon et Millen ont fourni des règles empiriques pour la prédiction de structures de complexe à liaison hydrogène, B...HX en phase gazeuse. Ici B est une molécule possédant une région nucléophile (doublets non liants ou électrons  $\pi$ ). "H-X" est la molécule engageant son hydrogène dans la liaison, l'hydrogène étant la région électrophile de la molécule. Les règles énoncées dans la référence [178] peuvent être résumées de la façon suivante :

- l'axe de la molécule H-X coïncide avec l'axe du doublet non-liants de la molécule B.
- Si le groupement B n'a pas de doublet non-liant, mais des électrons  $\pi$ -polarisables

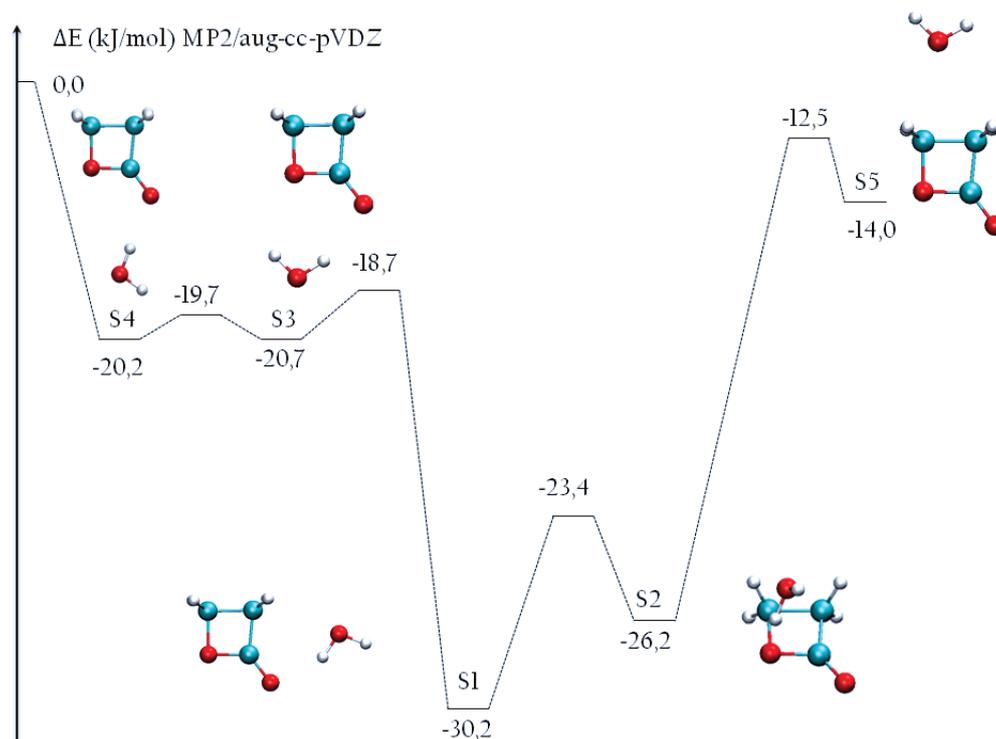


FIGURE 5.11 – Étude de la surface d'énergie potentielle associée au complexe BPL  $:(\text{H}_2\text{O})$ . Les différences énergétiques sont calculées par rapport à la somme des énergies électroniques des molécules formant le complexe. ( $D_e$ )

alors il existe une intersection entre l'axe H-X et l'axe inter-nucléaire des atomes formant la liaison  $\pi$ . De plus, l'axe de la molécule H-X est plus ou moins perpendiculaire au plan de symétrie de la liaison  $\pi$ .

Pour pouvoir appliquer rigoureusement les règles de Legon et Millen, il faut donc localiser les doublets non-liants. Pour cela nous allons étudier les minima de la fonction MEP<sup>2</sup>. Ainsi, nous pouvons rendre compte d'un schéma de type Lewis mettant en évidence des doublets non-liants non équivalents comme l'indique la figure 5.12 :

2. Il ne s'agit en aucun cas de la valeur du MEP projetée sur une surface iso-densité.

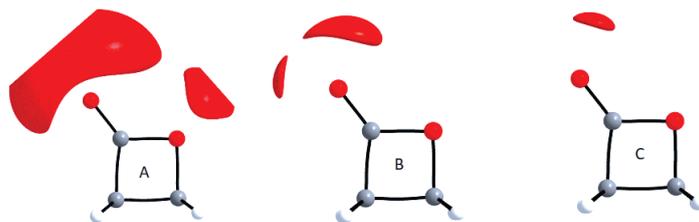


FIGURE 5.12 – Mise en évidence d’un schéma de type Lewis par l’étude des minima de la fonction MEP. Les valeurs des iso-surface sont (de la gauche vers la droite)  $-0,040$  ;  $-0,060$  et  $-0,065$  u.a.

Ces régions nucléophiles coïncident avec l’emplacement des doublets non-liant dans un schéma de type Lewis. Dans ce modèle, le poids de chaque doublet peut être estimé à partir de la valeur du MEP. Lorsque la valeur de l’iso-surface du MEP est de  $-0,040$  u.a. (figure 5.12-A), on met en évidence deux régions nucléophiles localisées à proximité des deux oxygènes de la BPL. Lorsque l’on diminue la valeur de l’iso-surface jusqu’à ce qu’elle atteigne une valeur de  $-0,060$  u.a. (figure 5.12-B), la région nucléophile relative à l’oxygène adjacent au groupement carbonyle disparaît. Pour cette même valeur de l’iso-surface, on met en évidence deux sous-régions nucléophiles à proximité de l’oxygène cétonique. Enfin, pour une valeur, de l’iso-surface égale à  $-0,065$  u.a. (figure 5.12-C), on peut visualiser la région la plus nucléophile. Dans le modèle de Legon et Millen, la molécule d’eau devrait établir une interaction via son hydrogène (le site le plus électrophile) avec la zone la plus nucléophile de la BPL aboutissant aux structures S3 et S4. Or, il ne s’agit pas des complexes les plus stables. A ce stade, nous souhaitons comprendre pourquoi la structure S1 correspond la plus stable. Cette problématique sera abordée dans la sous-section suivante dédiée à la caractérisation topologique des interactions non-covalentes dans le cadre la théorie QTAIM.

### 5.3.3 Caractérisation QTAIM des conformères BPL :H<sub>2</sub>O - étape 4

Le tableau 5.4 regroupe les deux informations topologiques, la densité électronique ( $\rho(r)$ ) et son laplacien ( $\nabla^2\rho(r)$ ) aux points critiques relatifs des interaction non covalentes.

Pour illustrer les points critiques, nous présentons les graphes moléculaires sur la figure 5.13.

Complexes 1 :1	$\rho(r)$ u.a.	$\nabla^2\rho(r)$ u.a.	Atomes impliqués dans l'interaction
S1	0,023	0,090	$H_{H_2O}-O_{\text{cétone}}$
	0,006	0,030	$O_{H_2O}-C_{\alpha}$
S2	0,016	0,035	$H_{H_2O}-O_{\text{éther}}$
	0,008	0,050	$O_{H_2O}-H_{\beta}$
S3	0,007	0,025	$H_{H_2O}-O_{\text{éther}}$
	0,016	0,051	$H_{H_2O}-O_{\text{cétone}}$
S4	0,020	0,071	$H_{H_2O}-O_{\text{cétone}}$
S5	0,007	0,030	$\sigma$ -hole

Tableau 5.4 – Caractéristiques AIM des complexes 1 : 1, l'indice éther réfère à l'oxygène adjacent

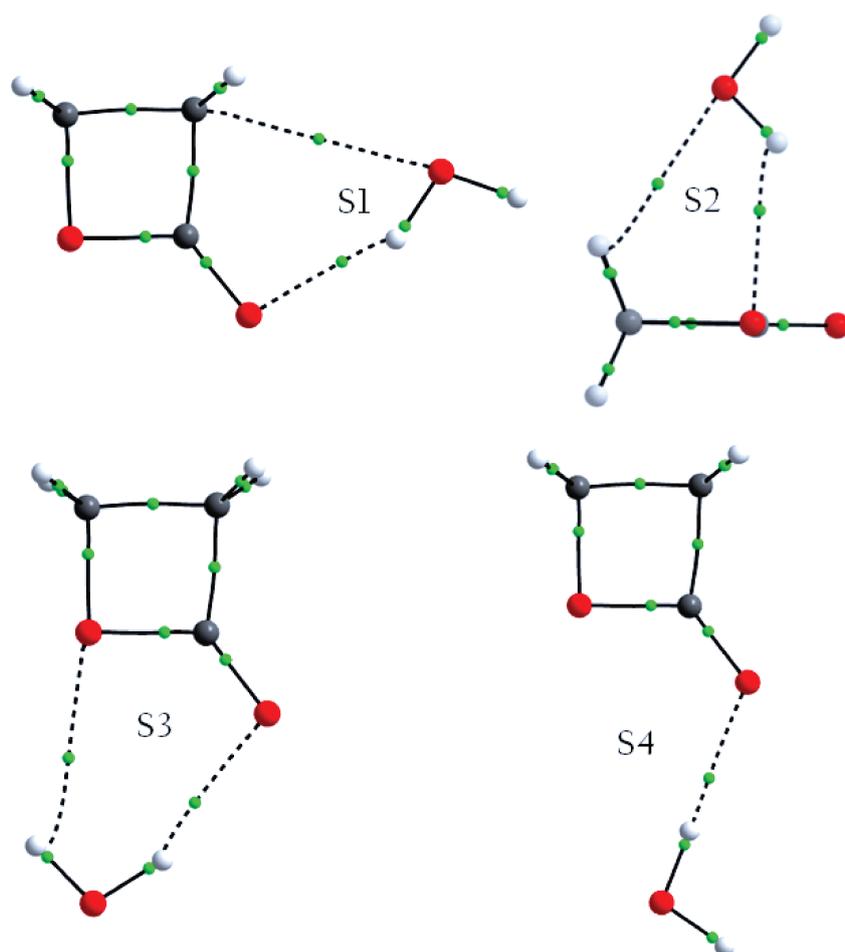


FIGURE 5.13 – Graphs AIM des complexes 1 : 1. Les sphères vertes indiquent les points critiques de liaisons (BCP) et les lignes noirs correspondent aux chemins de liaisons

L'ensemble des résultats nous permet d'identifier clairement les interactions intermoléculaires dans les conformères.

- Dans le complexe S1, il existe d'une part, une interaction de type liaison hydrogène, impliquant l'atome d'oxygène de la fonction carbonyle et l'un des atomes d'hydrogène de la molécule d'eau. Cette interaction est caractérisée par un BCP dont la densité électronique est de 0.023 u.a. avec une valeur du laplacien de 0.090 u.a. D'autre part, il existe une interaction secondaire impliquant l'atome d'oxygène de la molécule d'eau et l'atome C<sub>α</sub> de la BPL. La valeur de la densité électronique et de son laplacien au point critique de liaison sont respectivement de 0,006 et 0,030 u.a.
- Dans le complexe S2, nous avons une interaction entre l'un des atomes d'hydrogène de la molécule d'eau et l'atome d'oxygène adjacent à la fonction carbonyle caractérisé par une densité électronique de 0.016 et un laplacien valant 0.035 u.a. Il existe également une interaction secondaire impliquant l'atome d'oxygène de la molécule d'eau et l'hydrogène porté par l'atome de carbone<sub>α</sub> de la BPL. Le point critique correspondant caractérisé par une densité électronique de 0,008 u.a. et un laplacien valant 0,050 u.a.
- Dans le complexe S3, chaque atome d'hydrogène de la molécule d'eau interagit avec l'un des atomes d'oxygène de la BPL. L'interaction avec l'oxygène du groupement carbonyle est la plus forte, la valeur de la densité électronique au point critique de liaison étant de 0.016 u.a (la valeur du laplacien est 0.051 u.a). La valeur de la densité électronique au point critique de liaison caractérisant l'interaction avec l'autre oxygène est de 0.007 u.a. (la valeur du laplacien est 0.025 u.a).
- Dans le complexe S4, une interaction intermoléculaire de type liaison hydrogène est identifiée entre un atome d'hydrogène d'une molécule d'eau et l'atome d'oxygène de la fonction carbonyle. Les valeurs de la densité électronique et de son laplacien au point critique de liaison caractérisant l'interaction sont de 0.020 et 0.071 u.a., respectivement.

Les complexes S1 et S2 sont caractérisés d'une part, par une interaction primaire de type liaison hydrogène - impliquant l'hydrogène de la molécule d'eau avec l'un des oxygène de la BPL et d'autre part, par une interaction secondaire impliquant l'oxygène de la molécule d'eau. Dans ces deux complexes, la molécule d'eau cède une partie de sa densité électronique et reçoit une fraction de la densité électronique provenant de la BPL, ce qui est synonyme dans la littérature d'effet coopératif [179], la présence d'interaction secondaire semble renforcer la liaison hydrogène présente dans les complexes S1 et S2. En ce qui concerne le complexe S3 - complexe qui pourrait être obtenue suite à l'emploi des règles de Legon et Millen [166]. La présence d'interaction secondaire semble affaiblir la liaison hydrogène principale présente dans le complexe S3. Dans le complexe S4, il n'existe

pas d'interaction secondaire, et par conséquent, la présence d'effet coopératif ou d'effet anti-coopératif est impossible.

## 5.4 Etude du complexe BPL(H<sub>2</sub>O)<sub>2</sub>.

### 5.4.1 Caractérisation des régions nucléophiles et électrophiles des complexes BPL(H<sub>2</sub>O) et du dimère d'eau - étape 1

La recherche des directions privilégiées d'approches pour la formation des complexes 1 : 2 peut se faire de deux façons complémentaires : on peut considérer d'une part le dimère d'eau en interaction avec la BPL, ou d'autre part, l'un des conformères mono-hydratés en interaction avec une molécule d'eau supplémentaire. Cette sous-section se propose d'identifier les régions nucléophiles et électrophiles des complexes BPL(H<sub>2</sub>O) et du dimère d'eau à partir de l'étude du MEP sur une surface d'iso-densité de  $\rho = 0,001$  u.a. La figure 5.14 montre les points remarquables de la valeur du MEP du complexe S1 :

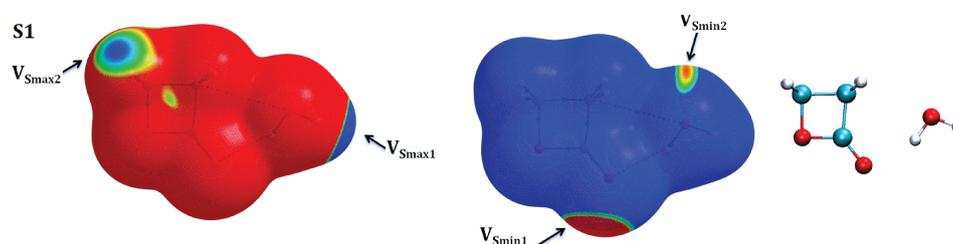


FIGURE 5.14 – Points remarquables sur le MEP du complexe S1.

La première zone électrophile est située sur l'axe de la liaison covalente O-H de la molécule d'eau (indiquée par une flèche  $V_{S_{max1}}$  sur la figure 5.14) dont la valeur est de 37,7 kcal/mol. La deuxième zone électrophile ( $V_{S_{max2}} = 24,8$  kcal/mol) est localisée au niveau de l'axe la liaison covalente C-H du carbone  $\beta$ . Quant aux zones nucléophiles, elles sont également identifiées sur la figure 5.14. La zone la plus nucléophile ( $V_{S_{min1}} = -30,2$  kcal/mol), se situe à proximité de l'oxygène du groupement carbonyle alors que la deuxième zone la plus nucléophile est localisée au niveau de l'oxygène de la molécule d'eau ( $V_{S_{min2}} = -22,6$  kcal/mol).

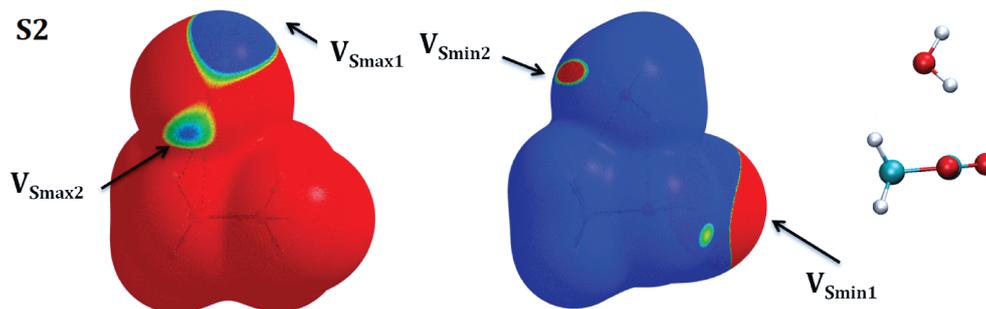


FIGURE 5.15 – Identification des deux zones les plus électrophiles et les plus nucléophiles sur le complexe S2.

Dans le cas du complexe S2 (figure 5.15), les zones électrophiles et nucléophiles les plus importantes sont indiquées sur la figure 5.15 :

- La zone la plus électrophile ( $V_{S_{\max 1}} = 43,4$  kcal/mol) est située le long de l'axe de la liaison covalente O-H de la molécule d'eau et est situé à proximité de l'hydrogène libre (celui qui n'est pas impliqué dans la liaison hydrogène). Cette zone est indiquée par une flèche  $V_{S_{\max 1}}$  sur la figure 5.15. La deuxième zone la plus électrophile ( $V_{S_{\max 2}} = 27,0$  kcal/mol) est localisée au niveau de l'axe la liaison covalente O-H de la molécule d'eau. Il s'agit cette fois-ci de l'autre hydrogène (impliqué dans l'interaction de type liaison hydrogène). Cette zone est indiquée par une flèche  $V_{S_{\max 2}}$  sur la figure 5.15.
- La zone indiquée par une flèche  $V_{S_{\min 1}}$  (-31,4 kcal/mol) est la zone la plus nucléophile. Quant à la deuxième zone la plus nucléophile, elle se situe au niveau de l'oxygène de la molécule d'eau ( $V_{S_{\min 2}} -22,6$  kcal/mol).

Nous avons également indiqué dans la figure 5.16, les zones les plus électrophiles et les plus nucléophiles des complexes S3, S4 et S5 :

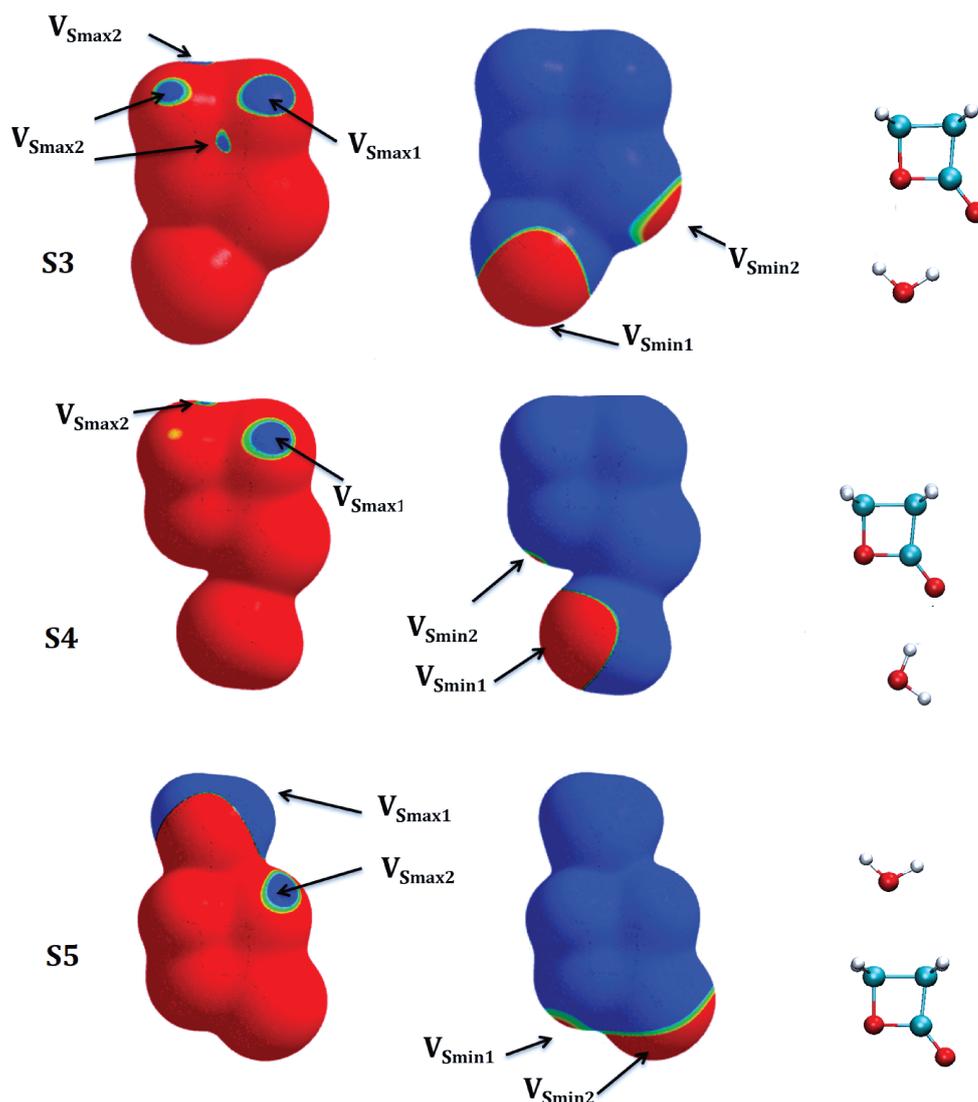


FIGURE 5.16 – Identification des zones les plus électrophiles et les plus nucléophiles sur les complexes S3, S4 et S5.

Les valeurs du MEP des points remarquables sont données dans le tableau 5.5 :

	S1	S2	S3	S4	S5	(H <sub>2</sub> O) <sub>2</sub>	BPL	H <sub>2</sub> O
			régions nucléophiles					
$V_{Smin1}$ :	-30,2	-31,4	-45,4	-46,0	-38,9	-39,9	-34,5	-32,3
$V_{Smin2}$ :	-26,7	-22,6	-24,6	-24,5	-30,7	-20,3	-26,4	
			régions électrophiles					
$V_{Smax1}$ :	37,7	43,4	31,5	30,0	53,8	51,3	25,9	41,9
$V_{Smax2}$ :	24,8	27,0	(30,1 - 30,9)	(27,0 - 27,3)	21,7	28,3	24,8	

Tableau 5.5 – Valeurs des MEP sur une surface d'iso-densité (0,001 u.a.) en kcal/mol.

Nous allons maintenant nous intéresser aux minima et maxima du MEP sur le dimère d'eau. Les zones les plus nucléophiles et électrophiles sont indiquées ci-contre dans la figure 5.17.

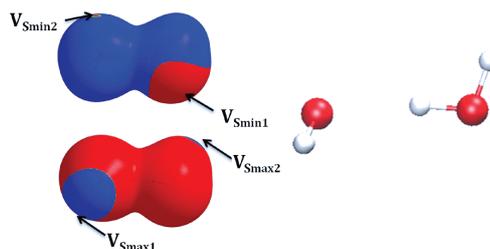


FIGURE 5.17 – Identification des maxima et des minima du MEP du dimère d'eau sur une surface d'iso-densité électronique de 0,001 u.a.

Le tableau 5.5 recense également les valeurs des minima et maxima du MEP projetées sur une surface d'iso-densité des points remarquables indiqués sur la figure 5.17. Nous avons également rappelé sur ce même tableau les valeurs des minima et maxima du MEP de la BPL et de la molécule d'eau. Il est intéressant de voir les effets de la présence d'une molécule d'eau sur la nucléophilicité et l'électrophilicité de la BPL et réciproquement :

Pour le conformère S1, la présence de la molécule d'eau diminue la nucléophilicité de l'oxygène du groupement carbonyle (la valeur est de -34,5 kcal/mol dans la BPL isolée et est de -30,2 kcal/mol dans le complexe S1). Quant à la BPL, elle induit une diminution légère de la nucléophilicité de l'oxygène de la molécule d'eau (la valeur est de -32,3 kcal/mol dans la molécule d'eau isolée et est de -26,7 kcal/mol dans le complexe S1). La BPL induit également une diminution de l'électrophilicité de l'hydrogène non engagé dans la liaison hydrogène (la valeur est de 41,9 kcal/mol dans la molécule d'eau isolée et est de 37,7 kcal/mol dans le complexe S1)

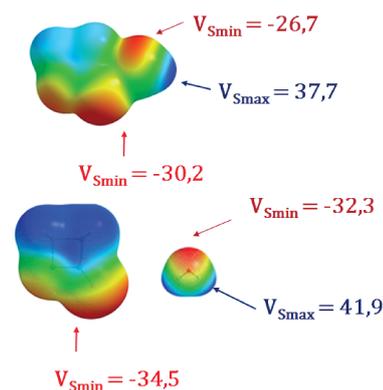


FIGURE 5.18 – Comparaison des points remarquables du MEP sur une surface d'iso-densité de 0.001 u.a. pour le complexe S1 et les partenaires isolés. Les valeurs des MEP sont données en kcal/mol.

En ce qui concerne le complexe S2, l'effet de la molécule d'eau se traduit par une diminution de la nucléophilicité de l'oxygène du groupement carbonyle. En effet, la valeur est de -34,5 kcal/mol dans la BPL isolée et est de -31,4 kcal/mol dans le complexe S2. Quant à la BPL, elle induit une diminution de la nucléophilicité de l'oxygène de la molécule d'eau puisque la valeur est de -32,3 kcal/mol dans la molécule d'eau isolée et est de -22,6 kcal/mol dans le complexe S2. Enfin, la BPL induit également une augmentation légère de l'électrophilicité de l'hydrogène non engagé dans la liaison hydrogène (la valeur est de 41,9 kcal/mol dans la molécule d'eau isolée et est de 43,4 kcal/mol dans le complexe S2).

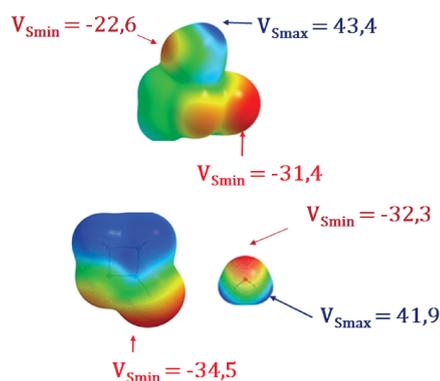


FIGURE 5.19 – Comparaison des points remarquables du MEP sur une surface d'iso-densité de 0.001 u.a. pour le complexe S2 et les partenaires isolés. Les valeurs des MEP sont données en kcal/mol.

Intéressons-nous au complexe S3, la présence de la molécule d'eau cause une diminution de la nucléophilicité de l'oxygène du groupement carbonyle puisque la valeur du  $V_{Smin}$  est de -34,5 kcal/mol dans la BPL isolée alors qu'elle est de -24,6 kcal/mol dans le complexe S3. En revanche, une augmentation de l'électrophilicité des hydrogènes portés par le carbone  $\alpha$  due à la présence de la molécule d'eau est observée (la valeur du  $V_{Smax}$  croît de 5,7 kcal/mol). En revanche, la présence de la BPL est responsable d'une augmentation de la nucléophilicité de l'oxygène de la molécule d'eau (la valeur est de -32,3 kcal/mol dans la molécule d'eau isolée et est de -45,4 kcal/mol dans le complexe S3).

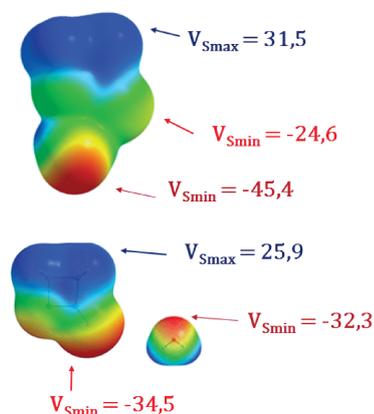


FIGURE 5.20 – Comparaison des points remarquables du MEP sur une surface d'iso-densité de 0.001 u.a. pour le complexe S3 et les partenaires isolés. Les valeurs des MEP sont données en kcal/mol.

Pour le cas du complexe S4, la présence de la molécule d'eau entraîne une augmentation de l'électrophilicité des hydrogènes portés par le carbone  $\alpha$  (la valeur du  $V_{S_{max}}$  croît de 4,1 kcal/mol). Néanmoins, la molécule d'eau est responsable d'une diminution légère de la nucléophilicité de l'oxygène adjacent à la fonction carbonyle (la valeur du  $V_{S_{max}}$  croît de 1,9 kcal/mol). La BPL induit une augmentation de nucléophilicité de l'oxygène de la molécule d'eau (la valeur du  $V_{S_{min}}$  décroît de 13,7 kcal/mol).

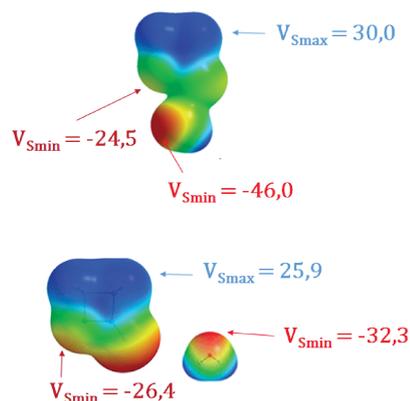


FIGURE 5.21 – Comparaison des points remarquables du MEP sur une surface d'iso-densité de 0.001 u.a. pour le complexe S4 et les partenaires isolés. Les valeurs des MEP sont données en kcal/mol.

Enfin, pour le conformère S5, la présence de la molécule d'eau semble impacter de manière significative la nucléophilicité des deux oxygènes de la BPL. Pour le cas de l'oxygène du groupement carbonyle, la valeur du  $V_{S_{min}}$  décroît de 6,6 kcal/mol. La valeur du  $V_{S_{min}}$  associé à l'autre oxygène décroît de 4,3 kcal/mol. La présence de la molécule d'eau entraîne une diminution de l'électrophilicité des hydrogènes portés par le carbone  $\alpha$  (diminution de 4,1 kcal/mol de valeur du MEP). La présence de la BPL rend les hydrogènes de la molécule d'eau plus électrophiles (augmentation de 11,9 kcal/mol de la valeur du  $V_{S_{max}}$ ).

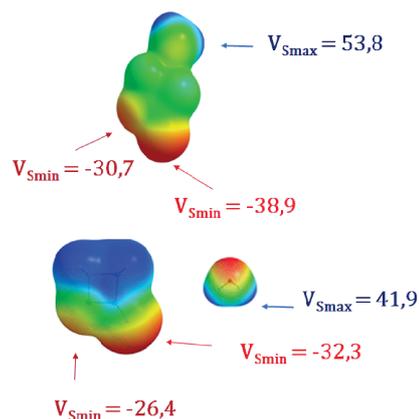


FIGURE 5.22 – Comparaison des points remarquables du MEP sur une surface d'iso-densité de 0.001 u.a. pour le complexe S5 et les partenaires isolés. Les valeurs des MEP sont données en kcal/mol.

## 5.4.2 Proposition de structures initiales pour le complexe dihydraté - étape 2

L'étape 2 de notre stratégie consiste à considérer toutes les interactions plausibles entre les zones complémentaires des deux partenaires. La mise en application de cette démarche nous a permis d'identifier vingt-et-une structures initiales que nous pouvons

classer en trois catégories :

- BPL + (H<sub>2</sub>O)<sub>2</sub> : en partant de la BPL et du dimère d'eau comme "briques" de départ, nous proposons quatre structures guess' (voir figure 5.23),
- BPL : (H<sub>2</sub>O) + (H<sub>2</sub>O) : l'addition d'une seconde molécule d'eau à un site nucléophile ou électrophile du complexe mono-hydraté, nous a permis d'identifier huit structures que nous avons regroupées dans la figure 5.24,
- BPL : (H<sub>2</sub>O) + (H<sub>2</sub>O) : l'addition d'une seconde molécule d'eau à un site nucléophile ou électrophile de la molécule d'eau présente dans le complexe 1 : 1, nous a permis d'identifier neuf structures que nous avons regroupées dans la figure 5.25.

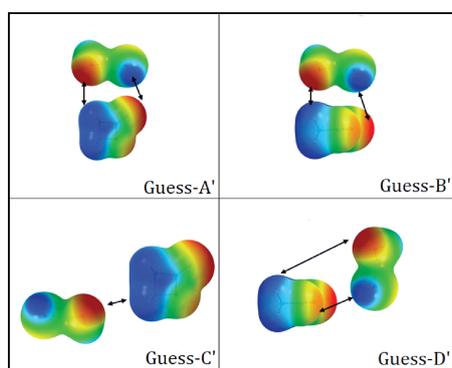


FIGURE 5.23 – Identification de quatre directions d'approches d'un dimère d'eau en interaction avec la BPL à partir l'analyse MEP.

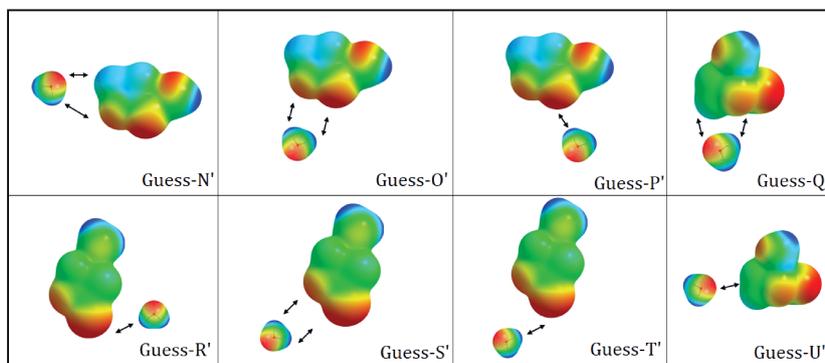


FIGURE 5.24 – Identification de huit directions d'approches à partir d'un complexe mono-hydraté en interaction avec une molécule d'eau.

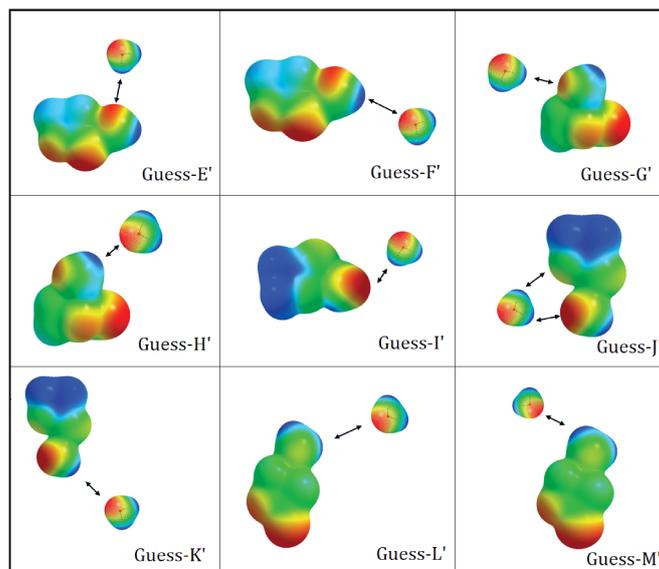


FIGURE 5.25 – Identification de neuf directions d’approches à partir d’un complexe mono-hydraté en interaction avec une molécule d’eau.

### 5.4.3 Etude de la PES pour les complexes di-hydratés - étape 3

La poursuite de notre étude sur les complexes di-hydratés s’est faite en optimisant la géométrie des structures proposées à l’issue de l’étape 2 au niveau MP2/aug-cc-pVDZ. Les calculs de fréquences ont permis de mettre en évidence quinze minima sur la surface d’énergie potentielle. Ces quinze minima ont été classés en trois catégories (voir tableau 5.6). La classification a été faite en tenant compte de l’erreur de la méthode MP2/aug-cc-pVDZ sur les énergies de formation des liaisons hydrogènes dans l’ensemble S22 [151], estimée à 7,1 kJ/mol. On regroupe dans une même famille les complexes dont la différence énergétique est de l’ordre de 7 kJ/mol, ce qui nous conduit à distinguer trois familles de conformères :

- la première famille (de 0,0 à 4,0 kJ/mol) est constituée d’un dimère d’eau en interaction avec le soluté via la fonction carbonyle et chaque molécule formant le complexe est à la fois donneuse et accepteuse de liaison hydrogène. Elle est constituée par les complexes D1 à D4. Cette famille de conformères contient les complexes les plus stables. Un regard attentif au tableau 5.6 permet de constater que les écarts énergétiques entre les complexes formant cette famille sont inférieurs à l’erreur de la méthode. Ces quatre conformères peuvent donc être considérer comme étant iso-énergétique à ce niveau de théorie.
- La deuxième famille de conformères (de 16,9 à 22,9 kJ/mol) contient des complexes dans lesquels les deux molécules d’eau n’interagissent pas entre elles, mais

seulement avec le soluté, via des sites autres que le  $\sigma$ -hole de la BPL. Il s'agit des complexes S1-S2, S2-S2, S1-S3 et S1-S4. Cette famille contient également le complexe D5, dans lequel un dimère d'eau interagit avec le soluté via la fonction carbonyle mais dans lequel une molécule d'eau joue le rôle de double donneur d'hydrogène. En se référant au tableau 5.6, on constate que les écarts énergétiques entre les complexes formant cette famille sont inférieurs à l'erreur de la méthode.

- Enfin, dans la troisième famille de conformères (de 25,9 à 36,0 kJ/mol), une molécule d'eau interagit avec le  $\sigma$ -hole, c'est-à-dire qu'il existe une molécule d'eau en position S5. Dans cette famille de complexes, certains conformères sont constitués d'un dimère d'eau en interaction avec le  $\sigma$ -hole, les autres étant constitués de deux molécules d'eau séparées interagissant avec la BPL. Nous avons regroupé tous les complexes restants dans cette famille.

Nous nous sommes uniquement intéressés à la première famille de conformères puisqu'elle contient les complexes les plus stables qui sont donc susceptibles d'être observés expérimentalement.

Un regard attentif au tableau 5.7 permet de confirmer que le conformère D1 constitue le minimum global. Il semblerait que la méthode MP2/aug-cc-pVDZ suffise à rendre compte de l'écart entre les énergies électroniques des complexes di-hydratés. Nous pouvons donc étudier les états de transition permettant le passage des conformères D2, D3 et D4 au conformère D1 au niveau MP2/aug-cc-pVDZ.

Les résultats présentés à la figure 5.26 permettent de constater que la barrière énergétique la plus élevée n'est que de 4,4 kJ/mol (barrière permettant le passage de la structure D1 à la structure D3). Or, les énergies de formations des complexes (voir tableau 5.8) sont nettement supérieures à cette barrière énergétique. Nous pouvons donc supposer que différents conformères pourraient se former, puis se réarranger pour conduire au conformère le plus stable, dans les conditions expérimentales, comme dans le cas du complexe mono-hydraté

Nous avons également comparé les moments dipolaires et les constantes rotationnelles obtenues expérimentalement à ceux des complexes D1 à D4 obtenus aux niveaux MP2/aug-cc-pVDZ et MP2/aug-cc-pVTZ. Le tableau 5.9 consigne les erreurs moyennes en valeur absolue par rapport aux résultats expérimentaux.

Le conformère D1 est le seul complexe 1 : 2 qui a été identifié expérimentalement [34]. Il s'avère que nos données théoriques mettent en évidence quatre complexes quasi-isoénergétiques. Est-il possible d'expliquer sur la base de notre étude théorique l'observation d'un seul conformère? Nous notons tout de même que le complexe D1 correspond au minimum global obtenu au niveau MP2/aug-cc-pVDZ. Afin de conforter ces résultats, nous avons effectué un calcul d'optimisation des quatre complexes au niveau MP2/aug-cc-pVTZ. L'erreur de cette méthode sur les énergies de formation de complexe à liaisons hydrogènes dans l'ensemble S22 [151] est de 3,1 kJ/mol. Nous allons également mesurer les écarts énergétiques de ces quatre conformères avec l'une des méthodologies développées dans le chapitre 4.

La comparaison entre les données expérimentales et théoriques (tableau 5.9) met en évidence que les deux niveaux de calcul (MP2/aug-cc-pVDZ et MP2/aug-cc-pVTZ) reproduisent de manière satisfaisante les propriétés électriques et rotationnelles des complexes. Cependant, l'erreur moyenne en valeur absolue (MAE) au niveau MP2/aug-cc-pVDZ sur les constantes rotationnelles par rapport à l'expérience est la plus petite pour le complexe D3. On note toutefois que l'erreur sur le complexe D1 est légèrement supérieure à celle du complexe D3 (de 36 MHz). Cette inversion est corrigée lors de l'emploi d'une base plus étendue. L'erreur moyenne en valeur absolue (MAE) étant minimale pour le complexe D1, par conséquent, notre analyse énergétique qui est faite dans les tableaux 5.8 et 5.7 est validée. Nous allons décrire les caractéristiques géométriques des conformères appartenant à la première famille de conformères. Les structures des autres conformères sont reportées en annexe.

Stabilité relative kJ/mol		
MP2/aVDZ	Te	Te + ZPE
1ère catégorie		
D1	0,0	0,0
D2	2,7	2,2
D3	3,3	2,2
D4	3,9	3,0
2ème catégorie		
S1-S2	16,9	13,0
S2-S2	20,4	15,5
S1-S3	21,9	17,6
D5	22,5	19,1
S1-S4	22,9	18,7
3ème catégorie		
S1-S5	25,9	20,1
D6-hole	29,7	24,0
S2-S5	31,3	23,7
D7-hole	31,4	25,7
S3-S5	34,7	27,1
S4-S5	36,0	28,1

Tableau 5.6 – Différences énergétiques (Te) et différences énergétiques corrigées de l'énergie de point zéro (Te + ZPE)

Différence d'énergie par rapport au conformère le plus stable (Te) en kJ/mol				
	MP2	MP2	RSDH MP2	CCSD(T)-F12
	aVDZ	aVTZ	aVDZ	aVDZ
D1	0,0	0,0	0,0	0,0
D2	2,7	2,3	2,1	2,3
D3	3,3	3,2	3,0	3,1
D4	3,9	3,2	1,9	3,6

Tableau 5.7 – Ecart énergétique des complexes D2,D3 et D4 par rapport au complexe D1 à différent niveaux de théories. Les géométries utilisées pour les méthodes RSDH et CCSD(T)-F12 sont celles qui ont été utilisé suite à un calcul d'optimisation MP2/aVTZ. L'approximation 3 est utilisée pour les méthodes RSDH. L'erreur moyenne en valeur absolue de la méthode RSDH MP2 avec l'approximation 3 avec la base aVDZ est de 0,26 kcal/mol (soit environ 1,1 kJ/mol) sur les énergies de formation de complexe à liaisons hydrogènes dans l'ensemble S22.

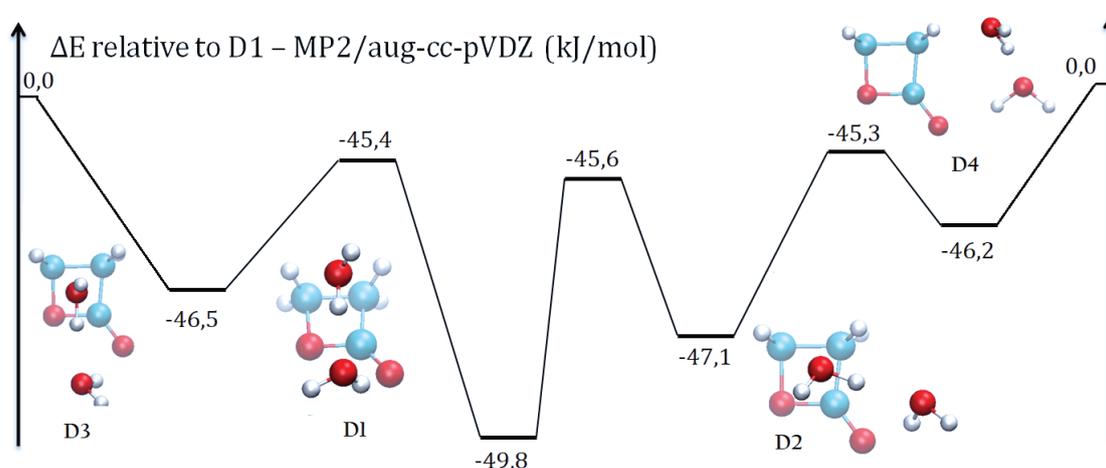


FIGURE 5.26 – Étude de la surface d'énergie potentielle associée au complexe BPL  $:(\text{H}_2\text{O})_2$ . Les différences énergétiques sont calculées par rapport à la somme des énergies électroniques des molécules formant le complexe.

MP2/aVDZ	Caractéristiques énergétiques					
	Te	Te+ZPE	Ce	Ce <sup>CP</sup>	De	De <sup>CP</sup>
D1	0,0	0,0	-52,6	-42,4	-49,8	-39,6
D2	2,7	2,2	-52,8	-42,4	-47,1	-36,8
D3	3,3	2,2	-51,1	-41,0	-46,5	-36,4
D4	3,9	3,0	-50,4	-41,5	-46,2	-37,1

Tableau 5.8 – Caractéristiques énergétiques des complexes D1 à D4 au niveau MP2/aug-cc-pVDZ en kJ/mol.

Les structures D1 à D4 sont toutes caractérisées par une liaison hydrogène entre l'hydrogène d'une molécule d'eau et l'oxygène du groupement carbonyle. Cette distance est

	Moment dipolaire (D)					Constantes rotationnelles (MHz)			
	$\mu_a$	$\mu_b$	$\mu_c$	$\mu_{total}$	MAE	A	B	C	MAE
Exp	2,16	1,54	0,33	2,67		2857	1730	1378	
MP2/aVDZ									
D1	2,40	1,56	0,48	2,91	0,65	2899	1775	1430	139
D2	0,05	0,54	1,18	1,29	5,33	3030	1676	1362	243
D3	1,19	1,56	1,54	2,49	2,38	2935	1742	1391	103
D4	0,19	1,01	0,72	1,25	4,31	3071	1348	1034	940
MP2/aVTZ									
D1	2,38	1,51	0,42	2,85	0,52	2940	1779	1426	180
D2	0,21	0,82	1,41	1,65	4,78	3054	1674	1347	284
D3	1,79	2,09	1,32	3,05	2,29	3028	1659	1314	306
D4	0,04	1,49	0,80	1,69	3,61	3122	1356	1039	978

Tableau 5.9 – Comparaison des moments dipolaires et constantes rotationnelles expérimentaux à ceux obtenus aux niveaux MP2/aug-cc-pVDZ et MP2/aug-cc-pVTZ, où MAE correspond à l’erreur moyenne en valeur absolue.

comprise entre 1,86 et 1,99 Å. Pour rappel, dans le conformère S1 du complexe 1 : 1 cette distance est de 1,98 Å. Dans le conformère S3, elle est de 2,14 Å et enfin, dans le conformère S4, elle est de 2,01 Å.

Plus précisément, dans le complexe D1, la molécule d’eau en interaction avec l’oxygène du groupement carbonyle occupe une position proche de la molécule d’eau du conformère S3 ; l’ajout d’une seconde molécule d’eau a eu pour effet de raccourcir cette distance de 0,15 Å. En ce qui concerne le complexe D3, la molécule d’eau en interaction avec l’oxygène du groupement carbonyle occupe la position S4 ; l’ajout d’une seconde molécule d’eau a eu pour effet de raccourcir cette distance de 0,04 Å.

De même, dans le complexe D4, la molécule d’eau en interaction avec l’oxygène du groupement carbonyle occupe une position proche de la molécule d’eau du conformère S1 ; l’ajout d’une seconde molécule d’eau a eu pour effet de raccourcir cette distance de 0,12 Å.

Dans le complexe D2, la molécule d’eau en interaction avec l’oxygène du groupement carbonyle occupe une position similaire au complexe mono-hydraté S1. La forte interaction avec la seconde molécule d’eau a eu pour conséquence d’élever la première molécule d’eau hors du plan de la BPL. Quant à l’angle caractéristique du dimère d’eau  $\widehat{O-H...O} = 171,3^\circ$ , nous observons une légère diminution de celui-ci dans l’ordre suivant  $169,5^\circ$  (D1),  $166,8^\circ$  (D3),  $162,8^\circ$  (D4) et  $159,2^\circ$  (D2).

La distance liaison hydrogène entre les deux molécules d’eau est comprise entre 1,86 et 1,90 Å, soit une légère diminution par rapport au dimère d’eau isolé (1,95 Å).

En somme, la présence de la seconde molécule d’eau a renforcé la liaison hydrogène

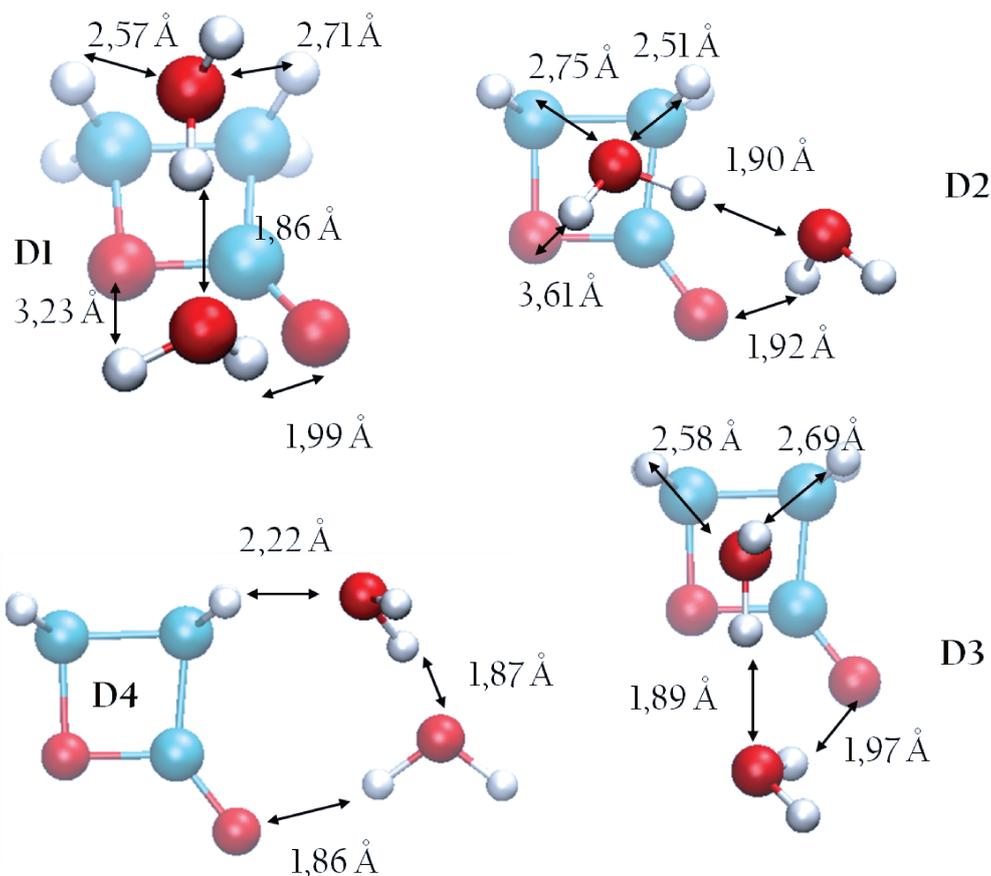


FIGURE 5.27 – Principales distances d’interaction entre les molécules formant les complexes D1 à D4.

avec l’oxygène du groupement carbonyle dans les conformères D1 à D4 et la présence de la BPL a renforcé la liaison hydrogène dans le dimère d’eau.

#### 5.4.4 Caractérisation QTAIM des conformères $BPL(H_2O)_2$ - étape 4

Le tableau 5.10 regroupe les deux informations topologiques pertinentes, la densité électronique ( $\rho(r)$ ) et son laplacien ( $\nabla^2\rho(r)$ ) aux points critiques relatifs des interaction non covalentes présent dans le conformère D1 :

$\rho(\mathbf{r})$	$\nabla^2\rho(\mathbf{r})$	$\varepsilon$	Atomes liés
Conformère D1			
0,022	0,08	0,06	H <sub>eau</sub> —O <sub>cétone</sub>
0,007	0,03	2,20	O <sub>eau</sub> —O <sub>éter</sub>
0,008	0,03	0,80	O <sub>eau</sub> —H <sub><math>\alpha</math></sub>
0,009	0,03	0,90	O <sub>eau</sub> —H <sub><math>\beta</math></sub>
0,031	0,11	0,02	H <sub>eau</sub> —O <sub>eau</sub>
Dimère d'eau isolé			
0,024	0,086	0,02	H <sub>eau</sub> —O <sub>eau</sub>
Conformère S2			
0,016	0,052	0,14	H <sub>eau</sub> —O <sub>éter</sub>
0,008	0,040	0,50	O <sub>eau</sub> —H <sub><math>\beta</math></sub>
Conformère S3			
0,016	0,050	0,03	H <sub>eau</sub> —O <sub>cétone</sub>
0,007	0,028	0,11	H <sub>eau</sub> —O <sub>éter</sub>

Tableau 5.10 – Valeurs de la densité électronique aux des points critiques et de leurs laplaciens pour le complexe D1.

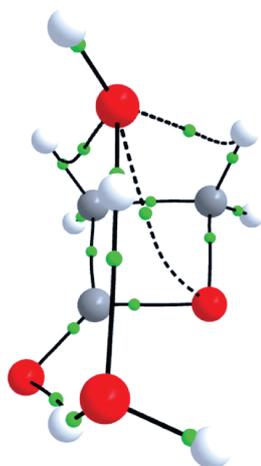


FIGURE 5.28 – Graphe moléculaire du complexe 1 : 2 observé expérimentalement.

L'analyse AIM montre que le complexe D1 est caractérisé par cinq BCPs : deux décrivant une liaison hydrogène de type C-H ...O, un décrivant une interaction de type O...O, un caractérisant la liaison hydrogène entre les deux molécules d'eau et une entre l'oxygène de la fonction carbonyle et une molécule d'eau. Le BCP(O...O) est un point particulier présentant une valeur élevée de l'ellipticité ( $\varepsilon = 2,2$ ). Cette propriété est une indication claire du caractère ambivalent d'un point critique entre O...O et O-H<sub>eau</sub> ... O<sub>BPL</sub>. La valeur de la densité électronique au BCP est un indicateur de la force de l'interaction intermoléculaire. Les valeurs présentées dans le tableau 5.10 montrent que les interactions de type liaison hydrogène sont plus fortes les autres types d'interaction que l'on qualifiera de secondaires (interactions C-H ... O et O...O).

Le complexe D1, peut être vu comme un complexe où une molécule d'eau occupe la même position que dans le complexe S3 et l'autre occupe la même position que dans le

complexe S2. Dès lors, il est intéressant de comparer les valeurs de la densité électronique aux points critiques et leurs laplaciens. Nous allons également comparer la valeur du BCP dans le dimère d'eau isolée et dans le complexe D1.

- Dans le dimère d'eau isolé, la valeur de la densité électronique au BCP et de son laplacien, caractérisant la liaison hydrogène sont respectivement de 0,024 et 0,086 u.a. alors que dans le complexe D1, ces valeurs sont respectivement de 0,030 et 0,105 u.a. indiquant que la BPL renforce légèrement la liaison hydrogène entre deux molécules d'eau.
- Dans le complexe S3, la valeur de la densité électronique au BCP et de son laplacien, caractérisant la liaison hydrogène entre l'oxygène de la BPL et l'hydrogène de la molécule d'eau sont respectivement de 0,016 et 0,051 u.a. Dans le complexe D1, ces valeurs sont respectivement de 0,022 et 0,080 u.a. soit une augmentation 0,006 u.a.
- La densité électronique au BCP entre l'oxygène de l'eau et l'hydrogène porté par le carbone  $\beta$  passe de 0,008 u.a. à 0,009 u.a. lorsque l'on passe du S2 au complexe D1, soit un changement négligeable.

Il en ressort que la présence de la BPL a pour conséquence de renforcer la liaison hydrogène présente dans le dimère d'eau. L'ajout d'une seconde molécule d'eau a également eu pour effet de renforcer les interactions présentes dans les complexes mono-hydratés. Cette tendance est connue sous le nom d'effet coopératif.

## 5.5 Etude du complexe BPL(H<sub>2</sub>O)<sub>3</sub>.

### 5.5.1 Caractérisation des régions nucléophiles et électrophiles des complexes BPL(H<sub>2</sub>O)<sub>2</sub> et du trimère d'eau - étape 1

Cette sous-section se propose d'identifier les minima et maxima du MEP sur une surface d'iso-densité de  $\rho = 0,001$  u.a. dans le but de proposer des directions privilégiées d'approches pour la formation des complexes 1 : 3. on peut considérer d'une part le trimère d'eau en interaction avec la BPL, ou d'autre part, l'un des conformères di-hydratés en interaction avec une molécule d'eau supplémentaire. La figure 5.29 indique les points remarquables du potentiel électrostatique moléculaire du complexe D1 :

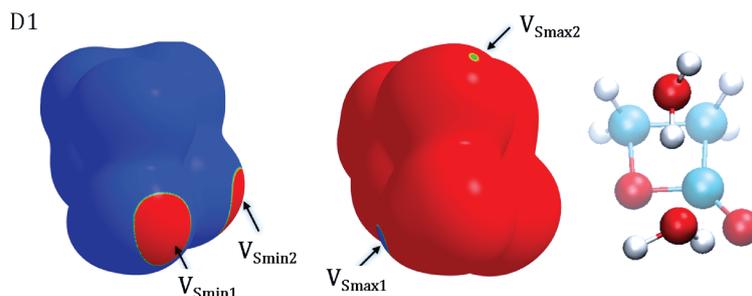


FIGURE 5.29 – Points remarquables du MEP sur le complexe D1

La première zone électrophile se situe sur l'axe de la liaison covalente O-H de la molécule d'eau (indiquée par une flèche  $V_{S_{max1}}$  sur la figure 5.29) dont la valeur est de 37,7 kcal/mol. Il s'agit de la molécule d'eau en interaction avec le groupement carbonyle. La deuxième zone électrophile ( $V_{S_{max2}} = 37,0$  kcal/mol) est localisée au niveau de l'axe de la liaison covalente O-H de l'autre molécule d'eau. Dans les deux cas, il s'agit de l'hydrogène de la molécule d'eau qui est libre : non engagé dans une liaison hydrogène. Les zones nucléophiles sont également identifiées sur la figure 5.29. La zone la plus nucléophile ( $V_{S_{min1}} = -30,4$  kcal/mol), se situe à proximité de l'oxygène de la molécule d'eau en interaction avec le groupement carbonyle alors que la deuxième zone la plus nucléophile est localisée au niveau de l'oxygène de l'autre molécule d'eau ( $V_{S_{min2}} = -22,6$  kcal/mol).

Dans le cas du complexe D2, les zones électrophiles et nucléophiles les plus importantes sont indiquées sur la figure 5.30. La zone la plus électrophile ( $V_{S_{max1}} = 43,9$  kcal/mol) est située le long de l'axe de la liaison covalente O-H de la molécule d'eau (il s'agit de la molécule d'eau en interaction avec l'oxygène du groupement carbonyle). La deuxième zone la plus électrophile ( $V_{S_{max2}} = 36,3$  kcal/mol) est à proximité de l'hydrogène libre de l'autre molécule d'eau. Nous allons décrire la localisation des zones nucléophiles sur le complexe D2. La zone indiquée par une flèche  $V_{S_{min1}}$  (-30,2 kcal/mol) est la zone la plus nucléophile. Quant à la deuxième zone la plus nucléophile, elle se situe au niveau de l'oxygène de la molécule d'eau ( $V_{S_{min2}} = -26,4$  kcal/mol).

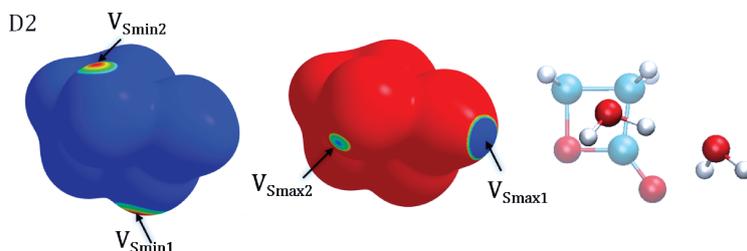


FIGURE 5.30 – Points remarquables du MEP sur le complexe D2

En ce qui concerne le complexe D3, les zones électrophiles et nucléophiles les plus importantes sont indiquées respectivement par des flèches  $V_{Smax}$  et  $V_{Smin}$  sur la figure 5.31. La zone la plus électrophile est située sur l'axe de la liaison covalente O-H de la molécule d'eau qui est en interaction avec l'oxygène du groupement carbonyle ( $V_{Smax1} = 40,8$  kcal/mol). La deuxième zone la plus électrophile est située à proximité de l'autre hydrogène libre de la seconde molécule d'eau ( $V_{Smax2} = 37,7$  kcal/mol).

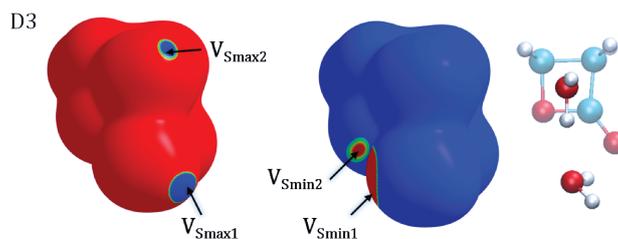


FIGURE 5.31 – Points remarquables du MEP sur le complexe D3

La région la plus nucléophile se situe à proximité de l'oxygène de la molécule d'eau qui est en interaction avec la BPL ( $V_{Smin1} = -33,8$  kcal/mol). La deuxième région la plus nucléophile est située au voisinage de l'oxygène adjacent à la fonction carbonyle ( $V_{Smin2} = -29,9$  kcal/mol).

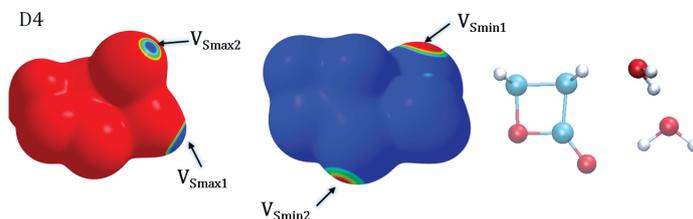


FIGURE 5.32 – Points remarquables du MEP sur le complexe D4

Les points remarquables du MEP sur la surface d'iso-densité de 0,001 u.a. du complexe D4 sont indiqués par des flèches  $V_{Smax}$  et  $V_{Smin}$  sur la figure 5.32. La zone nucléophile la plus importante est située sur l'oxygène de la molécule d'eau ( $V_{Smin1} = -31,7$  kcal/mol) la seconde zone la plus nucléophile est située au niveau de l'oxygène du groupement carbonyle ( $V_{Smin2} = -28,6$  kcal/mol). Les zones les plus électrophiles sont situées à proximité des hydrogènes libres des deux molécules d'eau ( $V_{Smax1} = 43,3$  kcal/mol et  $V_{Smax2} = 36,4$  kcal/mol).

Les points remarquables du MEP du trimère d'eau sont montrés par des flèches  $V_{Smax}$  et  $V_{Smin}$  sur la figure 5.33. Les zones nucléophiles les plus importantes sont situés sur les oxygènes ( $V_{Smin1} = -30,1$  kcal/mol et  $V_{Smin2} = -28,9$  kcal/mol). Les zones les plus électrophiles sont situées à proximité des hydrogènes libres des molécules d'eau ( $V_{Smax1} = 42,7$  kcal/mol et  $V_{Smax2} = 42,0$  kcal/mol).

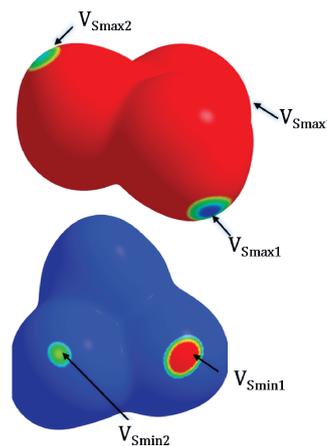


FIGURE 5.33 – Points remarquables du MEP sur le trimère d'eau

	S1	S3	S4	D1	D2	D3	D4	(H <sub>2</sub> O) <sub>2</sub>	(H <sub>2</sub> O) <sub>3</sub>	H <sub>2</sub> O	BPL
				régions nucléophiles							
$V_{Smin1}$ :	-30,2	-45,4	-46,0	-30,4	-30,2	-33,8	-31,7	-39,9	-30,1	-32,3	-34,5
$V_{Smin2}$ :	-26,7	-24,6	-24,5	-22,6	-26,4	-29,9	-28,6	-20,3	-28,9		-26,4
				régions électrophiles							
$V_{Smax1}$ :	37,7	31,5	30,0	37,7	43,9	40,8	43,3	51,3	42,7	41,9	25,9
$V_{Smax2}$ :	24,8	(30,1 - 30,9)	(27,0 - 27,3)	37,0	36,3	37,7	36,4	28,3	42,0		24,8

Tableau 5.11 – Valeurs des MEP sur une surface d'iso-densité (0,001 u.a.) en kcal/mol

Dans le tableau 5.11 nous avons rassemblé les valeurs du potentiel électrostatique sur les points remarquables identifiés dans les figures 5.29 à 5.33. Nous avons également rappelé sur ce même tableau les valeurs des minima et maxima du MEP de la BPL et de la molécule d'eau des complexes mono-hydratés et du dimère d'eau. Ce travail de recensement des points remarquables du MEP va nous permettre de proposer des directions d'approche privilégiées pour la formation des complexe tri-hydratés.

## 5.5.2 Proposition de structures initiales pour le complexe tri-hydraté - étape 2

La caractérisation des régions nucléophiles et électrophiles du soluté, de la molécule d'eau, du dimère d'eau, du trimère d'eau et des complexes mono et di-hydratés va nous permettre de proposer des structures de départ pour l'optimisation géométrique des complexes comprenant trois molécules d'eau. Le nombre de structures que l'on pourrait envisager croît de manière exponentielle. Il s'avère que si l'on tient compte des résultats des études portant sur les complexes mono et di-hydratés, le nombre de complexes que l'on pourrait envisager diminue drastiquement. Ces études sur les complexes mono et

di-hydratés ont montré que :

- tout d'abord, la rupture du réseau de liaison hydrogène entraîne une déstabilisation du complexe (de l'ordre de 10 kJ/mol sur les complexes di-hydratés).
- Ensuite, la présence d'une interaction de type liaison hydrogène entre l'hydrogène d'une molécule d'eau et l'oxygène du groupement carbonyle est présente systématiquement dans les complexes les plus stables.
- Puis, la présence d'une interaction de type  $\sigma$ -hole est systématiquement présente dans les complexes les moins stables. Il faut donc proposer des structures initiales où les molécules d'eau sont éloignées du  $\sigma$ -hole de la BPL.
- Enfin, chaque molécule présente dans le complexe est simultanément donneur de liaison hydrogène et accepteur de liaison hydrogène.

En tenant compte des règles énoncées précédemment, nous avons proposé neuf structures Guess'' (voir figure 5.34) :

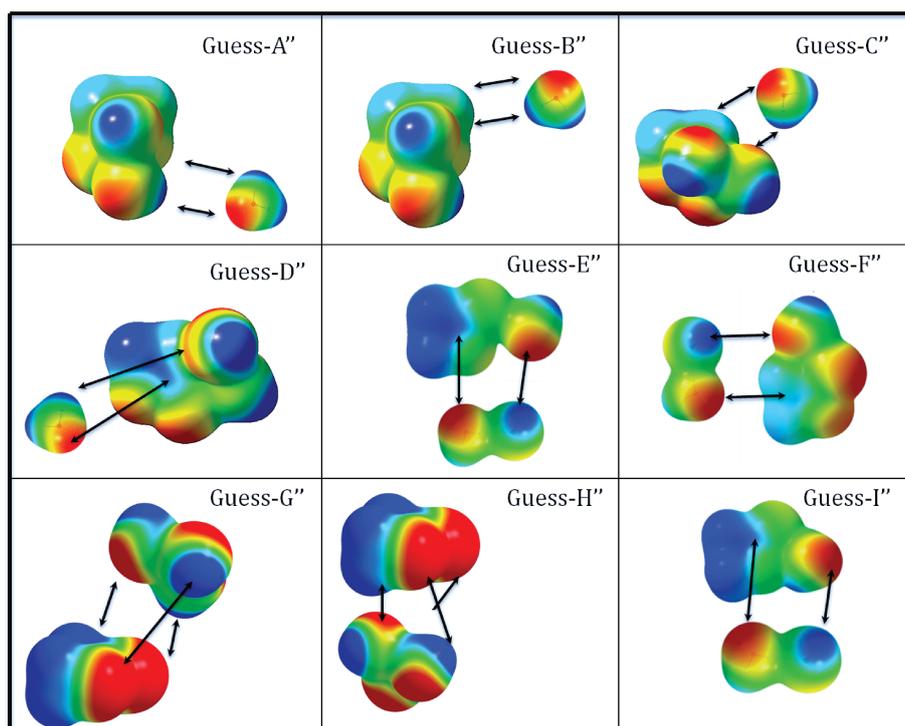


FIGURE 5.34 – Proposition de structures initiales pour l'optimisation géométrique des complexes tri-hydratés

### 5.5.3 Etude de la PES pour les complexes tri-hydratés - étape 3

La poursuite de notre étude sur les complexes tri-hydratés s'est faite en optimisant la géométrie des structures proposées à l'issue de l'étape 2 au niveau MP2/aug-cc-pVDZ (voir sous-section 5.5.2). Les calculs de fréquence ont mis en évidence sept minima sur la surface d'énergie potentielle. Encore une fois, nous avons retenu uniquement les conformères dont la différence d'énergie entre deux minima est inférieure à l'erreur de la méthode [151], estimée à 7,1 kJ/mol. Nos données théoriques mettent en évidence sept conformères quasi-isoénergétiques. Nous remarquons que le conformère T1, le conformère le plus stable au niveau MP2/aug-cc-pVDZ est le seul complexe identifié expérimentalement [34].

Pour conforter ces résultats, nous avons effectué d'une part un calcul d'énergie électronique avec les méthodes développées dans le chapitre 4 sur la géométrie obtenue au niveau MP2/aug-cc-pVDZ. Nous rappelons que les erreurs de cette méthode avec l'approximation 3 et la base aVDZ est respectivement de 1,1 kJ/mol. Le tableau 5.12 recense les différences énergétiques obtenues au niveau MP2/aug-cc-pVDZ :

	Différence d'énergie par rapport au conformère le plus stable (Te) en kJ/mol		
	MP2/aVDZ	MP2/aVDZ + ZPE	RSDH MP2/aVDZ
T1	0,0	0,0	0,0
T2	1,4	1,0	1,7
T3	1,7	1,6	1,7
T4-A	2,4	4,0	5,0
T4-B	3,0	4,5	5,5
T5	3,4	3,3	5,9
T6	5,4	4,4	6,3

Tableau 5.12 – Ecart énergétique des complexes tri-hydratés par rapport au complexe T1 à différent niveau de théorie.

Les différences énergétiques obtenues avec la méthode MP2/aug-cc-pVDZ ne nous permettent pas de conclure quant à la possibilité que les conformères T2 à T6 puissent être observés. Cependant, on peut noter que la stabilité relative des conformères tri-hydratés reste inchangée lorsque l'on a recourt à des méthodologies plus précises. Nous avons également comparé les moments dipolaires et les constantes rotationnelles obtenues expérimentalement à ceux des complexes T1 à T6 obtenus aux niveaux MP2/aug-cc-pVDZ. Le tableau 5.13 consigne les erreurs moyennes en valeur absolue par rapport aux résultats expérimentaux.

	Moment dipolaire (D)					Constantes rotationnelles (MHz)			
	$\mu_a$	$\mu_b$	$\mu_c$	$\mu_{total}$	MAE	A	B	C	MAE
Exp	2,36	0,60	0,12	2,44		1861	1166	884	
MP2/aVDZ									
T1	2,33	0,44	0,04	2,37	0,08	1922	1186	909	35
T2	1,71	0,03	1,14	2,06	0,65	1949	1166	913	39
T3	0,63	0,11	2,40	2,48	1,14	1985	1125	873	59
T4-a	4,08	1,86	2,26	5,03	1,93	1968	1210	1080	116
T4-b	4,21	2,05	1,92	5,06	1,93	1999	1206	1053	116
T5	3,33	0,67	3,11	4,61	1,55	1851	1163	1096	75
T6	1,00	4,24	0,59	4,40	1,86	1993	1113	950	84

Tableau 5.13 – Comparaison des moments dipolaires et constantes rotationnelles expérimentaux à ceux obtenus aux niveaux MP2/aug-cc-pVDZ .

Les erreurs moyennes en valeur absolue des moments dipolaires et des constantes rotationnelles répertoriées dans le tableau 5.13 concordent avec les écarts énergétiques vu précédemment (voir tableau 5.12). Ces erreurs mettent en évidence que le niveau de calcul MP2/aug-cc-pVDZ reproduit de manière satisfaisante les propriétés électriques et rotationnelles et permettent de conclure quant au minima le minima le plus stable de la surface d'énergie potentielle. Le conformère T1 corrèle parfaitement avec les données expérimentales, l'erreur moyenne en valeur absolue sur le moment dipolaire et les constantes rotationnelles sont respectivement de 0,08 Debye et 35 MHz.

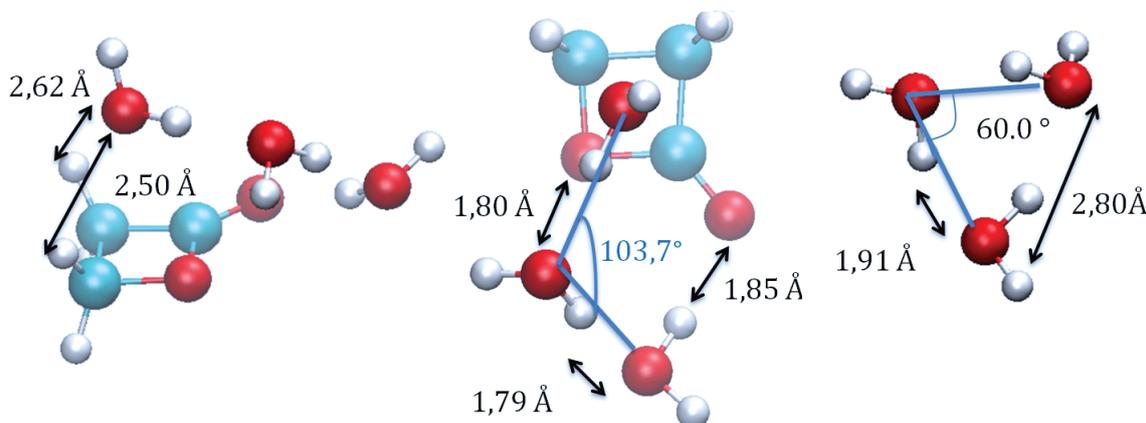


FIGURE 5.35 – Caractéristiques géométriques du complexe T1 et du trimère d'eau cyclique.

Nous allons décrire les caractéristiques géométriques de ce conformère. La figure 5.35 recense les distances de liaison hydrogène du complexe T1 et du trimère d'eau cyclique. Afin de clarifier notre discussion géométrique, nous avons labellisé chaque molécule d'eau

du complexe T1 et du trimère d'eau cyclique. Le conformère T1 est caractérisé par l'ouverture du trimère d'eau cyclique. L'angle  $\widehat{ABC}$  est de  $60.0^\circ$  dans le trimère d'eau; la présence BPL affecte considérablement cet angle dans le complexe T1, en effet cet angle est de  $103.7^\circ$  dans ce conformère.

Cet ouverture de cycle dans le complexe T1 s'accompagne d'un raccourcissement des liaisons hydrogènes dans le complexe T1. Dans le conformère T1, les distances de liaison hydrogène entre les molécules d'eau sont de  $1,80 \text{ \AA}$  alors que dans le trimère d'eau isolée, elles sont de  $1,91 \text{ \AA}$ . Ce qui correspond à une réduction de plus de  $0,1 \text{ \AA}$ .

On peut également s'intéresser aux distances de la liaison hydrogène présente à la fois dans le complexe D1 et dans le complexe T1. Cela nous permettrait d'apprécier l'impacte de l'addition d'une molécule d'eau sur les interactions présentes entre les molécules d'eau et la BPL. La distance de la liaison hydrogène entre l'oxygène du groupement carbonyle et la molécule d'eau est de  $1,99 \text{ \AA}$  alors que dans le complexe T1 elle est de  $1,85 \text{ \AA}$ . On observe également une réduction (de l'ordre de  $0,1 \text{ \AA}$ ) des distances de liaison hydrogène de type C-H...O lors de l'ajout d'une molécule d'eau supplémentaire (voir figure 5.35 et 5.27).

## 5.6 Caractérisation QTAIM des conformères BPL(H<sub>2</sub>O)<sub>3</sub> - étape 4

Nous avons regroupé dans le tableau 5.14 les deux informations pertinentes, soit la densité électronique ( $\rho(r)$ ) et son laplacien ( $\nabla^2\rho(r)$ ) aux points critiques relatifs aux interactions covalentes présent dans le complexe T1 :

Complexe T1		
$\rho(r)$	$\nabla^2\rho(r)$	Atomes impliqués dans l'interaction
0,029	0,107	H <sub>2</sub> O <sub>A</sub> —O <sub>ctone</sub>
0,008	0,035	O <sub>C</sub> —O <sub>ether</sub>
0,004	0,002	O <sub>B</sub> —O <sub>ether</sub>
0,010	0,040	O <sub>C</sub> —H(C <sub>β</sub> H)
0,009	0,004	O <sub>C</sub> —H(C <sub>α</sub> H)
0,034	0,125	H <sub>2</sub> O <sub>A</sub> —H(OH) <sub>B</sub>
0,034	0,125	H <sub>2</sub> O <sub>B</sub> —H(OH) <sub>C</sub>

Tableau 5.14 – Valeurs de la densité électronique aux des points critiques et de leurs laplaciens pour le complexe T1. "ether" réfère à l'oxygène adjacent à la fonction carbonyle.

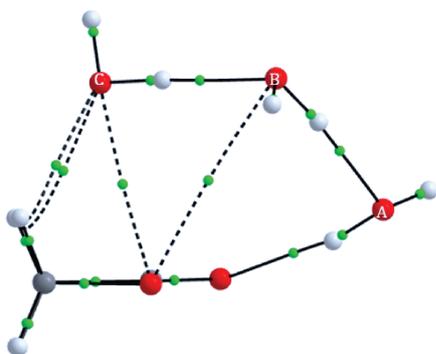


FIGURE 5.36 – Graphe moléculaire du complexe 1 : 3 observé expérimentalement.

L'analyse AIM montre que le complexe T1 comprend sept BCPs intermoléculaires : deux décrivant une liaison hydrogène de type C-H ...O, deux décrivant une interaction de type O...O, deux caractérisant les liaisons de type hydrogène entre les trois molécules d'eau et une caractérisant entre l'oxygène de la fonction carbonyle et une molécule d'eau. Le BCP O<sub>C</sub> --- O<sub>ether</sub> est un point particulier présentant une valeur élevée de l'ellipticité ( $\epsilon = 1,66$ ). Cette propriété indique clairement un caractère ambivalent d'un point critique entre O...O et O-H<sub>eau</sub> ... O<sub>BPL</sub>.

La valeur de la densité électronique au BCP indique la force de l'interaction intermoléculaire. Les valeurs présentées dans le tableau 5.14 attestent que les interactions de type liaison hydrogène sont plus forte que les autres types d'interaction que l'on qualifiera de secondaires (interactions C-H ... O et O...O). Le complexe T1 peut être vu comme un complexe où une molécule d'eau occupe la même position que dans le complexe S4 et les deux autres molécules d'eau occupent la même position que dans le complexe D1.

Dès lors, il est intéressant de comparer les valeurs de la densité électronique aux points critiques et leurs laplaciens. Nous allons également comparer les valeurs des BCP dans le trimère d'eau isolée et dans le complexe T1.

- Dans le trimère d'eau isolée, la valeur de la densité électronique des BCP et de leur laplacien, caractérisant les trois liaisons hydrogène sont respectivement de 0,030 et 0,100 u.a. alors que dans le complexe T1, ces valeurs sont respectivement de 0,034 et 0,125 u.a. indiquant que la BPL renforce légèrement les liaisons hydrogène entre les molécules d'eau. On note une absence de RCP dans le complexe T1 caractérisant une ouverture de cycle.
- Dans le complexe S4, la valeur de la densité électronique au BCP et de son laplacien, caractérisant la liaison hydrogène entre l'oxygène de la BPL et l'hydrogène de la molécule d'eau sont respectivement de 0,020 et 0,071 u.a. Dans le complexe T1, ces valeurs sont respectivement de 0,029 et 0,107 u.a.
- La densité électronique au BCP et son laplacien entre l'oxygène de l'eau et l'hydrogène porté par le carbone  $\beta$  dans le complexe D1 est respectivement de 0,009 et 0,03 u.a. Dans le complexe T1, ces valeurs sont de 0,010 et 0,040 u.a.

La présence de la BPL a eu pour conséquence de rompre une des trois liaisons hydrogène présente dans le trimère d'eau isolée d'une part et d'autre part elle a permis de renforcer les liaisons hydrogènes présente dans les complexes S4 et D1, ce qui est synonyme d'effet coopératif.

## Chapitre 6

# A theoretical investigation of watersolute interactions : from facial parallel to guesthost structures

Ce chapitre a été écrit sous la direction de M. E. Alikhani et a été publié dans *Theoretical Chemistry Accounts* [180]

## 6.1 Abstract

Encapsulation of small (bio-)organic molecules within water cages is governed by a subtle equilibrium between water-water and water-solute interactions. The competition between the formation of exohedral and endohedral complexes is investigated. The first step prior to a theoretical characterization of interactions involved in such complexes lies in the judicious choice of a level of theory. The  $\beta$ -propiolactone (BPL), a solute for which the micro-hydration was recently characterized by means of high resolution microwave spectroscopy (Angew. Chem. 2015, 127, 993), was selected for the present study, and a calibration step is carried out. It is shown that the dispersion-corrected density functional theory (DFT-D) suitably reproduce the geometric, energetic and spectroscopic features of the BPL  $:(\text{H}_2\text{O})_{1-5}$  complexes. The experimentally-deduced structures of the BPL  $:(\text{H}_2\text{O})_{4,5}$  species are fully understood in terms of the maximization of interactions between complementary sites in the MESP. DFTD calculations followed by the topological analysis within the Quantum Theory of Atoms in Molecules framework have shown that the solute could efficiently interact with  $(\text{H}_2\text{O})_{6,10}$  clusters in a similar manner that the  $(\text{H}_2\text{O})_{4,5}$  clusters do. The interaction of the solute with two larger water clusters is further investigated. The exohedral and endohedral BPL  $:(\text{H}_2\text{O})_{20}$  isomers are close in energy with each other, whereas the formation of an inclusion complex is energetically more favored than the facial interaction in the case of the BPL  $:(\text{H}_2\text{O})_{24}$  cluster. The topological analysis suggests that the substantial energetic stability is due to interactions between the solute and almost all oxygen atoms of the water cage.

## 6.2 Introduction

Due to their ubiquitous implications in almost all chemical fields, the hydration processes of organic molecules have consistently been the topic of experimental and theoretical investigations. Water-water and water-solute interactions compete and efficiently complement each other in important equilibria [17, 181–188].

For instance, hydrated species in which small water clusters are attached to a solute, were characterized for several systems [182, 189–191]. In a combined force-field and DFT approach, a clear segregation between the water molecules and the solute in hydrated complexes of 2-aminooxazole containing one to twenty water molecules has been reported by Calvo et al. [169] The authors underlined that the solute-solvent hydrogen bonding strength is weaker than the water-water ones, leading to a negligible distortion of the water cluster up to 2-aminooxazole  $:(\text{H}_2\text{O})_{15}$  cluster. For complexes containing more wa-

ter molecules, the solutesolvent interaction competes with the solvent-solvent interaction. Xantheas et al. recently reinvestigated the Benzene  $:(\text{H}_2\text{O})_8$  complex in a combined DFT and post-Hartree-Fock study [168]. They underlined that unlike to the popular B3LYP functional the dispersion-corrected functional ( $\omega$ B97-XD) is able to satisfactorily describe the Benzene  $:(\text{H}_2\text{O})_8$  complex interaction. They identified five isomers in which a cubic water octamer is attached to the aromatic ring, with similar energetic stability ( $E < 4$  kJ/mol at the MP2/aVTZ//MP2/aVDZ level of theory). They also showed that each isomer involves a cubic water octamer almost in facial parallel interaction with the aromatic ring. Furthermore, the formation of inclusion complexes in clathrate hydrates also underline the paramount importance of the complementary water-water and water-solute interaction. For instance it was shown at a molecular level that methane is efficiently encapsulated in a  $(\text{H}_2\text{O})_{20}$  cage, whereas larger  $(\text{H}_2\text{O})_{24,28}$  cages lead to more stable complexes when  $\text{SF}_6$  is the guest molecule [192]. Smaller guest species such as rare gases, diatomic or triatomic molecules may be efficiently trapped in the  $(\text{H}_2\text{O})_{20}$  cage with the concomitant formation of multiple water-solute interactions [193], characterized by bond critical points (BCPs) in the Quantum Theory of Atoms in Molecules (QTAIM) approach [4]. It was found that such small guest molecules do not lead to major distortion of the water cage. Some guest molecules however may dramatically affect water-water interactions, leading to so-called guest-induced Bjerrum defects [194].

A qualitative and quantitative understanding of interactions involved in the formation of endohedral and exohedral complexes at a molecular scale would help in understanding various processes such as hydration of small (bio-)organic molecules and clathrate formation. Endohedral complexes refer to the formation of inclusion compounds, whereas in exohedral complexes, the solute lies outside the water cluster. The present study is concerned with the mutual competition between endohedral and exohedral complexes.

A simple, powerful and quick method based on a measurable physical property of molecules and clusters was used to identify favorable interactions between two species : the molecular electrostatic potential (MESP) [182,195–202]. When investigating non-covalent interactions, studies based on the MESP analysis are mainly of two types :

- Explanatory approach A quantitative analysis of the isolated monomers MESP's may lead to the identification of electron-rich and electron-deficient regions ( $\sigma$ -holes or  $\pi$ -holes for instance). The identification of such local features provides a useful tool for interpreting various non-covalent interactions such as tetrel, chalcogen, halogen and hydrogen bondings, as demonstrated by Politzer, Murray *et al.* [196, 198].
- Predictive approach Among other approaches [1,203], we can mention the electro-

static potential for intermolecular complexes (EPIC) model of Pundlik et al. [109]. This model allows the study of weakly bond complexes between two species based on the calculation of the electrostatic interaction energy. It relies on the analysis of critical points of the electrostatic potential on isolated monomers, the MESP study of one species and the MESP-driven point charges of the other species. This model has been applied to the characterization of various complexes.

Herein another predictive approach is proposed :

- As a starting point, we investigate the possibility of a simple, qualitative analysis of the MESP for both isolated solute (S) and solvent clusters ( $W_n$ ) for proposing S : $W_s$  guess geometrical structures. Throughout the present work, the MESP analysis was used to unravel so-called electron-rich and electron-deficient regions. We should mention that such features only refer to surface properties, and not to volumes. The existence of 3D MESP maxima was indeed ruled out by Pathak et al. [106] in the 1990s. Initial geometries of complexes are built up considering the fact that electron-rich regions of the solute will be attracted toward an electron-deficient region of water clusters, and conversely. A visual inspection of the MESP of both isolated partners enables us to suggest several possible geometries which maximize the number of complementary MESP regions. Such approach has been already successfully applied to other cases [190, 204].
- In a second step, these structures are fully optimized using ab initio or DFT approaches.
- Finally, watersolute interactions are analyzed by means of the QTAIM topological tool applied on the optimized structures.

For such a study, the judicious choice of a level of theory for a consistent geometric and electronic description of complexes is an obvious and needed prerequisite. It was shown that dispersion-corrected density functionals may suitably describe such complexes provided that a calibration step is carried out, either with respect to the experimental data or to the results coming from high-level theoretical calculations [168, 193, 205]

Recently, Alonso *et al.* investigated the interaction between the  $\beta$ -propiolactone (BPL) and up to five water molecules [34]. A clear segregation between the water molecules and the solute was experimentally observed for the BPL : $(H_2O)_{4,5}$  complexes. This system was thus chosen for our study.

The paper is structured as follows : First of all, the choice of an exchange-correlation functional able to well reproduce energetic, geometric and electric properties of the water-BPL interaction in the BPL : $(H_2O)_n$  ( $n = 14$ ) cluster, is detailed. To further shed some

light on the interactions involved in the formation of facial parallel structures deduced from experimental data for the BPL  $:(\text{H}_2\text{O})_{4,5}$  clusters, a characterization based on the molecular electrostatic potential (MESP) and the Quantum Theory of Atoms in Molecules (QTAIM) topological tools is presented [4, 12, 206]. Stable structures involving the formation of similar watersolute interactions for the BPL  $:(\text{H}_2\text{O})_n$  ( $n = 6, 10$ ) clusters are then identified. In the last part, larger clusters involving 20 and 24 water molecules are studied. Two small prototypical  $5^{12}$  and  $5^{12}6^2$  water cages constituting elementary assemblies of some water clathrates were investigated. A particular attention is devoted to the transition from a facial parallel structure to the endohedral compound and to their mutual competition, until the completion of the first hydration shell around the BPL molecule.

### 6.3 Method calibration

All the geometry optimizations were carried out using the Gaussian 09 Rev D.01 software [207]. The D2 and D3 versions of the dispersion-corrected density functional theory within the framework of the BeckeJohnson damping function [17, 208] further denoted as GD2 and GD3BJ, were employed. The following functionals were selected : B3LYP-GD3BJ, [17, 48, 189] B2PLYP [59], B2PLYP-GD2, B2PLYP-GD3BJ, B97XD and LC-PBE-GD3BJ [209–211] in combination with the Pople’s 6-31++G(d,p), 6-311++G(d,p), 6-311++G(2d,2p) and the Dunning’s aug-cc-pVDZ and aug-cc-pVTZ basis sets, further referred to as Pop1, Pop2, Pop3, aVDZ and aVTZ, respectively. The CCSD(T)-F12/aVDZ method was selected as a reference method, and MP2/aVDZ as well as MP2/aVTZ calculations were also carried out. Dipole moments have been calculated with the "Density = Current" instruction for MP2, CCSD/aVDZ//CCSD(T)-F12/aVDZ (single-point with CCSD/aVDZ at the CCSD(T)-F12/aVDZ optimized geometry) and B2PLYP calculations. The molecular electrostatic potentials and the Quantum Theory of Atoms in Molecules (QTAIM) analyses were carried out using the AIMAll [212] software, for the identification of hydrogen bonding and non-covalent interactions [4].

The mono-hydrated BPL  $:(\text{H}_2\text{O})$  complex was chosen as a probe to evaluate the accuracy of the exchange-correlation functionals in describing the non-covalent interaction between water and BPL. The most relevant geometrical and energetic characteristics calculated with the 17 selected levels of theory for the BPL  $:(\text{H}_2\text{O})$  complex are gathered in the Table 6.1. The binding energy was calculated as follows :  $D_e = E_{\text{comp}} - (E_1 + E_2)$ , where  $E_{\text{comp}}$  is the total energy of the H-bonded complex and  $E_1$  and  $E_2$  are the total energies of the optimized free units (BPL molecule and water cluster).

	Distances (Å)		Energy (kJ/mol)	Dipole Moment (D) <sup>b</sup>	Rotational Constants (MHz)
	r <sub>OH...O</sub>	δr <sub>OH</sub> /δr <sub>CO</sub>	D <sub>e</sub> (D <sub>e</sub> <sup>CP</sup> )/D <sub>0</sub> (D <sub>0</sub> <sup>CP</sup> ) <sup>a</sup>	μ <sub>tot</sub> (μ <sub>a</sub> /μ <sub>b</sub> )	A/B/C
Experimental values <sup>c</sup>				2.731(0.996/2.53)	6793/2056/1614
CCSD(T)-F12/aVDZ	1.98	0.007/0.003	-29(-27)/ - ( - )		
CCSD/aVDZ <sup>d</sup>				2.69(0.97/2.51)	6772/2133/1656
MP2/aVDZ	1.98	0.008/0.007	-30(-24)/-23(-16)	2.66(0.94/2.49)	6637/2111/1636
MP2/aVTZ	1.96	0.009/0.007	-29(-26)/-21(-19)	2.65(1.05/2.44)	6774/2129/1654
B2PLYP/Pop2	1.99	0.008/0.008	-26(-23)/-20(-17)	2.65(1.03/2.45)	6877/2050/1611
B2PLYP/Pop3	1.97	0.008/0.008	-26(-23)/-19(-16)	2.73(1.20/2.45)	6870/2058/1616
B2PLYPD/Pop2	1.99	0.008/0.008	-31(-28)/-24(21)	2.65(0.80/2.53)	6729/2125/1649
B2PLYPD/Pop3	1.98	0.008/0.007	-30(-28)/-23(-20)	2.71(1.00/2.52)	6735/2127/1650
B2PLYP-GD3BJ/Pop2	1.98	0.008/0.008	-26(-23)/-19(-16)	2.65(0.95/2.48)	6824/2083/1629
B2PLYP-GD3BJ/Pop3	1.97	0.008/0.008	-25(-23)/-18(-15)	2.72(1.14/2.47)	6823/2089/1632
ωB97XD/Pop1	1.95	0.010/0.008	-31(-29)/-23(-21)	2.78(1.10/2.55)	6861/2115/1650
ωB97XD/Pop2	1.96	0.009/0.008	-30(-28)/-22(-21)	2.72(1.09/2.49)	6907/2103/1646
ωB97XD/Pop3	2.01	0.009/0.008	-29(-27)/-21(-19)	2.78(1.30/2.45)	6926/2095/1642
LC-ωPBE-GD3BJ/Pop1	1.95	0.009/0.007	-31(-29)/-24(-22)	2.76(1.08/2.54)	6905/2109/1650
LC-ωPBE-GD3BJ/Pop2	1.97	0.008/0.008	-29(-28)/-22(-21)	2.71(1.02/2.51)	6936/2075/1627
LC-ωPBE-GD3BJ/Pop3	1.96	0.009/0.007	-28(-27)/-21(-20)	2.76(1.25/2.46)	6948/2094/1643
B3LYP-GD3BJ/Pop2	1.96	0.009/0.008	-31(-29)/-23(-22)	2.71(1.10/2.49)	6871/2075/1627
B3LYP-GD3BJ/Pop3	2.01	0.009/0.008	-29(-28)/-22(-21)	2.77(1.30/2.45)	6893/2067/1623

Tableau 6.1 – Calculated geometries, energetics, dipole moments, and rotational constants associated to the experimental data of the 1 :1 complex. <sup>a</sup> D<sub>e</sub> is the binding energy, D<sub>0</sub> the binding energy corrected for the vibrational zero-point-energy. D<sub>e</sub> CP (respectively D<sub>0</sub><sup>CP</sup>) are the D<sub>e</sub> (respectively D<sub>0</sub>) corrected for BSSE using the CP method. <sup>b</sup> μ<sub>c</sub> = 0 experimentally and theoretically at all levels. <sup>c</sup> From Ref [34]. <sup>d</sup> Single point on the CCSD(T)-F12/aVDZ

The experimental rotational constants are described with a relative error (RE) less than 5% at all the chosen DFT levels. The experimental values (μ<sub>a</sub>, μ<sub>b</sub> and μ<sub>tot</sub>) of the dipole moments are accurately reproduced (RE = 5%) only at three levels of theory : B2PLYP-D/Pop3, B2PLYP-GD3BJ/Pop2 and LC-ωPBE-GD3BJ/Pop2.

Eight levels of theory accurately reproduce the binding and complexation energies calculated at the CCSD(T) level with a RE = 5% : the B2PLYP-D/Pop2, B2PLYP-D/Pop3, B97XD/Pop2, B97XD/Pop3, LC-PBE-GD3BJ/Pop2, LC-PBE-GD3BJ/Pop3, B3LYP-GD3BJ/Pop2 and B3LYP-GD3BJ/Pop3 levels. Furthermore, the MP2/aVDZ and MP2/aVTZ levels suitably describe the BPL : (H<sub>2</sub>O) interaction, the relative errors on the dipole moments, rotational constants, binding energy and complexation energy being less than 6% (Table 6.1).

It should be underlined that all the levels of theory selected for the present study satisfactorily reproduce the geometric, energetic and spectroscopic features of the 1 :1 complex. Accordingly, dispersion-corrected DFT is able to accurately describe the mono-hydrated complex. This feature is in line with other studies on hydrated species [1, 174].

The LC-PBE-GD3BJ/Pop2, that simultaneously reproduces the selected geometric, energetic and spectroscopic features with errors less than 5%, as well as the MP2/aVDZ levels, were selected for the study of larger hydrated complexes. The calculated rotational constants are in good agreement with the experimental data for the BPL  $:(\text{H}_2\text{O})_{24}$  (Table 6.2). Consequently, hereafter the LC-PBE-GD3BJ/Pop2 method will be used to study all the first-order properties of BPL  $:(\text{H}_2\text{O})_n$   $n \geq 4$ .

		Dipole Moment (D)	Rotational Constants (MHz)
		$\mu_a/\mu_b/\mu_c/\mu_{\text{tot}}$	A/B/C
1 :2	Exp	2.16/1.544/0.33/2.676	2857/1730/1378
	MP2/aVDZ	2.4/1.56/0.48/2.91	2905/1775/1430
	LC- $\omega$ PBE_D3/Pop2	2.3/1.56/0.43/2.82	2923/1792/1425
1 :3	Exp	2.357/0.6/0.12/2.435	1861/1166/884
	MP2/aVDZ	2.85/0.75/0.055/2.95	1922/1186/909
	LC- $\omega$ PBE_D3/Pop2	2.22/0.48/0.33/2.3	1927/1169/879
1 :4a	Exp	0.503/3.785/2.667/4.657	1234/932/831
	MP2/aVDZ	0.33/3.83/2.96/4.85	1260/956/846
	LC- $\omega$ PBE_D3/Pop2	0.44/4.16/2.98/5.14	1272/965/858
1 :4b	Exp	0.985/4.215/2.009/4.772	1264/934/828
	MP2/aVDZ	1.63/4.34/2.06/5.07	1321/981/840
	LC- $\omega$ PBE_D3/Pop2	0.89/4.59/2.44/5.27	1295/962/856

Tableau 6.2 – Experimental and theoretical dipole moments and rotational constants for the BPL  $:(\text{H}_2\text{O})_{24}$  clusters.

## 6.4 BPL $:(\text{H}_2\text{O})_4$ and BPL $:(\text{H}_2\text{O})_5$ complexes : A MESP picture to rationalize the structures experimentally deduced

Facial parallel structures were deduced from spectroscopic features for two BPL  $:(\text{H}_2\text{O})_4$  and one BPL  $:(\text{H}_2\text{O})_5$  isomers, denoted as 1 :4a, 1 :4b and 1 :5a, respectively. In an attempt to shed some light onto the aggregation pattern of water clusters around the solute, we carried out a MESP analysis. A MESP map of free BPL molecule revealed that both trapezoidal faces of the BPL is characterized by :

- two neighboring electron-deficient regions, at two consecutive corners, corresponding to the hydrogen atoms,
- two neighboring electron-rich regions, at two consecutive corners, corresponding to the oxygen atoms,
- an electron-deficient region in the middle of the face (Figure 6.1).

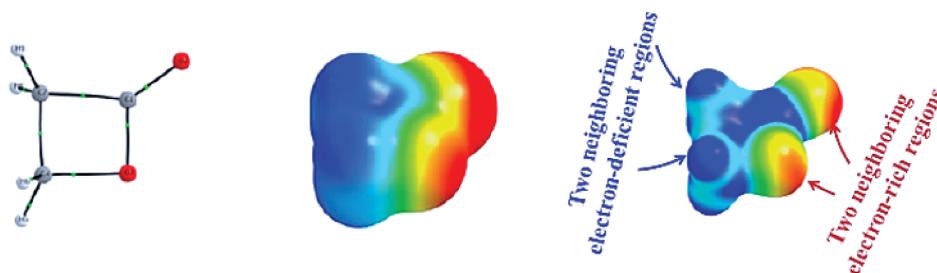


FIGURE 6.1 – Topological study of the BPL molecule : Structure of the  $\gamma$ -propiolactone (left), MESP map plotted at the 0.001 a.u. (center) and 0.04 a.u. (right) isodensity surface. Electronrich regions are depicted in red color, while electron-deficient regions are in blue color.

Electron-rich regions of the solute will be attracted toward an electron-deficient region of water clusters, and conversely. This is the starting point in the entrance channel of any non-covalent interaction. As a matter of fact, two species initially adopt the best spatial configuration to maximize the number of complementary MESP regions. The MESP analysis of water clusters (Fig. 6.2) is indeed in line with the structures earlier reported for the BPL  $:(\text{H}_2\text{O})_4$  and BPL  $:(\text{H}_2\text{O})_5$  isomers :

- in the case of the BPL  $:(\text{H}_2\text{O})_4$  complex, it was found that the water tetramer involved in the interaction with the solute is the uudd isomer, thus called with respect to the orientation of free OH groups. Both faces of the MESP of this cyclic

water tetramer presents two neighboring electron-deficient regions (corresponding to the hydrogen atoms) followed by two neighboring electron-rich regions (corresponding to the oxygen atoms). As shown in Figure 6.2, such a MESP can perfectly complement the one of the solute,

- in the case of the BPL  $:(\text{H}_2\text{O})_5$  complex, the cyclic ududu water pentamer was found to be distorted by the interaction with the solute. This distortion is in line with the MESP picture of both partners : four of the five water molecules have suitable, complementary MESP features for interaction with the solute, and the fifth one is puckered away from the solute.

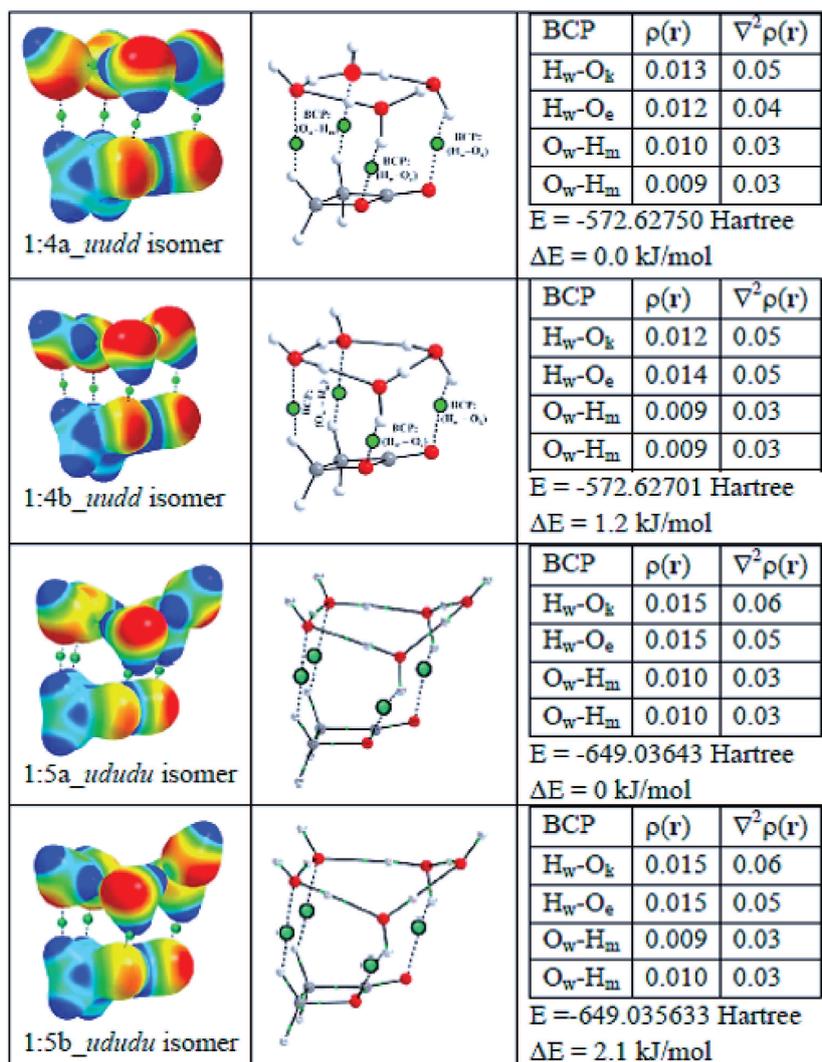


FIGURE 6.2 – MESP plot and BCPs of 1 :4<sub>uudd</sub> and 1 :4<sub>ududu</sub> isomers. See the caption of Figure 6.1 for more information.

As underlined by Alonso et al. [30], an inversion in the hydrogen bonding network in the water cluster leads to two similar isomers for the BPL  $:(\text{H}_2\text{O})_4$  complex. Similarly, our calculations also lead to the prediction of two almost isoenergetic isomers for the BPL  $:(\text{H}_2\text{O})_5$  complex, differing from one another by an inversion in the hydrogen bonding network in the water cluster.

It was found that both isomers of the BPL  $:(\text{H}_2\text{O})_4$  and of BPL  $:(\text{H}_2\text{O})_5$  complexes correspond to stable structures at the chosen level of theory. To further characterize the watersolute interaction in the equilibrium structures, the BCPs obtained from the QTAIM analysis were used to identify the interacting atoms, and to quantitatively describe the watersolute interaction by measure of the electron density  $\rho(\mathbf{r})$  and Laplacian  $\nabla^2\rho(\mathbf{r})$  at BCPs between water molecules and the solute. The interaction between water cluster and the solute for that both isomers of the BPL  $:(\text{H}_2\text{O})_4$  and of BPL  $:(\text{H}_2\text{O})_5$  complexes is characterized by four BCPs denoted as  $\text{H}_w\text{O}_k$ ,  $\text{H}_w\text{O}_e$ ,  $\text{O}_w\text{H}_m$  and  $\text{O}_w\text{H}_m$  with respect to the atoms involved in these interactions. Their energetic and topological features are reported in figure 6.2. The calculated dipole moments and rotational constants for these BPL  $:(\text{H}_2\text{O})_4$  and BPL  $:(\text{H}_2\text{O})_5$  isomers are reported in Table 6.3

		Dipole Moment (D)	Rotational Constants (MHz)
		$\mu_a/\mu_b/\mu_c/\mu_{\text{tot}}$	A/B/C
BPL $:(\text{H}_2\text{O})_4$	Experimental data	0.503/3.785/2.667/4.65	1234/932/831
		0.985/4.215/2.009/4.772	1264/934/828
	1 :4a <sub>u</sub> udd isomer	0.44/4.16/2.98/5.14	
	1 :4b <sub>u</sub> udd isomer	0.89/4.59/2.44/5.27	
BPL $:(\text{H}_2\text{O})_5$	Experimental data		946/654/642
	1 :5a <sub>u</sub> udd isomer	2.36/1.78/3.58/4.64	960/682/664
	1 :5b <sub>u</sub> udd isomer	2.92/2.25/4.44/5.78	952/698/640

Tableau 6.3 – Experimental and calculated dipole moments and rotational constants for the BPL  $:(\text{H}_2\text{O})_4$  and BPL  $:(\text{H}_2\text{O})_5$  complexes

It is remarkable that the experimentally identified BPL  $:(\text{H}_2\text{O})_4$  and BPL  $:(\text{H}_2\text{O})_5$  isomers involve similar water-solute interactions, as demonstrated by the QTAIM analysis. The above-described structures suggest that the following features of the solvent cluster may favor an efficient facial-parallel interaction with BPL : four almost coplanar water molecules with two consecutive electron-rich regions (oxygen atoms) that may interact with the electronefficient region of the BPL (hydrogen atoms of the  $-\text{CH}_2-$  groups), and two consecutive electron-deficient regions (hydrogen atoms) that may interact with the electron-rich region of the BPL (ester function). We thus selected appropriate isomers of  $(\text{H}_2\text{O})_6$  and  $(\text{H}_2\text{O})_{10}$  clusters sharing such features, and determined whether they may

attach the solute in a similar facial-parallel manner as the water tetramer and pentamer do.

## 6.5 BPL : $(\text{H}_2\text{O})_6$ complex

Different  $(\text{H}_2\text{O})_6$  isomers were identified for the water hexamer : the tridimensional prism, cage and book are almost isoenergetic, whereas the cyclic isomer is slightly higher in energy [213,214] . The MESP of each of these isomers are presented in Fig.6.3.

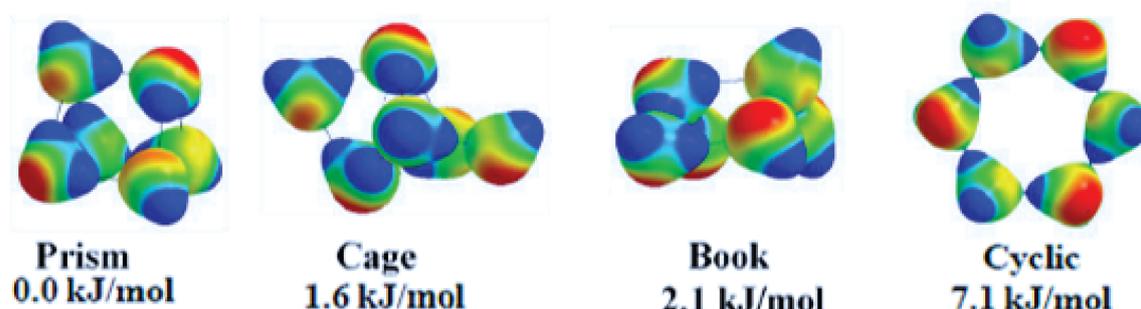


FIGURE 6.3 – MESP plot of the water hexamer isomers.

One face of the prism isomer presents a similar MESP to the one of the uudd water tetramer, with two consecutive electron-rich regions followed by two consecutive electron-deficient regions. The book isomer also contains four almost coplanar water molecules presenting a complementary MESP feature with respect to the one of the BPL (Fig.6.4). The two other isomers do not present two consecutive electron-rich regions followed by two consecutive electron-deficient regions. Thus, the possible interaction of BPL with the cage and cyclic isomers was not further investigated.

Geometry optimizations confirm that both prism and book isomers can efficiently attach the solute in a facial parallel arrangement. A QTAIM analysis carried out on relaxed structures led to the identification of four BCPs between the water cluster and the solute :  $\text{H}_w\text{-O}_k$ ,  $\text{H}_w\text{-O}_e$ ,  $\text{O}_w\text{-H}_m$ , and  $\text{O}_w\text{-H}_m$  (fig. 6.5). Thus, two isomers of the  $(\text{H}_2\text{O})_6$  cluster could attach the BPL through similar interactions to those involved in the experimentally identified BPL : $(\text{H}_2\text{O})_4$  and BPL : $(\text{H}_2\text{O})_5$  isomers.

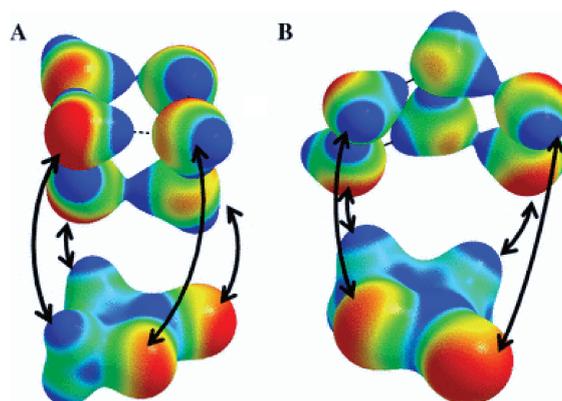


FIGURE 6.4 – Presence of two neighboring electron-deficient regions followed by two neighboring electron-rich regions on the rectangular base of the prism (A) and book (B) structures promote the electrostatic interaction with BPL in a facial parallel arrangement.

		BCP	$\rho(\mathbf{r})$	$\nabla^2\rho(\mathbf{r})$													
		H <sub>w</sub> -O <sub>k</sub>	0.019	0.08													
		H <sub>w</sub> -O <sub>e</sub>	0.014	0.05													
		O <sub>w</sub> -H <sub>m</sub>	0.013	0.04													
		O <sub>w</sub> -H <sub>m</sub>	0.010	0.03													
		$E = -725.44600$ Hartree															
		$\Delta E = 0$ kJ/mol															
		<table border="1"> <thead> <tr> <th>BCP</th> <th><math>\rho(\mathbf{r})</math></th> <th><math>\nabla^2\rho(\mathbf{r})</math></th> </tr> </thead> <tbody> <tr> <td>H<sub>w</sub>-O<sub>k</sub></td> <td>0.011</td> <td>0.04</td> </tr> <tr> <td>H<sub>w</sub>-O<sub>e</sub></td> <td>0.007</td> <td>0.03</td> </tr> <tr> <td>O<sub>w</sub>-H<sub>m</sub></td> <td>0.010</td> <td>0.04</td> </tr> <tr> <td>O<sub>w</sub>-H<sub>m</sub></td> <td>0.008</td> <td>0.03</td> </tr> </tbody> </table>			BCP	$\rho(\mathbf{r})$	$\nabla^2\rho(\mathbf{r})$	H <sub>w</sub> -O <sub>k</sub>	0.011	0.04	H <sub>w</sub> -O <sub>e</sub>	0.007	0.03	O <sub>w</sub> -H <sub>m</sub>	0.010	0.04	O <sub>w</sub> -H <sub>m</sub>
BCP	$\rho(\mathbf{r})$	$\nabla^2\rho(\mathbf{r})$															
H <sub>w</sub> -O <sub>k</sub>	0.011	0.04															
H <sub>w</sub> -O <sub>e</sub>	0.007	0.03															
O <sub>w</sub> -H <sub>m</sub>	0.010	0.04															
O <sub>w</sub> -H <sub>m</sub>	0.008	0.03															
$E = -725.44108$ Hartree																	
$\Delta E = 13.0$ kJ/mol																	

FIGURE 6.5 – MESP plot and BCPs of BPL  $:(\text{H}_2\text{O})_6$ .

## 6.6 BPL $:(\text{H}_2\text{O})_{10}$ complex

The most stable isomer for the water decamer is constituted of two stacked dudud pentameric plans [210]. The MESP of one of the rectangular face presents two consecutive electron-rich regions followed by two consecutive electron-deficient regions (Figure 6.6) that may efficiently attach the solute.

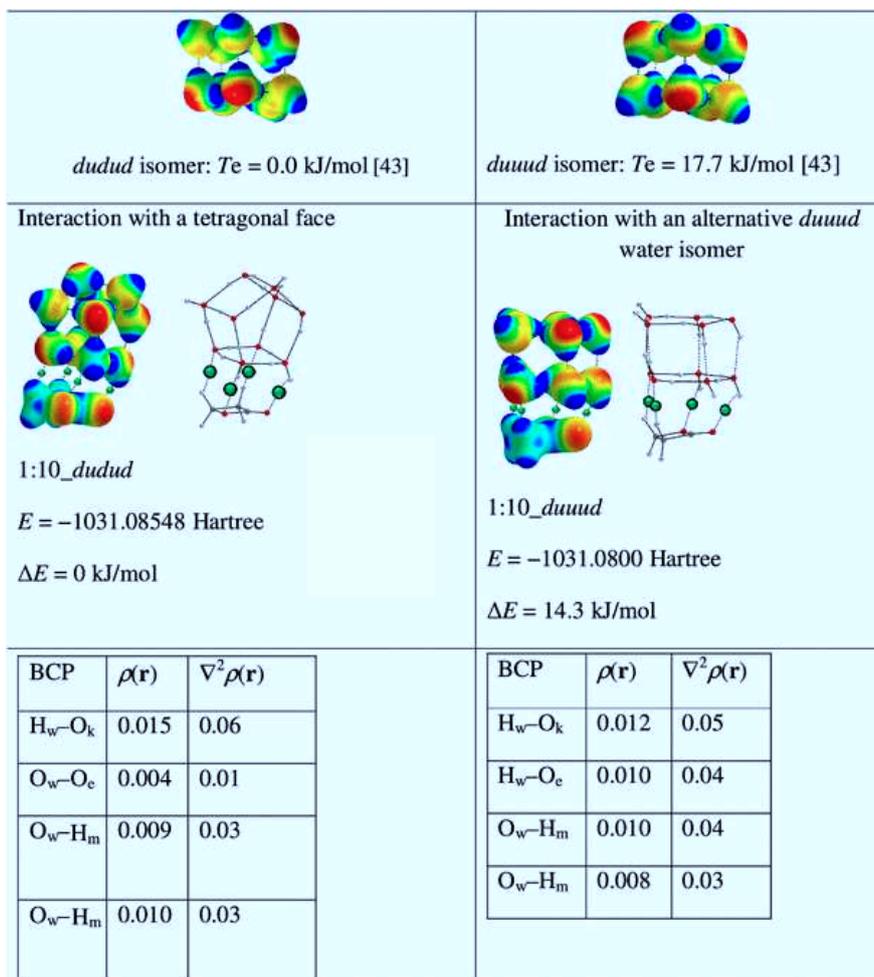


FIGURE 6.6 – Study of the BPL  $:(\text{H}_2\text{O})_{10}$  isomer.

From the above discussed MESP analysis of the water clusters that may lead to a favorable facial interaction with the solute, it can be anticipated that an alternative water decamer could also favorably attach the BPL. Indeed, the flip of some hydrogen atoms leads to a new isomer constituted of two pentagonal *duuud* faces. The MESP of the pentagonal face of such an isomer is characterized by two consecutive electron-rich regions followed by three consecutive electron-deficient regions. Thus, the pentagonal face of this isomer is suitable for interaction with the BPL (Figure 6.6). Geometry optimization indeed confirms that a water decamer cluster may attach the solute either by a rectangular face of the *dudud* isomer or by a pentagonal face of the *duuud* isomer. In both cases, a facial parallel structure characterized by four BCPs between the water cluster and the solute (H<sub>w</sub>-O<sub>k</sub>, H<sub>w</sub>-O<sub>e</sub>, O<sub>w</sub>-H<sub>m</sub>, and O<sub>w</sub>-H<sub>m</sub>) is obtained (Figure 6.6).

Thus, facial parallel structures characterized by four BCPs between the solute and a water cluster, and similar to the ones observed for the BPL  $:(\text{H}_2\text{O})_{4,5}$ , may exist for the

larger BPL  $:(\text{H}_2\text{O})_6, 10$  complexes.

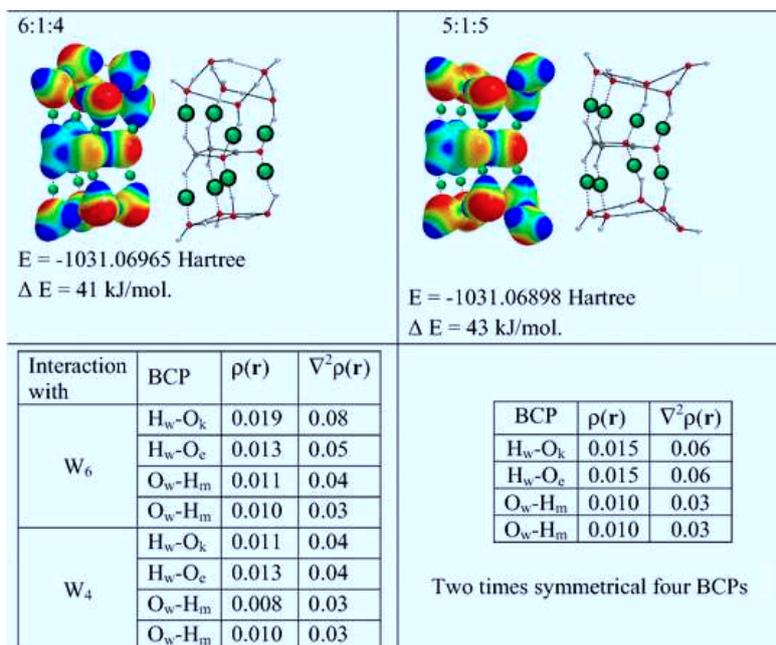


FIGURE 6.7 – Study of alternative BPL  $:(\text{H}_2\text{O})_{10}$  isomer involving two water clusters in interaction with the solute.

In the case of the BPL : $(\text{H}_2\text{O})_{6,10}$  complex, we alternatively considered two isomers in which two water clusters are simultaneously interacting with the BPL (Fig. 6.7) :

- A  $(\text{H}_2\text{O})_5$  :BPL : $(\text{H}_2\text{O})_5$  isomer in which each water pentamer is interacting with one face of the solute,
- A  $(\text{H}_2\text{O})_4$  :BPL : $(\text{H}_2\text{O})_6$  isomer in which a uudd water tetramer is attached to one face of the BPL, whereas a water hexamer in its book shape is interacting with the other face.

The QTAIM analysis on the relaxed 5 :1 :5 and 6 :1 :4 structures demonstrates that each water cluster is attached to the solute by means of four interactions :  $\text{H}_w\text{-O}_k$ ,  $\text{H}_w\text{-O}_e$ ,  $\text{O}_w\text{-H}_m$ , and  $\text{O}_w\text{-H}_m$ . However, both 5 :1 :5 and 6 :1 :4 isomers are significantly higher in energy than the 1 :10 isomer by  $\tilde{40}$  kJ/mol, thus suggesting that BPL : $(\text{H}_2\text{O})_n$  isomers in which a single water cluster is interacting with the solute would be preferentially formed rather than sandwich  $(\text{H}_2\text{O})_m$  :BPL : $(\text{H}_2\text{O})_{nm}$  arrangements.

## 6.7 BPL in or out of water cluster : BPL : $(\text{H}_2\text{O})_{20}$ and BPL : $(\text{H}_2\text{O})_{24}$ complexes

The study of larger BPL : $(\text{H}_2\text{O})_n$  clusters is particularly relevant in the context of clathrate hydrate formation. Endohedral and exohedral BPL : $(\text{H}_2\text{O})_{20,24}$  involving the  $5^{12}$  and  $5^{12}6^2$  standard water cages, respectively, were considered (Figure 6.8).

First of all, we remind the reader that the global minimum reported in the literature for both  $(\text{H}_2\text{O})_{20}$  and  $(\text{H}_2\text{O})_{24}$  cluster is cuboidal [195, 215–220] in which the insertion sites are too small to accept a guest polyatomic molecule without opening the water network. The  $(\text{H}_2\text{O})_{20}$  cuboidal structure was confirmed at the MP2 perturbative level of theory [212], while the  $(\text{H}_2\text{O})_{24}$  cuboidal structure has been identified based on molecular dynamic calculations [220]. However, it is also well known that the unit cell of a hostguest interaction in gas hydrates consists of two hollow  $5^{12}$   $(\text{H}_2\text{O})_{20}$  (dodecahedron) and six hollow  $5^{12}6^2$   $(\text{H}_2\text{O})_{24}$  (tetrakaidecahedron) cages. The simplest  $(\text{H}_2\text{O})_{20}$  dodecahedron structure used in this work corresponds to the lowest-lying isomers D-cage containing twelve pentagonal rings which is energetically less stable than the edge-sharing structure (singlet ground electronic state) by 46 kJ/mol [211]. Concerning the hollow  $(\text{H}_2\text{O})_{24}$  tetrakaidecahedron structure, Xantheas et al. identified theoretically three family isomers depending on the acceptor/donor characteristic of the bonded OH bonds and also upon the orientation of dangling OH bonds (noted as I, II and III) located within a narrow energy windows less than 2.5 kJ/mol [221–224]. The hollow  $(5^{12}6^2)$   $(\text{H}_2\text{O})_{24}$  T-cage consi-

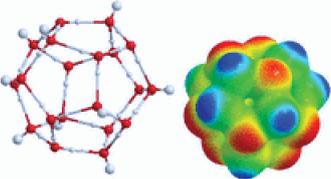
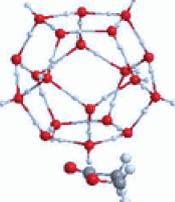
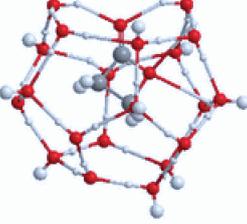
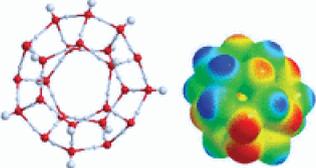
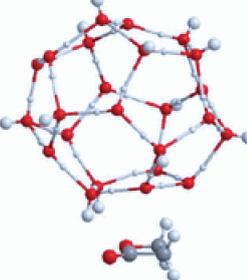
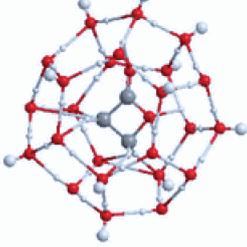
$n$	$(\text{H}_2\text{O})_n$	Exohedral 1:n	BPL@ $(\text{H}_2\text{O})_n$
20		 $E = -1795.763175$ Hartree	 $E = -1795.760004$ Hartree
24		 $E = -2101.497544$ Hartree	 $E = -2101.505807$ Hartree

FIGURE 6.8 – MESP of the  $5^{12}$  and  $5^{12}6^2$  water cages are plotted at  $\rho(\mathbf{r}) = 0.001$  a.u. Optimized geometries for both endo- and exohedral BPL  $:(\text{H}_2\text{O})_{20,24}$  complexes

dered in this work corresponds to the lowest-lying "family I isomers" of Xantheas. At the present level of theory, the  $5^{12}6^2$   $(\text{H}_2\text{O})_{24}$  T-cage was found to be 93.8 kJ/mol above the global minimum of cuboid  $(\text{H}_2\text{O})_{24}$  suggested by Takeuchi [220]. Furthermore, this T-cage structure has been used to study the stability and reactivity of  $\text{CH}_4$  encapsulated in the  $(\text{H}_2\text{O})_{24}$  cage [224]. Concerning the interaction of a BPL molecule with either  $5^{12}$   $(\text{H}_2\text{O})_{20}$  D-cage or  $5^{12}6^2$   $(\text{H}_2\text{O})_{24}$  T-cage, the exohedral structure obeys the same rules as for above-studied complexes : the  $5^{12}$  or  $(\text{H}_2\text{O})_{24}$  cage interacts with BPL by a pentagonal face offering the most favorable MESP feature, and four SW BCPs are identified.

The inclusion of BPL inside the  $5^{12}$  cage - BPL@ $(\text{H}_2\text{O})_{20}$  - generates 19 BCPs including two characteristics of the hydrogen bonding between the hydrogen of waters and the ketone oxygen of BPL, which leads to the formation of a solute-induced Bjerrum-like defect, as observed in the encapsulation of several solutes [194]. Nevertheless, the endohedral isomer is found to be slightly less stable than the exohedral form, by 8.3 kJ/mol.

Contrarily, the endohedral isomer is more stable than the exohedral one by 21.7

kJ/mol, for the BPL  $:(\text{H}_2\text{O})_{24}$  complex. A larger cavity of the  $5^{12}6^2$  structure ( $r = 4.33$  vs.  $3.95$  Å for the  $5^{12}$  dodecahedron cluster) [225] allows the BPL encapsulation without any defect on the water network. The guest molecule is found in sandwich between two hexagonal rings of the  $5^{12}6^2$  structure. According to the QTAIM analysis (Table 6.4, multiple oxygenoxygen interactions are formed between the solute and the cage, in addition to the hydrogen bonding involving hydrogen atoms of BPL.

BCP	$\rho(r)/\nabla^2\rho(r)$	Remarks
	0.005/+0.02	
	0.006/+0.02	
$\text{O}_{\text{ket}}-\text{O}_{\text{water}}$	0.008/+0.03	The ketone oxygen atom of BPL
	0.008/+0.03	
	0.008/+0.03	
	0.011/+0.05	
	0.003/+0.01	
$\text{O}_{\text{eth}}-\text{O}_{\text{water}}$	0.004/+0.02	The ether oxygen atom of BPL
	0.005/+0.02	
	0.006/+0.02	
$(\text{O}_{\text{eth}})\text{C}-\text{O}_{\text{water}}$	0.003/+0.01	The carbon atom close to $\text{O}_{\text{ether}}$
	0.003/+0.01	
$(\text{HC})\text{H}-\text{O}_{\text{water}}$	0.005/+0.02	The hydrogen atom of $\text{CH}_2$ group close to $\text{O}_{\text{ether}}$
	0.008/+0.03	
$(\text{HC})\text{H}-\text{O}_{\text{water}}$	0.005/+0.02	The hydrogen atom of $\text{CH}_2$ group close to $\text{O}_{\text{ether}}$
	0.006/+0.02	
	0.008/+0.03	
$(\text{HC})\text{H}-\text{O}_{\text{water}}$	0.005/+0.02	
	0.006/+0.02	
$(\text{HC})\text{H}-\text{O}_{\text{water}}$	0.006/+0.02	
	0.006/+0.02	
$(\text{HC})\text{H}-\text{O}_{\text{water}}$	0.003/+0.01	
	0.010/+0.03	

Tableau 6.4 – Electron density (in a.u.) and its Laplacian (in a.u.) at the BCPs between different atoms of BPL and oxygen atoms of the  $5^{12}6^2$  cage

A total of ten BCPs have been found between two oxygen atoms of BPL and those of water molecules. For these  $\text{O}_k\text{O}_w$  and  $\text{O}_e\text{O}_w$  BCPs, the electron density varies between 0.003 and 0.011 a.u. These values are very close to the ones of the oxygenoxygen interactions involved in the  $\text{CO}_2@(\text{H}_2\text{O})_{20}$  endohedral complex ( $\rho = [0.0040.008$  a.u.] [16]). We should underline that the  $\text{H}_m\text{O}_w$  interactions are far from being negligible with respect to the  $\text{O}_k\text{O}_w$  and  $\text{O}_e\text{O}_w$  ones. Consequently, both van der Waals  $\text{O}_{\text{BPL}}\text{O}_w$  interaction and hydrogen  $\text{H}_m\text{O}_w$  bonding contribute to the stability of the  $\text{BPL}@(\text{H}_2\text{O})_{24}$  endohedral complex.

## 6.8 Conclusion

In summary, the dispersion-corrected range-separated hybrid functional LC-wPBE-GD3BJ combined with a moderate size basis set (6-311++G(d,p)) was found to suitably describe the geometric, energetic and spectroscopic (dipole moments and rotational constants) features of micro-hydrated BPL  $:(\text{H}_2\text{O})_n$  complexes. Facial parallel structures previously reported for the BPL  $:(\text{H}_2\text{O})_{4,5}$  complexes were rationalized thanks to the MESP analysis of isolated partners. The QTAIM analysis further demonstrates that attachment of the water cluster on the BPL is similar in these BPL  $:(\text{H}_2\text{O})_{4,5}$  complexes. It was further shown that two isomers of water hexamer and water decamer could efficiently attach the solute by means of similar watersolute interactions characterized by four BPCs between water molecules and the solute.

Larger water clusters may also similarly interact with the solute : The attachment of the  $5^{12}$  and  $5^{12}6^2$  water D- and T-cages on the solute leads to a substantial stabilization of the system. Alternatively these water cages can also act as host for the solute. DFT-D calculations confirm that both endohedral and exohedral BPL  $:(\text{H}_2\text{O})_{20,24}$  complexes are stable with respect to the lowest-lying D- and T-cage added to free BPL molecule. Nevertheless, it is interesting to note that these endohedral compounds lie around 10 kJ/mol above the sum of free partners when we take into account the global minimum of water cluster of cuboid structure. In the BPL@ $(\text{H}_2\text{O})_{20}$  isomer, the BPL simultaneously acts as hydrogen-bond donor and hydrogen-bond acceptor to the water cluster, and the formation of solute-induced Bjerrum-like defects results in an opening of the water cage. As a result, the endohedral and exohedral BPL  $:(\text{H}_2\text{O})_{20}$  isomers are calculated to be close in energy. Such disruption of the waterwater interaction is not observed when BPL is inserted inside the larger  $5^{12}6^2$  water cage. The BPL@ $(\text{H}_2\text{O})_{24}$  isomer is much more stable than the exohedral BPL  $:(\text{H}_2\text{O})_{24}$  isomer. This can easily be explained from the QTAIM analysis that leads to the identification of 22 BCPs between oxygen atoms of water molecules and the solute in the inclusion complex, whereas only four BCPs are involved in the exohedral interaction.

## 6.9 Supporting Information

Cartesian coordinates of all the studied species are reported in the Supporting Information. This includes the Cartesian coordinates of the cuboidal (H<sub>2</sub>O)<sub>20</sub> and (H<sub>2</sub>O)<sub>24</sub> isomers. Spectroscopic features are also gathered for the BPL :(H<sub>2</sub>O)<sub>n</sub> complexes.

		Dipole Moment (D)	Rotational constants (MHz)
		$\mu_a/\mu_b/\mu_c/\mu_{tot}$	A/B/C
1 :4	1 :4-a_uudd	0.44/4.16/2.98/5.14 <b><i>0.503/3.785/2.667/4.657</i></b>	1272/965/858 <b><i>1234/932/831</i></b>
	1 :4-b_uudd	0.89/4.59/2.44/5.27 <b><i>0.985/4.215/2.009/4.772</i></b>	1295/962/856 <b><i>1264/934/828</i></b>
1 :5	1 :5-a_ududu	2.36/1.78/3.58/4.64 <b><i>946/654/642</i></b>	960/682/664 <b><i>946/654/642</i></b>
	1 :5-b_ududu	2.92/2.25/4.44/5.78	952/698/640
1 :6	1 :6-book	1.37/0.96/2.08/2.67	814/573/538
	1 :6-prism	0.81/2.24/0.1/2.38	1079/497/451
1 :10	1 :10_dudud	0.61/1.10/2.13/2.48	512/231/220
	1 :10_duuud	3.91/3.92/2.6/6.12	457/263/253
	5 :1 :5	0.00/3.64/0.73/3.71	482/218/209
1 :20	6 :1 :4	0.44/1.51/3.29/3.65	518/222/209
	1 :20 in	1.02/5.9/0.89/6.05	137/127/117
	1 :20 out	1.16/2.88/5.83/6.60	135/85/85
1 :24	1 :24 in	4.96/0.64/0.07/5.00	102/96/85
	1 :24 out	7.06/9.62/0.30/11.94	97/66/60

Tableau 6.5 – Available experimental (in bold and italic) and calculated (LC-PBE-GD3BJ/6-311++G(d,p)) rotational constants, and dipole moments of all the studied complexes.

SI-1 : Cartesian Coordinates of the isolated fragments at the LC-PBE-GD3BJ/6-311++G(d,p) level

BPL : Etot = -267.0724739 Hartree

C -0.6196280 0.0253840 0.0006140  
O -1.8027620 0.0080470 -0.0001330  
O 0.2316270 -1.0370570 -0.0003820  
C 0.4881250 1.0603350 0.0000280  
C 1.3938050 -0.1662420 0.0000410  
H 1.9899320 -0.3244060 -0.8972620  
H 1.9896400 -0.3247510 0.8974760  
H 0.5076600 1.6820210 -0.8932100  
H 0.5080340 1.6823520 0.8930200

(H2O) 2 : Etot = -152.8401088 Hartree

O 1.370072 -0.000330 -0.094734  
H 1.801758 0.768310 0.284862  
H 1.803464 -0.765780 0.289330  
H -0.556254 -0.000460 0.001657  
O -1.516018 0.000408 0.118599  
H -1.881400 -0.002692 -0.766770

(H2O) 3 : Etot = -229.2749735 Hartree

O 1.294641 0.932223 -0.085740  
O -1.461279 0.642465 -0.075881  
O 0.165275 -1.583463 0.103813  
H 1.228703 -0.037227 -0.020609  
H 1.990901 1.206656 0.512234  
H -0.602114 1.099057 -0.067346  
H -2.033349 1.109971 0.534099  
H -0.645217 -1.044793 0.069213  
H 0.071980 -2.263465 -0.565128

(H2O) 4 : Etot = -305.7124582 Hartree

H 0.229908 2.513837 0.786722

O 0.287915 1.900062 0.052919  
H 1.448091 0.581258 0.004136  
H -0.581171 1.447816 -0.004096  
O 1.900187 -0.287900 -0.052930  
H 2.514123 -0.229898 -0.786595  
H 0.581209 -1.447936 -0.004251  
O -0.287925 -1.900065 0.052931  
H -0.229897 -2.513898 0.786686  
H -1.448119 -0.581255 0.004130  
O -1.900187 0.287915 -0.052930  
H -2.514062 0.229976 -0.786653

(H2O) 5 : Etot = -382.144471 Hartree

H -0.183301 -1.923766 -0.076572  
H 2.807168 -0.088981 1.011519  
O 0.769842 -2.147138 -0.170485  
O -1.820134 -1.407032 0.042057  
H -1.919458 -0.429690 0.027878  
H -2.396239 -1.733291 0.734543  
O -1.885455 1.308573 -0.009314  
H -2.369459 1.898918 0.569340  
H -0.971680 1.667146 -0.073726  
O 0.667307 2.187180 -0.134697  
H 1.293509 1.446481 0.029457  
H 0.993250 2.645386 -0.910957  
O 2.283452 0.053551 0.221751  
H 1.778939 -0.776752 0.066346  
H 0.847179 -2.666537 -0.972317

(H2O) 6 (prism) : Etot = -458.577271 Hartree

H 0.060325 -1.031316 -1.592631  
O 0.961015 -1.330328 -1.399027  
H 1.481587 -0.520580 -1.305334  
H -0.598458 -0.670718 1.522375  
O -1.501369 -0.258622 1.426808

H -2.014146 -0.497356 2.200806  
H 0.968131 -1.586939 0.460794  
O 0.950349 -1.222997 1.367786  
H 1.536983 -0.457380 1.284109  
H -2.525906 -0.613781 -1.772492  
O -1.750259 -0.293361 -1.308794  
H -1.894376 -0.444643 -0.354456  
H -0.961976 1.602632 -0.816539  
O -0.534405 1.995851 -0.044352  
H -0.938388 1.513578 0.693174  
H 1.151507 1.536065 -0.059526  
O 2.003103 1.044229 -0.060569  
H 2.707242 1.692260 -0.115101

(H2O) 6 (book) : Etot = -458.576453 Hartree

O 2.3945390 -1.2086510 -0.4674750  
O 2.3174750 1.4615020 -0.4210000  
O -0.0245820 1.3876560 0.8609060  
O 0.1149620 -1.5436480 0.8767460  
O -2.3121640 -1.3550890 -0.4374540  
O -2.4961360 1.3892020 -0.4098070  
H -2.3313700 -1.9111770 -1.2176640  
H -1.4608700 -1.5274850 0.0093060  
H -2.5642040 0.4187580 -0.5021950  
H -3.3128300 1.6824040 -0.0031430  
H -0.8725660 1.5741510 0.4153000  
H -0.0710010 0.4500000 1.0952470  
H 0.2716650 -2.1054410 1.6382550  
H 0.9582050 -1.5151940 0.3601040  
H 1.4380950 1.5763840 0.0149500  
H 2.9250080 2.0585320 0.0178290  
H 2.5005480 -0.2274130 -0.4863480  
H 2.5665710 -1.5212960 -1.3569620

(H2O) 10 : Etot = -764.325813 Hartree

O 0.929120 2.430220 1.024884  
H 1.290363 3.203103 1.463099  
H 1.573860 1.697680 1.144758  
O -1.694223 1.429524 1.367564  
H -0.815317 1.844612 1.349542  
H -2.019601 1.465123 0.451201  
O 2.530890 0.235877 1.139691  
H 3.348238 0.123462 1.628876  
H 1.966640 -0.569489 1.323315  
O -1.636555 -1.148196 1.929049  
H -1.677962 -0.160054 1.817404  
H -2.191216 -1.378119 2.676719  
O -2.268605 1.023386 -1.355589  
H -2.195467 0.034717 -1.272432  
H -3.045296 1.210894 -1.885815  
O -1.965421 -1.555865 -0.872325  
H -1.098219 -1.911473 -1.139373  
H -1.981489 -1.614777 0.096560  
O 2.039670 -0.167762 -1.698348  
H 2.444729 0.025627 -0.840400  
H 1.457031 0.603313 -1.875146  
O 0.673179 -2.369288 -1.319554  
H 0.965604 -3.083056 -1.889327  
H 1.202763 -1.556263 -1.560117  
O 1.010143 -1.874801 1.429785  
H 0.103606 -1.684895 1.733246  
H 0.913242 -2.258034 0.540599  
O 0.357905 1.988203 -1.781274  
H -0.574880 1.715398 -1.769629  
H 0.524546 2.361837 -0.904156

(H<sub>2</sub>O)<sub>20</sub> : Cage Etot = -1528.674718 Hartree

O 1.4795980 -2.6392140 -2.1298040  
O 2.0931540 -0.3787350 -3.2834700  
O 3.3168930 1.0887390 -1.2706100

O 3.8812790 -0.4155460 0.7740900  
O 2.7293590 -2.8395660 0.4057930  
O -1.2384350 -3.1824470 -1.9269280  
O -1.8446880 -3.2456820 0.5932070  
O 0.4205190 -3.1469230 2.0768070  
O -2.3927190 -0.7869480 -2.8089120  
O -3.8064790 0.3472270 -1.0104920  
O -3.5542440 -1.0540250 1.2960580  
O -0.4089530 0.8719570 -3.6222240  
O -2.4806820 2.7214160 -0.3150320  
O -1.6168480 2.8350480 2.1591300  
O -1.9111090 0.2529430 3.1964780  
O 0.3227590 -0.9672780 3.8497730  
O 1.8614990 3.3951440 -0.5742200  
O 1.1792280 2.9996080 1.8570740  
O -0.3658460 3.2921510 -2.0981380  
O 2.3099580 0.7596450 2.8642500  
H 2.8307360 0.2863220 2.1970190  
H 1.6775580 0.1214070 3.2382280  
H 4.8159770 -0.2937550 0.9546800  
H 3.6678080 0.1669860 -0.0142340  
H 0.2944680 -1.1367370 4.7937160  
H -0.5306480 -0.5077660 3.6186620  
H 0.2231790 2.9545190 2.0318830  
H 1.5893360 2.2043510 2.2697060  
H 3.3719930 -3.5451310 0.5063030  
H 3.2081660 -1.9930550 0.5589400  
H 1.2184550 -3.0354290 1.5375740  
H 0.4329730 -2.4333860 2.7380810  
H 2.7919870 1.8871770 -1.0820470  
H 2.8982380 0.6282250 -2.0227830  
H -2.4878070 -0.1961570 2.5555640  
H -1.8181170 1.1716560 2.8850460  
H 2.4310730 4.1678570 -0.5675550  
H 1.5858670 3.2393750 0.3908930  
H -1.9663390 2.8068610 1.2256670

H -2.0667570 3.5581950 2.6009880  
H 1.8690630 -2.7672740 -1.2486590  
H 0.5304190 -2.8594190 -2.0624020  
H -2.4127130 -2.5173010 0.8887500  
H -1.0473750 -3.2415030 1.1704260  
H 1.8580260 -1.2531170 -2.8588590  
H 2.5894530 -0.5854570 -4.0782630  
H -4.4179170 -1.3316420 1.6090170  
H -3.7035060 -0.5403510 0.4683130  
H 0.4620060 3.3867230 -1.5704390  
H -0.4578050 4.0934500 -2.6180850  
H -2.9415860 1.9059420 -0.5842230  
H -1.7925890 2.8986480 -0.9791380  
H -1.4851240 -3.2268960 -0.9540190  
H -1.5426510 -4.0012020 -2.3242750  
H -0.3737350 1.7039220 -3.1253630  
H 0.4786860 0.4792430 -3.5608310  
H -4.6655780 0.5074630 -1.4076640  
H -3.2486970 -0.1171940 -1.7218400  
H -1.6759870 -0.2074340 -3.1536470  
H -1.9844800 -1.6382330 -2.5677640

(H2O)<sub>20</sub> : Cuboid Etot = -1528.691657 Hartree

O 1.271100 4.013562 1.519482  
H 0.283992 3.913375 1.569334  
H 1.504086 4.759769 2.075388  
O -1.452580 3.980715 -1.468644  
H -0.461601 3.934811 -1.520412  
H -1.723409 4.735568 -1.995005  
O -4.050469 -0.880280 -1.353623  
H -3.584869 -0.048775 -1.550192  
H -4.306269 -0.825929 -0.418534  
O 2.491457 1.370809 -1.365319  
H 3.171789 0.726167 -1.628423  
H 2.525346 1.397546 -0.384516

O 1.197731 3.887304 -1.335962  
H 1.714097 3.107164 -1.596406  
H 1.401360 4.035006 -0.397448  
O 2.857933 -3.070535 1.573760  
H 3.365429 -3.648928 2.146487  
H 3.310898 -2.185531 1.580493  
O 4.076496 -0.728511 1.322287  
H 4.306401 -0.669690 0.380421  
H 3.583705 0.085866 1.523107  
O -2.669848 -3.122068 1.298632  
H -2.678714 -3.344461 0.352758  
H -1.731006 -3.084656 1.547381  
O 0.073606 0.116782 -1.313938  
H 0.068770 0.113041 -0.329775  
H 0.912375 0.570818 -1.543766  
O -4.228200 -0.988782 1.503624  
H -4.994171 -1.175552 2.049846  
H -3.646683 -1.794385 1.529220  
O 0.016797 -2.595201 -1.411196  
H 0.020465 -1.623774 -1.551425  
H -0.878945 -2.901479 -1.642149  
O 4.206644 -0.830023 -1.534634  
H 3.663724 -1.662337 -1.543853  
H 4.979064 -0.995036 -2.078942  
O -0.036299 0.032857 1.408400  
H -0.933826 0.398654 1.562620  
H -0.049622 -0.936835 1.560022  
O -2.342896 1.337096 -1.413790  
H -2.213922 2.283427 -1.603807  
H -1.461664 0.925651 -1.549842  
O 2.778830 -3.055340 -1.287678  
H 1.832432 -3.046473 -1.511796  
H 2.821130 -3.269777 -0.341400  
O 0.108398 -2.674986 1.330538  
H 1.003638 -2.934062 1.614441  
H 0.122038 -2.727851 0.349639

O -1.367499 3.768562 1.380037  
H -1.828012 2.944454 1.608881  
H -1.578578 3.935336 0.446361  
O 2.285299 1.422093 1.373192  
H 2.100358 2.350665 1.601114  
H 1.442457 0.942893 1.523539  
O -2.713260 -3.159180 -1.560187  
H -3.214655 -2.300579 -1.579690  
H -3.177496 -3.766162 -2.140137  
O -2.495415 1.175465 1.329118  
H -3.196211 0.548857 1.583958  
H -2.536507 1.230482 0.349705v

(H2O) 24 : Cage Etot = -1834.4027769 Hartree

O 0.8500900 2.7064170 2.7605810  
O 2.3566260 -0.8282340 -2.9220160  
O 4.3888990 -1.2130790 -0.8818450  
O 4.3808940 1.0262810 0.6613160  
O 3.0278060 2.9804070 -0.9821510  
O 1.3942960 4.3644950 0.5802290  
O -1.1138910 4.2025300 -0.8278440  
O -0.9174940 2.4857490 -2.8683270  
O 1.8254040 1.7432260 -3.1623420  
O -1.7047410 2.1062670 3.1572970  
O -3.0085700 3.1586930 0.8609270  
O -2.3857760 -0.6198880 3.0674990  
O -2.8445370 0.7946700 -2.9276870  
O -4.3964750 1.2230600 -0.7306720  
O -4.1948700 -0.9598880 1.0764200  
O -3.3019650 -3.1340770 -0.4853770  
O -1.3346650 -4.1359890 0.8916880  
O 0.8902050 -4.2819780 -0.7021960  
O 0.3528870 -2.7059370 -2.9857250  
O -2.0934770 -1.8970090 -2.6546580  
O 2.8192460 0.6126680 2.7802740  
O 1.9521360 -1.8075250 2.6604290

O 3.2565860 -3.2481690 0.5815610  
O -0.6402060 -2.5073080 3.0990420  
H 1.9898160 0.7657080 -3.0929470  
H 2.1221730 2.0072700 -4.0360710  
H -2.3444320 -0.9568930 -2.6998100  
H -2.5713400 -2.3028900 -1.9057120  
H 2.6118400 2.5789060 -1.7694720  
H 3.4634170 2.2638280 -0.4982580  
H -2.5555020 -3.5355790 0.0514690  
H -3.9303430 -3.8410750 -0.6459120  
H 1.6334940 -1.4881520 -2.9379770  
H 2.9433210 -1.0499630 -2.1858840  
H -0.5767430 -2.3562570 -2.8463780  
H 0.3196500 -3.2193920 -3.7961170  
H 4.4419340 -0.3947530 -0.3379370  
H 5.1963270 -1.2587900 -1.3992060  
H 0.7356700 -3.7571390 -1.5065220  
H 1.7077230 -3.9461760 -0.3087560  
H 5.1918880 1.4508330 0.9510580  
H 3.8247950 0.8929350 1.4732940  
H -1.0856040 -3.6003360 1.6648610  
H -0.5233140 -4.2448520 0.3526450  
H 0.0089910 2.2061040 -2.9335510  
H -0.9729870 3.1511210 -2.1372460  
H -2.0628710 1.4365170 -2.8887800  
H -3.2350410 0.9159360 -3.7962550  
H -0.3050880 4.3117530 -0.3064950  
H -1.8032740 3.8932070 -0.2028550  
H -3.8732530 1.0843940 -1.5523680  
H -5.2577730 1.5484050 -1.0002020  
H 1.7327870 5.2560280 0.6838180  
H 2.0149010 3.8888750 -0.0297230  
H -3.9491010 -1.7231380 0.5261050  
H -4.3134980 -0.2165010 0.4644110  
H 1.0116610 -2.0517400 2.8041320  
H 2.3126110 -2.3764400 1.9624600

H -0.7474820 2.3188470 3.0100410  
H -1.9574380 2.5500530 3.9696640  
H 3.7444800 -2.5897970 0.0477140  
H 3.8867030 -3.9228580 0.8438280  
H -3.4982040 2.4640620 0.3945950  
H -2.6305170 2.7519370 1.6575590  
H 2.4351520 -0.3177350 2.7235440  
H 3.1789980 0.6915130 3.6673120  
H 1.4848540 1.9769870 2.6898530  
H 1.0368600 3.3185100 2.0248180  
H -1.2992150 -1.7426400 3.0754250  
H -0.7940540 -2.9587170 3.9321100  
H -2.1343450 0.3190890 3.1000960  
H -3.0758680 -0.7140600 2.3722410

(H2O) 24 : Cuboid Etot = - 1834.4385108 Hartree

O 1.085673 -0.142294 0.090043  
H 0.372031 0.214320 0.652058  
H 1.478355 -0.870954 0.624284  
O -1.873582 2.294355 -0.007079  
H -1.412299 1.931549 0.783287  
H -1.378803 3.100822 -0.256375  
O -0.570354 1.055028 1.991271  
H -1.130191 0.496910 2.567565  
H 0.178680 1.392944 2.522294  
O 1.954409 -2.142749 1.712334  
H 2.279216 -2.902087 1.195153  
H 1.031709 -2.367084 1.984924  
O -1.854271 0.207741 -1.848008  
H -1.270866 0.394385 -2.607840  
H -1.775550 0.976140 -1.247919  
O -1.748531 -2.259935 -0.474888  
H -1.672422 -1.400169 -0.932711  
H -2.702230 -2.453293 -0.458782  
O 2.830096 1.822588 -0.641289

H 2.214844 1.110171 -0.356235  
H 3.713834 1.428838 -0.535128  
O 1.351262 -1.161436 -2.431550  
H 1.104319 -0.743969 -1.581030  
H 0.994655 -0.571954 -3.116569  
O 3.476095 0.023091 2.514323  
H 4.052647 0.113718 1.736557  
H 3.006809 -0.822710 2.382154  
O 1.771560 1.920448 3.139065  
H 2.446325 1.224012 2.890322  
H 1.942420 2.144500 4.056159  
O -2.294306 -0.608636 3.336280  
H -2.559324 -0.590687 4.257735  
H -3.129209 -0.590565 2.792279  
O 0.012254 0.874448 -3.811757  
H -0.148325 1.038908 -4.742730  
H 0.467369 1.679428 -3.438366  
O 4.889779 0.090241 0.024259  
H 5.844461 0.105909 -0.065620  
H 4.562077 -0.679506 -0.516219  
O 3.888863 -1.856277 -1.448737  
H 3.600802 -2.683884 -1.027069  
H 3.126842 -1.567888 -1.982761  
O 1.238382 2.882093 -2.651755  
H 1.953874 2.537331 -2.084291  
H 0.714846 3.461235 -2.074434  
O 1.875993 3.787996 1.002200  
H 1.791881 3.309785 1.843663  
H 2.384947 3.175235 0.433043  
O -0.128890 4.342739 -0.580744  
H 0.594600 4.203673 0.095345  
H -0.282920 5.287861 -0.637093  
O -4.437530 -0.511868 1.820133  
H -4.537437 0.317194 1.322177  
H -4.544906 -1.228632 1.170855  
O 0.419589 -3.655049 -1.645035

H -0.405828 -3.377126 -1.211383  
H 0.698507 -2.863964 -2.141941  
O -0.654272 -2.575651 2.078041  
H -1.191144 -2.007207 2.650534  
H -1.039563 -2.474797 1.186903  
O 2.585763 -4.046758 -0.204909  
H 1.730073 -4.009008 -0.716136  
H 2.849978 -4.966994 -0.153945  
O -4.512256 1.841802 0.221223  
H -3.574813 2.140984 0.150305  
H -5.040928 2.611423 0.440297  
O -4.568709 -2.261789 -0.393973  
H -5.199311 -2.952149 -0.606813  
H -4.699854 -1.517051 -1.040186  
O -4.608847 -0.096496 -1.891864  
H -4.808255 0.645946 -1.301207  
H -3.657293 0.009418 -2.085806

SI-2 : Cartesian Coordinates of the BPL : (H<sub>2</sub>O)<sub>n</sub> complexes at the LC-PBE-GD3BJ/6-311++G(d,p) level

Complex 1 :1

Etot = -343.498826 Hartree  
C 0.2741890 0.5172250 -0.0001610  
O -0.5613840 1.3656200 -0.0002370  
O 1.6151190 0.6861120 0.0003170  
C 1.8377260 -0.7543020 0.0001920  
H 2.3754340 -1.0531310 -0.8977130  
H 2.3746820 -1.0533900 0.8984610  
C 0.3315840 -0.9920590 -0.0004490  
H -0.1024540 -1.4453440 -0.8890970  
H -0.1032140 -1.4457420 0.8876220  
O -2.7175620 -0.5102900 0.0002270  
H -2.2444840 0.3335970 0.0002490  
H -3.6503450 -0.2927080 0.0005280

Complex 1 :2

Etot = -419.931076 Hartree

C 0.5963160 -0.7911980 -0.2343810  
O -0.1522960 -1.6085970 -0.6758550  
O 0.6498090 -0.3490490 1.0391570  
C 1.6853130 0.6227110 0.7004490  
H 1.2938510 1.6291010 0.8235780  
H 2.5790450 0.4437660 1.2954200  
C 1.6970420 0.1081360 -0.7352900  
H 2.5913610 -0.4229630 -1.0565090  
H 1.3531230 0.8083930 -1.4919810  
O -2.5633110 -0.3673370 0.3031060  
H -2.6375550 -0.5262720 1.2464260  
H -1.9648880 -1.0537390 -0.0265840  
O -0.9432440 1.8160550 -0.4207390  
H -1.4205830 2.4055920 -1.0058360  
H -1.5940480 1.1496560 -0.1345240

Complex 1 :3

Etot = -496.36421222 Hartree

C -0.8963300 -0.9012380 -0.1593890  
O -0.1021720 -1.6515340 -0.6441500  
O -0.8314350 -0.3249880 1.0539580  
C -2.0865740 0.3885080 0.8256250  
H -1.9011310 1.4587680 0.8055830  
H -2.8216200 0.0993370 1.5746280  
C -2.2070990 -0.2690570 -0.5452200  
H -3.0203730 -0.9794220 -0.6826090  
H -2.1090260 0.4034600 -1.3932410  
O 2.5841950 -1.3285190 0.0489040  
H 1.6719660 -1.6026770 -0.1527980  
H 3.1431020 -1.7400520 -0.6122270  
O 2.2860430 1.3793780 0.3732970

H 2.3319810 1.5306850 1.3189800  
H 2.4856990 0.4290250 0.2491640  
O -0.1199550 2.0075780 -0.8224740  
H 0.0909590 2.6301900 -1.5191540  
H 0.7350550 1.8060840 -0.3907060

Complex 1 :4-a\_uudd

Etot = -572.800096 Hartree  
C -1.5035780 -0.6505530 -0.1926900  
O -1.3361960 -1.8255460 -0.1284030  
O -1.7483980 0.1972330 0.8399590  
C -1.7824240 1.3495220 -0.0614370  
H -2.7631100 1.8201810 -0.0261420  
H -0.9820230 2.0425790 0.1875830  
C -1.5118050 0.4290070 -1.2470200  
H -2.3097800 0.3235640 -1.9797520  
H -0.5460190 0.5479940 -1.7374050  
O 1.6494450 -1.8898110 -0.0031570  
H 1.5149150 -1.4126430 0.8417740  
H 0.8277050 -2.3723070 -0.1425480  
O 1.0381780 -0.1530660 2.0627240  
H 1.2020900 0.6901870 1.6036540  
H 0.0898650 -0.1677780 2.2239440  
O 1.4480720 2.0051300 0.3152360  
H 2.1203370 2.6765790 0.4422570  
H 1.7155370 1.4762470 -0.4668080  
O 1.8843930 0.1877310 -1.6505090  
H 2.5783770 0.0854510 -2.3033800  
H 1.8550090 -0.6512730 -1.1231020

Complex 1 :4-b\_uudd

Etot = -572.799661 Hartree  
C -1.6565850 0.4165940 -0.1007400

O -1.7457610 1.6005210 -0.0641580  
O -1.1381430 -0.3351090 -1.1087270  
C -1.3581600 -1.5699730 -0.3562210  
H -2.0350440 -2.2239240 -0.9028470  
H -0.4047380 -2.0519850 -0.1484950  
C -1.9650780 -0.7635410 0.7872440  
H -3.0263980 -0.9075220 0.9806490  
H -1.3899490 -0.7456870 1.7107220  
O 1.2036170 2.1574470 0.2474960  
H 1.1330490 1.5197920 0.9794330  
H 0.3106730 2.4959180 0.1242380  
O 1.7238090 0.3098830 -1.7286940  
H 1.6564560 1.0851620 -1.1317910  
H 0.8213170 0.1809470 -2.0398140  
O 2.0451920 -1.6569080 0.0349240  
H 2.8676300 -2.1413820 -0.0500190  
H 2.0262130 -0.9832790 -0.6922050  
O 1.0311270 -0.0662670 1.9859850  
H 1.4716380 -0.0823170 2.8373710  
H 1.4893520 -0.7207300 1.4164580

Complex 1 :5-a\_ududu

Etot = -649.2340543 Hartree

C -2.223518 0.736235 0.472304  
C -2.521315 -0.637472 -0.119095  
O -1.635452 -0.342827 -1.246403  
C -1.347065 0.888260 -0.746175  
O -0.615055 1.697034 -1.216766  
H -3.048799 1.444742 0.512412  
H -1.653464 0.746361 1.400700  
H -3.535412 -0.803751 -0.477374  
H -2.136638 -1.483986 0.444973  
O 1.890820 1.787658 0.360598  
H 2.429105 1.144777 -0.145299

H 1.238190 2.117200 -0.267935  
O 0.605307 0.598484 2.335532  
H 0.932688 0.922814 3.175663  
H 1.123852 1.066629 1.626042  
O 0.028417 -1.909695 1.528529  
H 0.288340 -1.047373 1.923965  
H 0.318458 -2.592225 2.135837  
O 3.010934 -0.204157 -1.107541  
H 2.314829 -0.893679 -1.179533  
H 3.849769 -0.603647 -1.336441  
O 0.919723 -1.925056 -1.115665  
H 0.162542 -1.545293 -1.574176  
H 0.630378 -2.006229 -0.187318

Complex 1 :5-b\_ududu

Etot = -649.2333373 Hartree

C 2.417335 -0.453893 0.494812  
C 2.376767 0.767317 -0.417890  
O 1.439700 0.073152 -1.302332  
C 1.444383 -1.027585 -0.504889  
O 0.820433 -2.026685 -0.660019  
H 3.360722 -0.993573 0.552544  
H 1.972145 -0.328583 1.480586  
H 3.292133 1.002836 -0.957337  
H 1.905170 1.654259 -0.000011  
O -1.723738 -1.787937 0.859247  
O -0.291461 0.097148 2.367367  
O -0.294951 2.359491 0.885668  
O -3.069624 -0.342864 -0.973876  
O -1.179775 1.531392 -1.458388  
H -2.678209 -0.922276 -0.280867  
H -1.027210 -2.250566 0.379816  
H -1.253968 -1.169964 1.449373  
H -0.525171 0.226664 3.287959

H -0.344251 0.977425 1.933027  
H -0.628458 3.234242 1.090170  
H -0.675311 2.105647 0.002248  
H -1.944450 0.923521 -1.368935  
H -4.022398 -0.406863 -0.910307  
H -0.466308 1.002628 -1.831804

Complex 1 :6-book

Etot = -725.671132 Hartree  
C -2.4356450 0.3504600 -0.9154130  
O -2.3237890 0.2149660 0.5407910  
C -1.6796250 -0.9709300 0.4169950  
O -1.2536050 -1.6424980 1.3032590  
C -1.7340560 -0.9930500 -1.0892670  
H -3.4844900 0.3912080 -1.2024870  
H -1.8807440 1.2251700 -1.2486100  
H -2.3382430 -1.8070160 -1.4866560  
H -0.7534330 -0.9869980 -1.5649160  
O 2.1381920 0.5088330 1.5826910  
O 2.1032100 0.9928030 -1.2440180  
O 1.5616640 -2.1407510 0.8326080  
O 1.5280110 -1.6663540 -1.8048420  
O -0.1115830 1.9481280 1.6081590  
O 0.0050780 2.6087980 -1.0606740  
H 1.3023450 1.0262860 1.7019620  
H 2.7399560 0.7826300 2.2775540  
H 2.3590130 0.8518110 -0.3188370  
H 1.9599710 0.0984430 -1.6060530  
H 1.8270270 -1.2865170 1.2178960  
H 0.6621080 -2.2889270 1.1513960  
H 1.5789450 -1.9674780 -0.8658720  
H 2.0545400 -2.2758260 -2.3243460  
H -0.1418820 2.3351660 0.7099640  
H -0.9490900 1.4837140 1.7092940

H 0.7956230 2.0397860 -1.2436710  
H 0.1868780 3.4682710 -1.4442940

Complex 1 :6-prism

Etot = -725.666098 Hartree

O 2.486212 0.050686 -1.010564  
C 2.321779 0.528986 0.247541  
C 2.635701 -1.317914 -0.515724  
O 2.127662 1.659088 0.563725  
C 2.454189 -0.819026 0.914771  
H 1.837235 -1.944593 -0.906249  
H 3.621844 -1.700709 -0.771806  
H 3.316271 -0.899564 1.574426  
H 1.547942 -1.137225 1.429605  
O -0.574824 -1.674851 -0.834572  
H -1.479669 -1.982325 -0.969079  
H -0.567607 -1.473822 0.127198  
O -0.924248 -0.866599 1.749809  
H -1.853417 -1.117035 1.640962  
H -0.912081 0.104714 1.660753  
O -3.297253 -1.370562 0.215585  
H -4.193551 -1.703806 0.282872  
H -3.362548 -0.443279 -0.104753  
O -3.129964 1.223811 -0.598322  
H -2.751395 1.723651 0.136504  
H -2.437314 1.257096 -1.274091  
O -0.470262 0.900194 -1.620145  
H -0.442666 -0.073963 -1.467130  
H 0.324025 1.121067 -2.113405  
O -0.787402 1.827294 0.969085  
H -0.041962 2.305231 1.344919  
H -0.514490 1.619794 0.052946

Complex 0,041666667 10\_dudud

Etot = -1031.4161351 Hartree

C 3.888164 0.945455 0.956628

O 3.720451 0.835698 -0.492138

C 3.475257 -0.490241 -0.401963

O 3.226939 -1.235958 -1.297775

C 3.623467 -0.549987 1.099082

H 4.895238 1.285256 1.191171

H 3.123273 1.598712 1.369918

H 4.458585 -1.164764 1.430422

H 2.707401 -0.833637 1.616840

O 0.400537 -1.756061 -1.161028

O -2.325834 -2.308951 -0.844556

O 0.553604 0.926811 -1.695246

O -3.688908 -0.134221 -1.466481

O -2.391414 -1.766526 1.949681

O -3.572958 0.506071 1.294782

O 0.625041 1.453192 1.060014

O -1.748601 2.623799 1.017209

O -1.859377 1.964734 -1.753715

O 0.338056 -1.211717 1.599313

H 1.249653 -2.141731 -1.405170

H 0.459228 -0.822493 -1.456625

H -1.382744 -2.283774 -1.080966

H 0.374633 -1.541990 0.683364

H -2.353886 -2.341147 0.126269

H -3.191445 -0.984850 -1.325022

H 1.179688 1.269934 -2.336810

H -0.335940 1.346047 -1.853452

H 0.768968 1.351678 0.101914

H -4.432524 -0.330267 -2.039541

H -3.815855 0.384453 0.361945

H -2.542916 1.270775 -1.782483

H -2.858701 -0.900826 1.803706

H -2.854571 -2.226908 2.651887

H -0.551282 -1.450561 1.912768

H -2.998363 1.291954 1.313910  
H 0.509409 0.532102 1.381537  
H -0.835582 2.245015 1.133594  
H -1.772431 3.466270 1.474691  
H -1.941439 2.374418 -0.876813

Complex 1 :10\_duuud

Etot = -1031.4098895 Hartree  
O 0.5539550 0.1250310 2.1032330  
O 0.3974240 1.9880670 0.2150410  
O -0.0168840 0.7910260 -2.1901920  
O -0.2796350 -1.9366290 -1.7381670  
O 0.1231560 -2.3007310 0.9142170  
O -2.7941100 -1.8962140 1.2950280  
O -3.0671910 -1.3821900 -1.4204440  
O -2.7625390 1.4619400 -1.8285890  
O -2.3733500 2.4468190 0.6806490  
O -2.2582640 0.5168310 2.4042140  
H 0.5491190 0.8266450 1.4011360  
H 0.2424210 -1.4291660 1.3543050  
H 0.3016280 1.5706650 -0.6704570  
H -0.1177340 -2.0887710 -0.7760460  
H 0.2643950 -2.5688600 -2.2124430  
H -3.0126050 -1.7632820 0.3504730  
H -3.0775730 -0.4288880 -1.5949520  
H -2.2118280 -1.6909320 -1.7522330  
H -2.3918390 1.7002200 1.3576560  
H -3.4127770 1.9504610 -2.3371210  
H -2.7201040 1.8761640 -0.9375050  
H -0.0683410 -0.1793570 -2.1034240  
H -0.9318860 1.0799060 -2.3344460  
H -2.5517520 -0.3595200 2.0595950  
H -1.3278850 0.3893410 2.6397450  
H -0.4651770 2.4006860 0.3849720

H -2.8417700 3.1866430 1.0725160  
H -1.9675770 -2.3934790 1.2938490  
H 0.9981320 -2.7015460 0.9536680  
O 2.8264740 -1.7034420 -0.0574290  
C 3.1013530 -0.5471880 -0.1188220  
O 3.3445980 0.2900330 0.9198560  
C 3.5659440 1.4180160 0.0129990  
C 3.3006110 0.5066340 -1.1811670  
H 1.4381210 0.1452510 2.4804640  
H 2.8204570 2.1895130 0.1907770  
H 4.5830050 1.7891760 0.1234050  
H 4.1427310 0.3084900 -1.8419010  
H 2.4023280 0.7215440 -1.7594440

Complex 5 :1 :5

Etot = - 1031.3944943

C 0.000001 -0.051803 -0.208069  
O -0.000005 1.101411 -0.516700  
O 0.000022 -0.539892 1.058634  
C 0.000018 -1.928368 0.556767  
C -0.000013 -1.398001 -0.873307  
H -0.901735 -1.580294 -1.457432  
H 0.901705 -1.580278 -1.457445  
H -0.910387 -2.429541 0.876062  
H 0.910438 -2.429536 0.876029  
H -1.781835 1.803791 -1.457999  
O -2.724363 1.673087 -1.607956  
H -3.164953 1.996651 -0.794114  
O -3.839907 2.228879 0.806268  
H -3.960429 3.024324 1.325498  
H -3.459048 1.542996 1.397569  
O -2.727683 0.166433 2.170750  
H -1.765591 0.146747 2.150781  
H -3.001751 -0.630271 1.676348

H -4.110580 -2.527972 0.700566  
O -3.324027 -2.002450 0.545308  
H -3.383126 -1.666327 -0.376848  
O -3.189913 -0.915209 -1.924870  
H -3.057568 0.066757 -1.851751  
H -3.736166 -1.064087 -2.698005  
H 1.781797 1.803803 -1.458022  
O 2.724324 1.673109 -1.607994  
H 3.164928 1.996695 -0.794169  
O 3.839971 2.228821 0.806199  
H 3.960597 3.024241 1.325444  
H 3.459028 1.542989 1.397502  
O 2.727743 0.166422 2.170743  
H 1.765652 0.146743 2.150796  
H 3.001783 -0.630289 1.676332  
H 4.110479 -2.528067 0.700573  
O 3.323985 -2.002460 0.545302  
H 3.383144 -1.666340 -0.376850  
O 3.189853 -0.915203 -1.924864  
H 3.057490 0.066763 -1.851745  
H 3.736082 -1.064051 -2.698024

Complex 6 :1 :4

Etot= - 1031.3965854 Hartree  
C 0.687813 -1.134069 -1.129792  
O 0.763321 -0.954317 0.334482  
C 0.438268 0.354877 0.219175  
O 0.323105 1.130971 1.121331  
C 0.328870 0.341593 -1.277739  
H 1.664084 -1.429307 -1.508156  
H -0.101282 -1.844110 -1.366711  
H 1.088196 0.967394 -1.744494  
H -0.664087 0.582831 -1.655818  
O -3.539169 0.130069 1.711381

H -2.889718 -0.612439 1.788839  
H -4.134279 0.062989 2.460580  
O -3.890494 -0.404373 -1.090309  
H -4.016810 -0.175044 -0.156425  
H -3.537006 0.405617 -1.503742  
O -2.246345 2.455599 0.797410  
H -2.729584 1.734655 1.239220  
H -1.320550 2.317043 1.034181  
O -2.645473 1.964070 -1.811757  
H -2.498608 2.282874 -0.890084  
H -3.005674 2.698377 -2.311159  
O -1.821828 -1.917185 1.621760  
H -1.979130 -2.305347 0.736072  
H -0.877898 -1.734540 1.653580  
O -2.335633 -2.552633 -0.990834  
H -2.939391 -1.781587 -1.144140  
H -2.784567 -3.328862 -1.329783  
O 3.607327 -1.263466 1.373543  
O 3.321832 1.462609 1.565074  
O 3.458983 1.503642 -1.237023  
O 4.145041 -1.124620 -1.235859  
H 3.591422 -0.310072 1.605158  
H 2.684850 -1.532659 1.422847  
H 4.039571 -1.281090 -0.264335  
H 2.422493 1.690288 1.820034  
H 3.379787 1.661874 0.613102  
H 4.005210 2.105376 -1.745522  
H 3.814249 0.600915 -1.382405  
H 4.933675 -1.592510 -1.514289

Complex 1 :20 (out)

Etot = -1795.7631755 Hartree  
C -5.217481 -0.233043 0.632457  
C -5.735393 0.841889 -0.291090

C -5.229842 -0.098286 -1.380334  
O -4.785256 -1.041971 -0.359488  
O -5.137647 -0.425230 1.805479  
O -1.282898 -2.814026 0.742213  
O -2.268091 -0.722814 2.040902  
O -1.908895 0.194951 -2.294387  
O -0.063593 1.561573 -3.449382  
O -1.170266 -2.480879 -2.063823  
O 1.384854 -2.535507 -2.670441  
O 2.109804 -0.028997 -3.771608  
O -2.541135 1.202577 0.115031  
O 3.124492 -3.278801 -0.599261  
O 4.554285 -1.230130 -0.053087  
O 1.352643 -3.619491 1.434521  
O 4.153712 0.714789 -1.906974  
O 3.999573 -0.154458 2.470616  
O 1.969009 -1.527558 3.291120  
O -0.161189 0.094541 3.657993  
O 3.104395 2.396297 1.857304  
O 3.311247 3.092428 -0.645020  
O 0.696927 3.438411 -1.479877  
O -0.928396 3.512676 0.549434  
O 0.456267 2.777633 2.784037  
H -2.023788 1.993679 0.336806  
H -2.394491 0.553597 0.837406  
H -1.393606 4.350298 0.496082  
H -0.302046 3.487870 -0.233830  
H -3.187815 -0.893119 2.283069  
H -1.934099 -1.533488 1.577954  
H -1.712415 -0.756851 -2.203911  
H -2.117466 0.535172 -1.391543  
H 0.323815 3.414090 3.489974  
H -0.074906 3.087706 2.017830  
H 0.015503 0.998351 3.358787  
H -0.969546 -0.189704 3.192648  
H 0.492217 2.787748 -2.174107

H 1.640710 3.336910 -1.247624  
H -0.385455 -3.101605 0.975936  
H -1.300156 -2.724692 -0.225607  
H -0.475826 1.999429 -4.197487  
H -0.790331 1.015507 -3.003945  
H -0.205160 -2.533539 -2.306014  
H -1.625330 -3.149744 -2.579503  
H 2.216556 2.493928 2.240441  
H 3.429641 1.507208 2.100332  
H 1.698609 -2.243140 2.695006  
H 1.171636 -0.974680 3.463607  
H 3.247899 2.829046 0.318945  
H 3.917960 3.834775 -0.685423  
H 1.428049 -4.501949 1.803814  
H 2.005847 -3.559887 0.700674  
H 1.325295 0.563714 -3.717473  
H 2.393672 -0.032900 -4.688281  
H 2.006258 -2.792818 -1.964525  
H 1.704734 -1.697526 -3.044821  
H 3.218937 -0.694667 2.805369  
H 4.682756 -0.217798 3.141475  
H 3.468910 0.486174 -2.553817  
H 3.892325 1.569301 -1.523359  
H 3.737536 -3.972432 -0.853073  
H 3.693453 -2.469883 -0.370041  
H 4.462876 -0.509837 -0.716212  
H 4.419321 -0.830607 0.825759  
H -5.989599 -0.560262 -2.009053  
H -4.379111 0.249296 -1.961471  
H -6.810490 1.002241 -0.235358  
H -5.185320 1.773606 -0.184047

Complex 1 :20 (in)  
BPL@(H2O)20

Etot = -1795.7600035Hartree  
C 0.922599 -0.234597 -0.035978  
C -0.449794 0.417370 0.018780  
O -1.068387 -0.882296 -0.243317  
C 0.126839 -1.478954 -0.316712  
O 0.333193 -2.638287 -0.552159  
O -2.230703 -3.814418 -0.848441  
O -2.507586 -2.673036 1.663313  
O -0.212055 -2.128340 2.776217  
O -0.411162 0.386046 3.712587  
O 1.727195 1.982734 2.725621  
O 0.793658 3.583687 0.775341  
O -1.762143 3.691687 0.698506  
O -2.378811 2.830984 -1.811209  
O 2.184067 3.054792 -1.594827  
O 0.166277 2.206040 -3.026422  
O 0.331259 -0.494168 -3.195392  
O -2.188936 -1.752092 -2.894552  
O -3.558236 0.206733 -1.679692  
O -3.923805 -0.247364 1.108264  
O -2.647919 1.658586 2.462077  
O 3.835316 0.405794 1.738725  
O 3.957761 0.996414 -0.843446  
O 3.110317 -1.199768 -2.420451  
O 3.083045 -3.104438 -0.592631  
O 2.795503 -2.225595 2.051437  
H -1.266100 -3.773262 -0.801960  
H -2.422070 -3.180834 -1.558390  
H 3.195898 -2.859133 0.345572  
H -2.668967 -1.906213 -3.711393  
H -2.666489 -1.013881 -2.437158  
H -2.999919 -1.853567 1.483865  
H -2.488629 -3.139912 0.800138  
H 3.669249 -1.461241 -3.155938  
H 3.186540 -1.925577 -1.741206  
H 1.163294 -0.958449 -3.026419

H -0.422148 -1.107613 -3.133892  
H 1.844063 -2.171769 2.211495  
H 3.141182 -1.317680 2.069579  
H -3.193114 1.103757 -1.726179  
H -3.693548 0.016830 -0.734157  
H -0.100515 -2.839389 3.412892  
H -1.069925 -2.350819 2.293385  
H -3.483666 0.475057 1.623168  
H -4.837782 -0.268716 1.400989  
H 3.767210 0.268531 -1.462353  
H 3.395488 1.742095 -1.122106  
H -0.722145 2.429833 -2.709093  
H 0.186826 1.236070 -3.174784  
H 3.904062 0.618850 0.768061  
H 4.711650 0.542067 2.106265  
H -2.930469 3.461992 -2.279271  
H -2.221677 3.202975 -0.913086  
H -0.353252 -0.556684 3.417343  
H -0.471928 0.366908 4.670770  
H -2.338014 2.403470 1.916898  
H -1.883171 1.315474 2.953538  
H 1.429848 2.741399 -2.171262  
H 2.590749 3.791688 -2.056393  
H 1.058597 1.396315 3.113673  
H 2.493946 1.436366 2.481082  
H -2.011784 4.570063 0.995781  
H -0.750818 3.667213 0.717029  
H 1.180051 3.035823 1.496161  
H 1.332147 3.447166 -0.024517  
H 1.564570 0.065440 -0.856687  
H 1.477823 -0.259319 0.895342  
H -0.678729 1.107361 -0.786514  
H -0.768476 0.797533 0.983612  
H 2.129499 -3.268685 -0.646803

Complex 1 :24 (out)

Etot = -2101.4975445 Hartree

O -0.585127 -2.578520 3.071933

O 0.752465 -0.459370 3.946315

O 0.709038 -4.198359 1.178609

O -0.911022 -4.229749 -0.909344

O -0.237278 -2.555340 -3.013193

O -2.253316 -0.902160 -3.355607

O -1.971935 1.823357 -3.376266

O 0.190014 2.882142 -2.428101

O -0.302595 4.560058 -0.254541

O -2.764171 4.025739 0.434951

O -3.998138 2.664231 -1.601273

O 2.428937 1.311805 -1.942950

O 2.070266 -1.281347 -1.935399

O 2.799548 -2.407000 0.422236

O 3.136754 1.941796 0.616298

O 2.943440 -0.349098 2.187019

O -0.428946 2.070278 3.437410

O -2.884421 2.339201 2.715936

O 1.225273 3.554699 1.903046

O -4.180028 0.106046 1.869963

O -3.320185 -2.281752 2.267740

O -3.588050 -3.637422 -0.011219

O -4.555694 -1.948849 -1.920642

O -5.434470 0.336762 -0.669761

H -0.971556 -1.899552 -3.156966

H -0.037970 -2.933572 -3.872425

H 0.995302 2.351063 -2.281511

H 0.081271 3.479610 -1.664221

H -0.655689 -3.675888 -1.672990

H -1.818504 -3.984773 -0.675849

H -1.249832 4.377036 0.012956

H -0.242702 5.505260 -0.408245

H -2.128714 0.070078 -3.351351

H -3.030146 -1.104292 -2.816913  
H -1.122155 2.197612 -2.989813  
H -2.052146 2.213461 -4.249764  
H -4.236014 -2.600605 -1.257397  
H -5.004070 -2.444968 -2.609436  
H -3.325949 2.370676 -2.240929  
H -4.486658 1.873581 -1.333919  
H -4.013211 -4.477676 0.176241  
H -3.502201 -3.163503 0.857290  
H -2.830589 3.456555 1.220495  
H -3.258534 3.573506 -0.280199  
H 1.273743 -1.706886 -2.289450  
H 2.290699 -1.725395 -1.086874  
H 2.235006 0.333127 -1.957932  
H 3.249650 1.409383 -2.440212  
H 2.156909 -3.064777 0.731322  
H 2.817322 -1.698856 1.103535  
H 2.832035 1.742970 -0.298087  
H 4.098510 1.979859 0.559972  
H 0.942175 -5.103518 1.394591  
H 0.126480 -4.237021 0.374874  
H 0.737104 3.953703 1.162417  
H 1.972023 3.082401 1.499874  
H -3.699515 0.910602 2.165362  
H -4.601150 0.297171 1.017118  
H 0.238978 -1.244250 3.624589  
H 0.919009 -0.606282 4.879715  
H -5.225931 -0.482731 -1.160594  
H -6.388023 0.443581 -0.687954  
H 2.899030 0.493692 1.698640  
H 2.231493 -0.330289 2.849541  
H -3.610980 -1.327608 2.112127  
H -3.790067 -2.570186 3.054235  
H -1.481251 -2.441022 2.726573  
H -0.121503 -3.158090 2.440531  
H -1.922452 2.199679 2.986159

H -3.299774 2.791896 3.453476  
H -0.026296 1.200299 3.591892  
H 0.210626 2.607733 2.913598  
O 5.299653 0.675582 -2.225056  
C 5.543075 0.110905 -1.207298  
C 5.727844 -1.301354 -0.712054  
O 5.737628 0.683750 0.011006  
C 5.941132 -0.625727 0.638362  
H 5.175368 -0.803204 1.390784  
H 6.944511 -0.680840 1.056561  
H 4.828588 -1.911164 -0.790359  
H 6.584315 -1.812383 -1.148015

Complex 1 :24 (in)

BPL@(H2O)24

Etot = -2101.5058073 Hartree

C 0.866421 0.058189 0.013994  
O 2.055120 -0.036773 -0.034047  
O 0.148990 1.198274 0.012407  
C -1.107101 0.451873 0.120773  
C -0.326225 -0.854255 0.106121  
H -1.607227 0.697429 1.051085  
H -1.733308 0.660634 -0.740131  
H -0.362320 -1.446919 1.015167  
H -0.464017 -1.488904 -0.765303  
O 4.658307 -0.046895 -0.891697  
O 3.929623 2.075600 0.766476  
O 2.191859 3.533477 -0.985235  
O 0.189242 4.380243 0.518022  
O -2.131172 3.632419 -0.920037  
O -1.480367 2.013020 -2.941032  
O 2.693129 -0.263303 -3.056450  
O 1.331368 1.995402 -3.183285  
O -3.036322 -0.002209 -3.084915

O -4.614957 0.152241 -0.859062  
O -3.665891 -1.795822 0.955236  
O -2.442794 -3.813440 -0.566604  
O -0.363380 -4.322103 0.904269  
O 1.821194 -3.740810 -0.584285  
O 1.055855 -2.543478 -3.038904  
O -1.566441 -2.385655 -2.786203  
O -3.874658 2.377696 0.840665  
O -2.371409 1.665869 3.201896  
O -0.025251 2.770636 2.780559  
O -2.200359 -1.152159 3.145686  
O 2.374741 1.219913 2.865082  
O 2.319761 -1.387196 2.855882  
O 3.901095 -2.264967 0.649690  
O -0.112888 -2.699106 3.240930  
H 1.824045 1.131519 -3.142441  
H 1.572096 2.396926 -4.021382  
H -2.095516 -1.573736 -2.885459  
H -1.949297 -2.909314 -2.056526  
H 1.876919 2.993710 -1.733580  
H 2.770147 2.959639 -0.458353  
H -1.646201 -4.027860 0.007011  
H -2.920235 -4.637857 -0.681875  
H 2.133890 -1.059703 -3.011359  
H 3.268056 -0.271733 -2.275195  
H 0.068643 -2.445240 -2.979512  
H 1.220232 -3.103161 -3.801247  
H 4.361480 0.709042 -0.342870  
H 5.544381 0.161446 -1.194714  
H 1.610134 -3.327155 -1.437519  
H 2.541268 -3.222321 -0.191208  
H 4.556010 2.672514 1.181513  
H 3.346444 1.740329 1.485991  
H -0.283529 -3.828594 1.737449  
H 0.465910 -4.160308 0.401065  
H -0.509823 1.962041 -2.990161

H -1.711155 2.654153 -2.222297  
H -2.395146 0.784612 -3.053875  
H -3.454184 0.034270 -3.948614  
H -1.364008 3.922493 -0.399042  
H -2.774594 3.268416 -0.280827  
H -4.104207 0.117197 -1.699221  
H -5.542772 0.201722 -1.098765  
H 0.319553 5.326069 0.612479  
H 0.934088 4.050506 -0.055273  
H -3.310965 -2.520267 0.411504  
H -4.053251 -1.151413 0.342405  
H 1.483858 -1.877830 2.967394  
H 2.788352 -1.750740 2.085866  
H -1.482747 2.098630 3.066009  
H -2.743314 2.054359 3.996784  
H 4.130720 -1.486688 0.103306  
H 4.720182 -2.736622 0.814929  
H -4.154813 1.580342 0.367959  
H -3.453806 2.090076 1.666423  
H 2.306990 0.221025 2.836089  
H 2.699008 1.425091 3.745508  
H 0.747544 2.183416 2.738443  
H 0.029993 3.342165 1.990064  
H -0.897772 -2.076711 3.252404  
H -0.154741 -3.198631 4.059510  
H -2.234664 -0.186162 3.252686  
H -2.790078 -1.373745 2.392046

# Conclusion générale et perspectives

## Conclusion

Au cours de cette thèse, nous avons mis en place une méthodologie permettant une étude rigoureuse de la micro-hydratation sur les complexes  $\text{BPL}(\text{H}_2\text{O})_{1-3}$ . Celle-ci se décompose en quatre étapes. La première étape consiste à localiser les régions nucléophiles et électrophiles des partenaires isolés, grâce à l'étude du potentiel électrostatique moléculaire sur une surface d'iso-densité électronique. La deuxième étape consiste à orienter les partenaires de façon à faire coïncider une zone électrophile d'un monomère avec une zone nucléophile de l'autre monomère et réciproquement. La troisième étape repose sur l'optimisation géométrique des structures proposées à l'issue de la deuxième étape. Enfin, la quatrième et dernière étape permet une caractérisation qualitative et quantitative des interactions non-covalentes par la théorie AIM.

Les études des complexes  $\text{BPL}(\text{H}_2\text{O})_{1-3}$  au niveau MP2/aug-cc-pVDZ ont montré entre autres que la rupture du réseau de liaison hydrogène au sein du cluster d'eau entraîne une déstabilisation du complexe. De plus, la présence d'une interaction de type liaison hydrogène entre l'hydrogène d'une molécule d'eau et l'oxygène du groupement carbonyle de la BPL est présente systématiquement dans les complexes les plus stables. Enfin, chaque molécule présente dans le complexe est simultanément donneur de liaison hydrogène et accepteur de liaison hydrogène. Pour pouvoir envisager une étude avec un nombre croissant de molécule d'eau, nous avons évalué la fiabilité d'un certain nombre de méthodes DFT corrigées de la dispersion. La méthode LC- $\omega$ PBE-GD3BJ/6-311++G(d,p) a été retenue pour l'étude des complexes  $\text{BPL}(\text{H}_2\text{O})_{4-24}$ .

La structure des complexes  $\text{BPL}(\text{H}_2\text{O})_{4,5}$  a été rationalisée grâce à l'analyse du potentiel électrostatique moléculaire des partenaires isolées. L'analyse AIM démontre que la fixation du cluster d'eau sur la BPL se fait par l'intermédiaire de quatre interactions entre le cluster d'eau et le soluté. Cette analyse a également montré que les clusters d'eau  $(\text{H}_2\text{O})_{6,10}$  interagissent d'une façon similaire avec le soluté que dans les complexes  $\text{BPL}(\text{H}_2\text{O})_{4,5}$ . La scission du cluster  $(\text{H}_2\text{O})_{10}$  en deux clusters  $(\text{H}_2\text{O})_5$  pour la multiplication des interactions eau-soluté sur les deux faces de la BPL dans une structure de type "sandwich"  $(\text{H}_2\text{O})_5\text{BPL}(\text{H}_2\text{O})_5$  aboutit à une déstabilisation du complexe.

Des clusters d'eau comprenant 20 à 24 molécules d'eau peuvent interagir avec l'une des deux faces trapézoïdales de la BPL. Alternativement, ces cages d'eau  $(\text{H}_2\text{O})_{20,24}$  peuvent également servir d'hôte pour le soluté. Pour ce qui est du cas du complexe  $\text{BPL}(\text{H}_2\text{O})_{20}$ , le conformère  $\text{BPL}@(\text{H}_2\text{O})_{20}$  est moins stable que le conformère dans lequel la cage interagit avec l'une des faces trapézoïdales de la BPL. L'encapsulation de la BPL par une cage composée de 20 molécules d'eau mène à une rupture du réseau de liaison hydrogène

causant une déstabilisation du complexe. Une telle rupture du réseau de liaison hydrogène n'est pas observé avec la cage  $5^{12} 6^2$  composée de 24 molécule d'eau. Enfin, et c'est un point important, il faut mentionner que les complexes  $\text{BPL}(\text{H}_2\text{O})_{20,24}$  sont moins stable que la somme des monomères dans leur minimum global. Les clusters d'eau  $(\text{H}_2\text{O})_{20,24}$  adoptent une structure cuboïde et non sphérique dans leur conformation la plus stable [219].

Nous avons également développé au cours de cette thèse une nouvelle classe de méthodes hybrides fonction d'onde/DFT basée sur une décomposition de l'interaction bi-électronique de type CAM. Nous avons considéré une approximation de type MP2 pour la description de l'interaction  $w_{ee}^{\text{lr},\mu}(\mathbf{r}_{12}) + \lambda w_{ee}^{\text{sr},\mu}(\mathbf{r}_{12})$ , l'interaction complémentaire étant approximée par une fonctionnelle de la densité de courte portée. Ce schéma que l'on nomme RSDH est une généralisation des doubles hybrides usuelles reposant sur une séparation linéaire de l'interaction bi-électronique (DS1DH) et comprenant un paramètre. Dans ce cas le paramètre  $\mu$  vaut zéro. Ce schéma est également une généralisation des méthodes hybrides à séparation de portée (RSH+MP2). Dans ce cas le paramètre  $\lambda$  vaut zéro.

Nous avons proposé un certain nombre d'approximations pour la fonctionnelle d'échange-corrélation de courte portée. Pour cela nous avons considéré les limites  $\mu = 0$  et  $\mu \rightarrow \infty$ . Les approximations pour la fonctionnelle de corrélation de courte portée ont été évaluées dans un premier temps sur le gaz d'électrons uniforme. Nous avons dans un second temps évalué le méthode RSDH avec une fonctionnelle d'échange-corrélation de courte portée de type PBE. Les approximations 1 à 5 présentées dans cette thèse ont pu être appliquées sur un ensemble d'énergies d'atomisation et de barrières réactionnelles (les ensembles AE6 et BH6). Cela nous a permis de trouver les paramètres  $\mu$  et  $\lambda$  optimaux. Avec ces paramètres optimaux, la méthode RSDH a une convergence relativement rapide en fonction de la taille de la base, ce qui peut s'expliquer par le fait qu'il n'existe qu'une fraction modeste de corrélation MP2 à courte distance qui est d'environ  $\lambda = 0.35$ . Nous avons ensuite testé les deux meilleures approximations GGA sur un ensemble plus grand d'énergie d'atomisation et de barrières réactionnelles, les ensembles AE49 et DBH24. Ces approximations ont également été testées sur un ensemble d'interaction non-covalentes, l'ensemble S22.

Les résultats montrent que la méthode RSDH est globalement plus précise que les méthodes RSH + MP2, DS1DH et MP2. Parmi les approximations testées, nous recommandons l'approximation 3 avec les paramètres  $(\mu, \lambda) = (0.46, 0.58)$  pour des applications générales.

## Perspectives

En ce qui concerne l'étude de la micro-hydratation des composés organiques, nous envisageons de poursuivre cette étude :

- en tenant compte des effets des dispersions à plusieurs corps, des effets de température et de pression afin de pouvoir se rapprocher davantage de la réalité expérimentale.
- Il faudrait également mettre en place une méthodologie permettant de prédire le nombre minimal de molécule d'eau nécessaire à l'encapsulation du soluté. Pour cela, nous pouvons envisager de prédire les dimensions de la cage d'eau en se basant sur les points stationnaires des complexes mono-hydratés.
- Nous souhaitons également comprendre les mécanismes de micro-hydratation d'autres composés neutres ne présentant pas de moment dipolaire permanent (tel que le dioxyde de carbone ou le méthane qui présente un intérêt environnemental) ou de composés organiques chargés ou radicalaires.

En ce qui concerne les perspectives sur le développement de fonctionnelle doubles hybrides à séparation de portée :

- on pourrait développer une fonctionnelle de corrélation de courte portée pure avec polarisation de spin (en vue de l'application sur des systèmes à couches ouverte), dans le but d'appliquer l'approximation 7. Cette approximation est la plus précise sur le gaz d'électrons homogène.
- On pourrait développer des meta-GGA de courte portée et satisfaire davantage de contraintes exactes connues comme par exemple les relations de transformations d'échelle non-uniforme ("non-uniform density scaling"), le fait que l'énergie de corrélation doit être nulle pour un système mono-électronique, la borne de Lieb-Oxford pour les densités à deux électrons.
- Remplacer l'approximation de type MP2 par une approximation de type RPA en vue d'une meilleure description des interactions non-covalentes. Cette perspective est actuellement en cours de réalisation.
- Étendre la méthode RSDH à des systèmes périodiques en vue de l'étude de clathrates d'eau.

Enfin, nous devons travailler sur l'efficacité du code pour réduire le coût calculatoire. Nous devons optimiser le code pour avoir une meilleure parallélisation, utiliser des méthodes basées sur le "*density fitting*" qui est un moyen moins coûteux d'évaluer les intégrales bi-électroniques.

# Bibliographie

- [1] A. Stone, *The theory of intermolecular forces* (OUP Oxford, 2013).
- [2] B. Jeziorski, R. Moszynski and K. Szalewicz, *Chemical Reviews* **94**, 1887 (1994).
- [3] A. Martín Pendás, M. Blanco and E. Francisco, *The Journal of Chemical Physics* **125**, 184112 (2006).
- [4] R. F. Bader, *Encyclopedia of computational Chemistry* **1** (2002).
- [5] K. Morokuma, *Accounts of Chemical Research* **10**, 294 (1977).
- [6] D. Philp and J. F. Stoddart, *Angewandte Chemie International Edition* **35**, 1154 (1996).
- [7] D. Dargent, E. Zins, B. Madebène and M. E. Alikhani, *Journal of Molecular Modeling* **21**, 214 (2015).
- [8] E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg, P. Hobza et al., *Pure and applied chemistry* **83**, 1637 (2011).
- [9] G. A. Jeffrey and G. A. Jeffrey, *An introduction to hydrogen bonding*, vol. 32 (Oxford university press New York, 1997).
- [10] S. J. Grabowski, *Chemical Reviews* **111**, 2597 (2011).
- [11] J. Kraitchman, *American Journal of Physics* **21**, 17 (1953).
- [12] P. Popelier, *The Journal of Physical Chemistry A* **102**, 1873 (1998).
- [13] E. R. Johnson, S. Keinan, P. Mori-Sanchez, J. Contreras-Garcia, A. J. Cohen and W. Yang, *Journal of the American Chemical Society* **132**, 6498 (2010).
- [14] M. Alikhani, F. Fuster and B. Silvi, *Structural Chemistry* **16**, 203 (2005).
- [15] F. Weinhold, JE Carpenter In *The Structure of Small Molecules and Ions*; R. Naaman and Z. Vager, Eds (1989).
- [16] P. Politzer, J. S. Murray, T. Clark and G. Resnati, *Physical Chemistry Chemical Physics* **19**, 32166 (2017).

- [17] A. S. Mahadevi and G. N. Sastry, *Chemical reviews* **116**, 2775 (2016).
- [18] J. Hoja, A. F. Sax and K. Szalewicz, *Chemistry-A European Journal* **20**, 2292 (2014).
- [19] A. Szabo and N. S. Ostlund, *Modern quantum chemistry : introduction to advanced electronic structure theory* (Courier Corporation, 2012).
- [20] M. O. Sinnokrot and C. D. Sherrill, *The Journal of Physical Chemistry A* **110**, 10656 (2006).
- [21] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen and A. K. Wilson, *Chemical Physics Letters* **286**, 243 (1998).
- [22] W. Kohn and L. J. Sham, *Physical review* **140**, A1133 (1965).
- [23] Y. Zhao and D. G. Truhlar, *Theoretical Chemistry Accounts* **120**, 215 (2008).
- [24] J.-D. Chai and M. Head-Gordon, *The Journal of chemical physics* **128**, 084106 (2008).
- [25] G. A. DiLabio and A. Otero-de-la Roza, *Noncovalent interactions in density-functional theory*, vol. 29 (John Wiley and Sons, Hoboken, NJ, 2016).
- [26] S. Grimme, A. Hansen, J. G. Brandenburg and C. Bannwarth, *Chemical reviews* **116**, 5105 (2016).
- [27] J. Toulouse, F. Colonna and A. Savin, *Physical Review A* **70**, 062505 (2004).
- [28] O. Franck, *A closer look at wave-function/density-functional hybrid methods*, Ph.D. thesis, Université Pierre et Marie Curie-Paris VI (2016).
- [29] O. Franck, B. Mussard, E. Luppi and J. Toulouse, *The Journal of chemical physics* **142**, 074107 (2015).
- [30] J. G. Angyán, I. C. Gerber, A. Savin and J. Toulouse, *Physical Review A* **72**, 012510 (2005).
- [31] J. Toulouse, W. Zhu, A. Savin, G. Jansen and J. G. Angyán, *The Journal of chemical physics* **135**, 084119 (2011).
- [32] D. E. Taylor, J. G. Ángyán, G. Galli, C. Zhang, F. Gygi, K. Hirao, J. W. Song, K. Rahul, O. Anatole von Lilienfeld, R. Podeszwa et al., *The Journal of chemical physics* **145**, 124105 (2016).
- [33] G. A. DiLabio, E. R. Johnson and A. Otero-de-la Roza, *Physical Chemistry Chemical Physics* **15**, 12821 (2013).
- [34] C. Pérez, J. L. Neill, M. T. Muckle, D. P. Zaleski, I. Peña, J. C. Lopez, J. L. Alonso and B. H. Pate, *Angewandte Chemie* **127**, 993 (2015).

- [35] D. Cremer, Wiley Interdisciplinary Reviews : Computational Molecular Science **1**, 509 (2011).
- [36] P. Hohenberg and W. Kohn, Physical review **136**, B864 (1964).
- [37] M. Levy, Proceedings of the National Academy of Sciences **76**, 6062 (1979).
- [38] M. Levy, Physical Review A **26**, 1200 (1982).
- [39] E. H. Lieb, in *Inequalities* (Springer, 2002), pp. 269–303.
- [40] P. Dirac, in *Proc. Cambridge Phil. Soc.* (1931), vol. 27, p. 240.
- [41] J. C. Slater, Physical Review **81**, 385 (1951).
- [42] D. M. Ceperley and B. Alder, Physical Review Letters **45**, 566 (1980).
- [43] S. H. Vosko, L. Wilk and M. Nusair, Canadian Journal of Physics **58**, 1200 (1980).
- [44] J. P. Perdew and Y. Wang, Physical Review B **45**, 13244 (1992).
- [45] P.-F. Loos and P. M. Gill, Wiley Interdisciplinary Reviews : Computational Molecular Science **6**, 410 (2016).
- [46] J. P. Perdew and S. Kurth, *Density functionals for non-relativistic Coulomb systems in the new century* (Springer, 2003).
- [47] A. D. Becke, Physical review A **38**, 3098 (1988).
- [48] C. Lee, W. Yang and R. G. Parr, Physical review B **37**, 785 (1988).
- [49] G. Oliver and J. Perdew, Physical Review A **20**, 397 (1979).
- [50] E. H. Lieb and S. Oxford, International Journal of Quantum Chemistry **19**, 427 (1981).
- [51] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou and K. Burke, Physical Review Letters **100**, 136406 (2008).
- [52] J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, Physical Review Letters **91**, 146401 (2003).
- [53] J. Sun, A. Ruzsinszky and J. P. Perdew, Physical review letters **115**, 036402 (2015).
- [54] A. D. Becke, The Journal of Chemical Physics **98**, 5648 (1993).
- [55] C. Adamo and V. Barone, The Journal of Chemical Physics **110**, 6158 (1999).
- [56] M. Ernzerhof and G. E. Scuseria, The Journal of Chemical Physics **110**, 5029 (1999).
- [57] Y. Zhao, B. J. Lynch and D. G. Truhlar, The Journal of Physical Chemistry A **108**, 4786 (2004).
- [58] F. Neese, T. Schwabe and S. Grimme, The Journal of chemical physics **126**, 124115 (2007).

- [59] S. Grimme, *The Journal of chemical physics* **124**, 034108 (2006).
- [60] A. Görling and M. Levy, *Physical Review B* **47**, 13105 (1993).
- [61] A. Görling and M. Levy, *Physical Review A* **50**, 196 (1994).
- [62] A. Tarnopolsky, A. Karton, R. Sertchook, D. Vuzman and J. M. Martin, *The Journal of Physical Chemistry A* **112**, 3 (2008).
- [63] A. Karton, A. Tarnopolsky, J.-F. Lamère, G. C. Schatz and J. M. Martin, *The Journal of Physical Chemistry A* **112**, 12868 (2008).
- [64] D. C. Graham, A. S. Menon, L. Goerigk, S. Grimme and L. Radom, *The Journal of Physical Chemistry A* **113**, 9861 (2009).
- [65] J. Sancho-García and A. Pérez-Jiménez, *The Journal of chemical physics* **131**, 084108 (2009).
- [66] C. Adamo and V. Barone, *The Journal of chemical physics* **108**, 664 (1998).
- [67] A. Mohajeri and M. Alipour, *The Journal of chemical physics* **136**, 124111 (2012).
- [68] E. Fromager, *The Journal of chemical physics* **135**, 244106 (2011).
- [69] T. Schwabe and S. Grimme, *Physical Chemistry Chemical Physics* **9**, 3397 (2007).
- [70] O. A. Vydrov and T. Van Voorhis, *The Journal of chemical physics* **133**, 244103 (2010).
- [71] J. Aragó, E. Ortí and J. C. Sancho-García, *Journal of chemical theory and computation* **9**, 3437 (2013).
- [72] S. Grimme, *The Journal of chemical physics* **118**, 9095 (2003).
- [73] J.-D. Chai and M. Head-Gordon, *The Journal of chemical physics* **131**, 174105 (2009).
- [74] S. Kozuch and J. M. Martin, *Journal of computational chemistry* **34**, 2327 (2013).
- [75] S. Kozuch, D. Gruzman and J. M. Martin, *The Journal of Physical Chemistry C* **114**, 20801 (2010).
- [76] S. Kozuch and J. M. Martin, *Physical Chemistry Chemical Physics* **13**, 20104 (2011).
- [77] B. Chan and L. Radom, *Journal of chemical theory and computation* **7**, 2852 (2011).
- [78] Y. Jung, R. C. Lochan, A. D. Dutoi and M. Head-Gordon, *The Journal of chemical physics* **121**, 9793 (2004).
- [79] J. Almlöf, *Chemical physics letters* **181**, 319 (1991).
- [80] T. Benighaus, R. A. DiStasio, R. C. Lochan, J.-D. Chai and M. Head-Gordon, *The Journal of Physical Chemistry A* **112**, 2702 (2008).

- [81] L. Goerigk and S. Grimme, *Journal of Chemical Theory and Computation* **7**, 291 (2010).
- [82] K. Sharkas, J. Toulouse and A. Savin, *The Journal of chemical physics* **134**, 064113 (2011).
- [83] J. Toulouse, K. Sharkas, E. Brémond and C. Adamo, *Communication : rationale for a new class of double-hybrid approximations in density-functional theory* (2011).
- [84] E. Brémond and C. Adamo, *The Journal of chemical physics* **135**, 024106 (2011).
- [85] J.-D. Chai and S.-P. Mao, *Chemical Physics Letters* **538**, 121 (2012).
- [86] S. M. Souvi, K. Sharkas and J. Toulouse, *The Journal of chemical physics* **140**, 084107 (2014).
- [87] R. Haunschuld and W. Klopper, *Theoretical Chemistry Accounts* **131**, 1112 (2012).
- [88] I. Y. Zhang and X. Xu, *The journal of physical chemistry letters* **4**, 1669 (2013).
- [89] Y. Cornaton and E. Fromager, *International Journal of Quantum Chemistry* **114**, 1199 (2014).
- [90] M. Levy and J. P. Perdew, *Physical Review A* **32**, 2010 (1985).
- [91] M. Levy, W. Yang and R. G. Parr, *The Journal of chemical physics* **83**, 2334 (1985).
- [92] M. Levy, *Physical Review A* **43**, 4637 (1991).
- [93] M. Levy and J. P. Perdew, *Physical Review B* **48**, 11638 (1993).
- [94] J. G. Ángyán, *Physical Review A* **78**, 022510 (2008).
- [95] E. Fromager and H. J. A. Jensen, *Physical Review A* **78**, 022504 (2008).
- [96] A. Savin, *In Recent Developments of Modern Density Functional Theory ; Seminario, JM, Ed* (Elsevier : Amsterdam, 1996).
- [97] J. Toulouse, F. Colonna and A. Savin, *The Journal of chemical physics* **122**, 014110 (2005).
- [98] R. Pollet, A. Savin, T. Leininger and H. Stoll, *The Journal of chemical physics* **116**, 1250 (2002).
- [99] E. Fromager, J. Toulouse and H. J. A. Jensen, *The Journal of chemical physics* **126**, 074111 (2007).
- [100] E. Goll, H.-J. Werner and H. Stoll, *Physical Chemistry Chemical Physics* **7**, 3917 (2005).
- [101] W. Zhu, J. Toulouse, A. Savin and J. G. Ángyán, *The Journal of chemical physics* **132**, 244108 (2010).

- [102] J. Toulouse, W. Zhu, J. G. Angyán and A. Savin, *Physical Review A* **82**, 032502 (2010).
- [103] E. Fromager, R. Cimiraglia and H. J. A. Jensen, *Physical Review A* **81**, 024502 (2010).
- [104] R. F. Stewart, *Chemical Physics Letters* **65**, 335 (1979).
- [105] R. F. Bader, M. T. Carroll, J. R. Cheeseman and C. Chang, *Journal of the American Chemical Society* **109**, 7968 (1987).
- [106] R. K. Pathak and S. R. Gadre, *The Journal of Chemical Physics* **93**, 1770 (1990).
- [107] P. Sjoberg and P. Politzer, *Journal of Physical Chemistry* **94**, 3959 (1990).
- [108] P. Politzer, S. J. Landry and T. Waernheim, *The Journal of Physical Chemistry* **86**, 4767 (1982).
- [109] S. R. Gadre and S. S. Pundlik, *The Journal of Physical Chemistry B* **101**, 3298 (1997).
- [110] C. F. Matta and R. J. Boyd, *An introduction to the quantum theory of atoms in molecules* (Wiley Online Library, 2007).
- [111] R. F. Bader, *The Journal of Physical Chemistry A* **113**, 10391 (2009).
- [112] S. J. Grabowski, W. A. Sokalski, E. Dyguda and J. Leszczyński, *The Journal of Physical Chemistry B* **110**, 6444 (2006).
- [113] I. Rozas, *Physical Chemistry Chemical Physics* **9**, 2782 (2007).
- [114] C. Kalai and J. Toulouse, *The Journal of Chemical Physics* **148**, 164105 (2018).
- [115] H. S. Yu, S. L. Li and D. G. Truhlar, *The Journal of chemical physics* **145**, 130901 (2016).
- [116] B. Mussard and J. Toulouse, *Molecular Physics* **115**, 161 (2017).
- [117] L. Goerigk and S. Grimme, *Wiley Interdisciplinary Reviews : Computational Molecular Science* **4**, 576 (2014).
- [118] Y. Cornaton, A. Stoyanova, H. J. A. Jensen and E. Fromager, *Physical Review A* **88**, 022516 (2013).
- [119] T. Yanai and D. Tew, *Chem. Phys. Lett.* **393**, 51 (2004).
- [120] M. J. Peach, T. Helgaker, P. Salek, T. W. Keal, O. B. Lutnæs, D. J. Tozer and N. C. Handy, *Physical Chemistry Chemical Physics* **8**, 558 (2006).
- [121] M. J. Peach, A. J. Cohen and D. J. Tozer, *Physical Chemistry Chemical Physics* **8**, 4543 (2006).

- [122] A. J. Garza, I. W. Bulik, T. M. Henderson and G. E. Scuseria, *Physical Chemistry Chemical Physics* **17**, 22412 (2015).
- [123] K. Sharkas, A. Savin, H. J. A. Jensen and J. Toulouse, *The Journal of chemical physics* **137**, 044104 (2012).
- [124] P. M. Gill, R. D. Adamson and J. A. Pople, *Molecular Physics* **88**, 1005 (1996).
- [125] H. Iikura, T. Tsuneda, T. Yanai and K. Hirao, *The Journal of Chemical Physics* **115**, 3540 (2001).
- [126] J. Heyd, G. E. Scuseria and M. Ernzerhof, *The Journal of chemical physics* **118**, 8207 (2003).
- [127] J. Toulouse, A. Savin and H.-J. Flad, *International journal of quantum chemistry* **100**, 1047 (2004).
- [128] E. Goll, H.-J. Werner, H. Stoll, T. Leininger, P. Gori-Giorgi and A. Savin, *Chemical physics* **329**, 276 (2006).
- [129] E. Goll, M. Ernst, F. Moegle-Hofacker and H. Stoll, *The Journal of chemical physics* **130**, 234112 (2009).
- [130] S. Pazziani, S. Moroni, P. Gori-Giorgi and G. B. Bachelet, *Physical Review B* **73**, 155111 (2006).
- [131] M. Levy **337** (2013).
- [132] K. Burke, J. P. Perdew and M. Ernzerhof, *The Journal of chemical physics* **109**, 3760 (1998).
- [133] M. Seidl, P. Gori-Giorgi and A. Savin, *Physical Review A* **75**, 042511 (2007).
- [134] J. Toulouse, *Extension multidéterminantale de la méthode de Kohn-Sham en théorie de la fonctionnelle de la densité par décomposition de l'interaction électronique en contributions de longue portée et de courte portée*, Ph.D. thesis, Université Pierre et Marie Curie-Paris VI (2005).
- [135] J. Toulouse and A. Savin, *Journal of Molecular Structure : THEOCHEM* **762**, 147 (2006).
- [136] L. Zecca, P. Gori-Giorgi, S. Moroni and G. B. Bachelet, *Physical Review B* **70**, 205127 (2004).
- [137] A. Grüneis, M. Marsman, J. Harl, L. Schimka and G. Kresse, *The Journal of chemical physics* **131**, 154115 (2009).
- [138] D. L. Freeman, *Physical Review B* **15**, 5512 (1977).

- [139] H. Werner, P. Knowles, G. Knizia, F. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut et al., Google Scholar (2015).
- [140] B. J. Lynch and D. G. Truhlar, *The Journal of Physical Chemistry A* **107**, 8996 (2003).
- [141] T. H. Dunning Jr, *The Journal of chemical physics* **90**, 1007 (1989).
- [142] R. Haunschild and W. Klopper, *The Journal of Chemical Physics* **136**, 164102 (2012).
- [143] P. L. Fast, J. C. Corchado, M. L. Sánchez and D. G. Truhlar, *The Journal of Physical Chemistry A* **103**, 5129 (1999).
- [144] L. Curtiss, *J. chem. Phys.* **94**, 7221 (1991).
- [145] L. A. Curtiss, K. Raghavachari, P. C. Redfern and J. A. Pople, *The Journal of Chemical Physics* **106**, 1063 (1997).
- [146] B. Mussard, P. Reinhardt, J. G. Ángyán and J. Toulouse, *The Journal of chemical physics* **142**, 154123 (2015).
- [147] J. Zheng, Y. Zhao and D. G. Truhlar, *Journal of Chemical Theory and Computation* **3**, 569 (2007).
- [148] J. Zheng, Y. Zhao and D. G. Truhlar, *Journal of chemical theory and computation* **5**, 808 (2009).
- [149] D. E. Woon and T. H. Dunning Jr, *The Journal of chemical physics* **98**, 1358 (1993).
- [150] P. Jurečka, J. Šponer, J. Černý and P. Hobza, *Physical Chemistry Chemical Physics* **8**, 1985 (2006).
- [151] T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall and C. D. Sherrill, *The Journal of chemical physics* **132**, 144104 (2010).
- [152] M. Schütz, G. Hetzer and H.-J. Werner, *The Journal of chemical physics* **111**, 5691 (1999).
- [153] J. P. Perdew, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [154] T. Helgaker, W. Klopper, H. Koch and J. Noga, *The Journal of Chemical Physics* **106**, 9639 (1997).
- [155] S. Grimme and F. Neese, *The Journal of chemical physics* **127**, 154116 (2007).
- [156] E. Rebolini, A. Savin and J. Toulouse, *Molecular Physics* **111**, 1219 (2013).
- [157] E. D. Hedegård, F. Heiden, S. Knecht, E. Fromager and H. J. A. Jensen, *The Journal of Chemical Physics* **139**, 184308 (2013).

- [158] E. Rebolini and J. Toulouse, *The Journal of Chemical Physics* **144**, 094107 (2016).
- [159] J. Toulouse, P. Gori-Giorgi and A. Savin, *International journal of quantum chemistry* **106**, 2026 (2006).
- [160] E. Goll, T. Leininger, F. R. Manby, A. Mitrushchenkov, H.-J. Werner and H. Stoll, *Physical Chemistry Chemical Physics* **10**, 3353 (2008).
- [161] V. Vaida, *The Journal of chemical physics* **135**, 020901 (2011).
- [162] C. Li, J. Chen, H.-B. Xie, Y. Zhao, D. Xia, T. Xu, X. Li and X. Qiao, *Environmental Science & Technology* **51**, 5043 (2017).
- [163] R. B. Gerber, M. E. Varner, A. D. Hammerich, S. Riikonen, G. Murdachaew, D. Shemesh and B. J. Finlayson-Pitts, *Accounts of Chemical Research* **48**, 399 (2015).
- [164] K. Liu, J. Cruzan, R. Saykally et al., *SCIENCE-NEW YORK THEN WASHINGTON-* pp. 929–932 (1996).
- [165] G. Maes and J. Smets, *The Journal of Physical Chemistry* **97**, 1818 (1993).
- [166] J. M. Guevara-Vela, E. Romero-Montalvo, V. A. M. Gómez, R. Chávez-Calvillo, M. García-Revilla, E. Francisco, Á. M. Pendás and T. Rocha-Rinza, *Physical Chemistry Chemical Physics* **18**, 19557 (2016).
- [167] T. Steiner, *Angewandte Chemie International Edition* **41**, 48 (2002).
- [168] E. Miliordos, E. Aprà and S. S. Xantheas, *Journal of chemical theory and computation* **12**, 4004 (2016).
- [169] F. Calvo, M.-C. Bacchus-Montabonel and C. Clavaguéra, *The Journal of Physical Chemistry A* **120**, 2380 (2016).
- [170] V. Vaquero, M. E. Sanz, I. Pena, S. Mata, C. Cabezas, J. C. Lopez and J. L. Alonso, *The Journal of Physical Chemistry A* **118**, 2584 (2014).
- [171] S. Mata, V. Cortijo, W. Caminati, J. L. Alonso, M. E. Sanz, J. C. López and S. Blanco, *The Journal of Physical Chemistry A* **114**, 11393 (2010).
- [172] T. Emilsson, H. Gutowsky, G. De Oliveira and C. Dykstra, *The Journal of Chemical Physics* **112**, 1287 (2000).
- [173] M. Cirtog, M. Alikhani, B. Madebene, P. Soulard, P. Asselin and B. Tremblay, *The Journal of Physical Chemistry A* **115**, 6688 (2011).
- [174] Y. Liu, J. Zhao, F. Li and Z. Chen, *Journal of computational chemistry* **34**, 121 (2013).
- [175] Z. P. Shields, J. S. Murray and P. Politzer, *International Journal of Quantum Chemistry* **110**, 2823 (2010).

- [176] S. R. Gadre, K. Babu and A. P. Rendell, *The Journal of Physical Chemistry A* **104**, 8976 (2000).
- [177] S. F. Boys and F. d. Bernardi, *Molecular Physics* **19**, 553 (1970).
- [178] A. Legon and D. Millen, *Accounts of Chemical Research* **20**, 39 (1987).
- [179] J. M. Guevara-Vela, R. Chávez-Calvillo, M. García-Revilla, J. Hernández-Trujillo, O. Christiansen, E. Francisco, Á. Martín Pendás and T. Rocha-Rinza, *Chemistry-A European Journal* **19**, 14304 (2013).
- [180] C. Kalai, E.-L. Zins and M. E. Alikhani, *Theoretical Chemistry Accounts* **136**, 48 (2017).
- [181] M. Rodgers and P. Armentrout, *Chemical reviews* **116**, 5642 (2016).
- [182] P. Hobza and K. Müller-Dethlefs, *Non-covalent interactions : theory and experiment*, vol. 2 (Royal Society of Chemistry, 2010).
- [183] Z. Wei, D. Chen, H. Zhao, Y. Li, J. Zhu and B. Liu, *The Journal of chemical physics* **140**, 02B621\_1 (2014).
- [184] A. Bouchet, M. Schütz and O. Dopfer, *ChemPhysChem* **17**, 232 (2016).
- [185] G. Feng, Q. Gou, L. Evangelisti, L. Spada, S. Blanco and W. Caminati, *Physical Chemistry Chemical Physics* **18**, 23651 (2016).
- [186] G. Fogarasi and P. G. Szalay, *Physical Chemistry Chemical Physics* **17**, 29880 (2015).
- [187] D. Chen, Z. Wei and B. Liu, *Journal of Molecular Modeling* **21**, 234 (2015).
- [188] V. Riffet, G. Bouchoux and G. Frison, *The Journal of Physical Chemistry B* **119**, 11527 (2015).
- [189] C. Pérez, D. P. Zaleski, N. A. Seifert, B. Temelso, G. C. Shields, Z. Kisiel and B. H. Pate, *Angewandte Chemie International Edition* **53**, 14368 (2014).
- [190] D. Dargent, E. Zins, B. Madebène and M. Alikhani, *Theoretical Chemistry Accounts* **135**, 32 (2016).
- [191] T. Zelený, P. Hobza and M. Kabeláč, *Physical Chemistry Chemical Physics* **11**, 3430 (2009).
- [192] S. S. Park, S. Lee, Y. S. Won and Y. J. Ahn, *Chemical Physics* **441**, 128 (2014).
- [193] P. Kumar and N. Sathyamurthy, *The Journal of Physical Chemistry A* **115**, 14276 (2011).
- [194] S. Alavi, R. Susilo and J. A. Ripmeester, *The Journal of chemical physics* **130**, 174501 (2009).

- [195] S. R. Gadre, S. D. Yeole and N. Sahu, *Chemical reviews* **114**, 12132 (2014).
- [196] J. Murray and K. Sen, Amsterdam, The Netherlands (1996).
- [197] M. H. Kolar and P. Hobza, *Chemical reviews* **116**, 5155 (2016).
- [198] J. S. Murray, P. Lane, T. Clark, K. E. Riley and P. Politzer, *Journal of Molecular Modeling* **18**, 541 (2012).
- [199] A. Bauzá, T. J. Mooibroek and A. Frontera, *ChemPhysChem* **16**, 2496 (2015).
- [200] A. Kumar, S. R. Gadre, N. Mohan and C. H. Suresh, *The Journal of Physical Chemistry A* **118**, 526 (2014).
- [201] G. Singh, R. Verma and S. R. Gadre, *The Journal of Physical Chemistry A* **119**, 13055 (2015).
- [202] M. Prakash, V. Subramanian and S. R. Gadre, *The Journal of Physical Chemistry A* **113**, 12260 (2009).
- [203] A. J. Misquitta, A. J. Stone and S. L. Price, *Journal of chemical theory and computation* **4**, 19 (2008).
- [204] E. Zins and M. Alikhani, *Molecular Physics* **114**, 1317 (2016).
- [205] K. L. Copeland and G. S. Tschumper, *Journal of chemical theory and computation* **8**, 1646 (2012).
- [206] U. Koch and P. L. Popelier, *The Journal of Physical Chemistry* **99**, 9747 (1995).
- [207] M. Frisch, Gaussian, Inc., Pittsburgh PA (1992).
- [208] S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *The Journal of Chemical Physics* **132**, 154104 (2010).
- [209] O. A. Vydrov and G. E. Scuseria, *The Journal of chemical physics* **125**, 234109 (2006).
- [210] O. A. Vydrov, J. Heyd, A. V. Krukau and G. E. Scuseria, *The Journal of Chemical Physics* **125**, 074106 (2006).
- [211] O. A. Vydrov, G. E. Scuseria and J. P. Perdew, *The Journal of Chemical Physics* **126**, 154109 (2007).
- [212] T. A. Keith, TK Gristmill Software, Overland Park KS, USA (2014).
- [213] E. Miliordos, E. Aprà and S. S. Xantheas, *The Journal of Chemical Physics* **139**, 114302 (2013).
- [214] R. M. Shields, B. Temelso, K. A. Archer, T. E. Morrell and G. C. Shields, *The Journal of Physical Chemistry A* **114**, 11725 (2010).

- [215] G. S. Fanourgakis, E. Apra and S. S. Xantheas, *The Journal of Chemical Physics* **121**, 2655 (2004).
- [216] J. P. Furtado, A. P. Rahalkar, S. Shanker, P. Bandyopadhyay and S. R. Gadre, *The journal of physical chemistry letters* **3**, 2253 (2012).
- [217] P. Parkkinen, S. Riikonen and L. Halonen, *The Journal of Physical Chemistry A* **117**, 9985 (2013).
- [218] P. D. Mezei, A. Ruzsinszky and G. I. Csonka, *Journal of chemical theory and computation* **12**, 4222 (2016).
- [219] H. Kabrede and R. Hentschke, *The Journal of Physical Chemistry B* **107**, 3914 (2003).
- [220] H. Takeuchi, *Journal of chemical information and modeling* **48**, 2226 (2008).
- [221] M. V. Kirov, G. S. Fanourgakis and S. S. Xantheas, *Chemical Physics Letters* **461**, 180 (2008).
- [222] S. Yoo, M. V. Kirov and S. S. Xantheas, *Journal of the American Chemical Society* **131**, 7564 (2009).
- [223] S. Y. Willow and S. S. Xantheas, *Chemical Physics Letters* **525**, 13 (2012).
- [224] K. Ramya and A. Venkatnathan, *The Journal of Physical Chemistry A* **116**, 7742 (2012).
- [225] E. D. Sloan Jr and C. Koh, *Clathrate hydrates of natural gases* (CRC press, 2007).

## Elements de valorisation de la thèse

### Liste des formations scientifiques ou transversales suivies :

- Formation à l'enseignement dans le cadre du contrat doctoral - Novembre 2015 (2 jours)
- Initiation à la programmation en fortran - Mai 2015 (2 jours)
- Participation à l'école d'été autour des méthodes de calcul de structure électronique (école d'été Roscoff) Juillet-Aout 2015 (1 mois)

### Articles scientifiques :

- Cairedine Kalai, Emilie-Laure Zins, and Mohammad Esmal Alikhani " A theoretical investigation of water-solute interactions : From facial parallel to guest-host structures" *Theoretical Chemistry Accounts* (2017)
- Cairedine Kalai and Julien Toulouse "A general range-separated double-hybrid density-functional theory" *Journal of Chemical Physics* (2018)

### Conférences :

- Conférence nationale : "15<sup>ème</sup> édition de la rencontre des chimistes théoriciens francophones" (communication par affiche) Lyon (France) 27 juin - 1er juillet
- Conférence internationale : "European Symposium on Chemical Bonding" (ESCB1) (communication orale) 29 août - 2 septembre 2016
- Conférence internationale : "7th International Conference on Density-Functional Theory and its Applications" (communication par affiche) Tällberg (Dalarna) Suède 21 août - 25 août 2017
- Conférence internationale : "43rd Congress of Theoretical Chemists of Latin Expression (CHITEL)" (communication orale) Paris (France), July 3 - 7, (2017)

### Chargé de mission d'enseignement :

- chimie : Structure et réactivité (1C001)
- chimie des solutions et structure des cristaux (1C002)
- spectroscopies : TP d'infrarouge (MTX3)
- Initiation à la programmation en python (3C007)

## Résumé/Abstract

**Description topologique des phénomènes d'hydratation et développement méthodologique de fonctionnelles doubles hybrides à séparation de portée.**

Résumé : cette thèse s'intéresse aux phénomènes d'hydratation de composés organiques à l'échelle moléculaire. Il est admis que l'équation de Schrödinger dans le cadre de l'approximation de Born-Oppenheimer et dans un contexte non relativiste contient toute la physique nécessaire pour décrire entre autres les phénomènes de solvatation de composés organiques. Des méthodes basées sur une fonction d'onde multi-déterminante sont capables de rendre compte des phénomènes de  $\mu$ -hydratation avec une précision approchant la réalité expérimentale. Or, ces méthodes sont limitées par la taille du système. L'utilisation de la DFT semble indispensable à une étude de complexes, même pour un nombre limité de molécules d'eau. Il s'avère que ces méthodes ne prennent pas en compte les interactions de nature dispersive. Des corrections empiriques ont été proposées récemment pour palier à ce problème. Cependant, ces corrections ne s'appliquent qu'à l'énergie et sur la géométrie des complexes hydratés, la fonction d'onde n'étant pas affectée par la correction. D'autres alternatives pour la prise en compte des effets de dispersion reposent sur l'emploi de méthodes hybrides fonction d'onde/DFT. Ceci peut s'effectuer en introduisant une séparation de portée dans le traitement des interactions électroniques. L'un des objectifs de cette thèse consiste à proposer une nouvelle méthode double hybride à séparation de portée permettant une bonne description des phénomènes d'hydratation. L'autre objectif de cette thèse consiste à utiliser des outils topologiques permettant la prédiction de composés organiques hydratés par l'étude du potentiel électrostatique moléculaire et la caractérisation de ces interactions non covalentes par la théorie "*Atoms in molecules*".

**Mots clés : AIM, MEP, double hybride à séparation de portée, micro-hydratation.**

**Topological description of hydration phenomena and development of range-separated double-hybrid methods.**

Abstract : this thesis deals with hydration phenomena of organic compounds at the molecular scale. The Schrödinger equation considered within the Born-Oppenheimer approximation and within a non-relativistic context contains all the physics necessary to describe in particular the micro-solvation of organic compounds. Methods that are based on a multi-determinant wave function are able to account for micro-hydration phenomena with a precision approaching the experimental reality. These methods are limited by the size of the system. The use of DFT seems necessary for a study of complexes, even for a limited number of water molecules. It turns out that these methods do not take into account dispersive interactions. Empirical corrections have recently been proposed to address this problem. However, these corrections apply only to the energy and to the geometry of the hydrated complexes, the wave function not being affected by the correction. Other alternatives for taking into account dispersion effects using double-hybrid methods should thus be considered. This can be done by introducing a range separation on the electronic interactions. There are two main objectives in this thesis. The first one is to propose a new double-hybrid method with range separation allowing a satisfactory description of the hydration phenomena at the molecular scale. The second objective consists in using topological tools allowing the prediction of hydrated organic compounds using the electrostatic molecular potential and the characterization of these non-covalent interactions by the "*Atoms in molecules*" theory.

**Keywords : AIM, MEP, range-separated double-hybrid methods, micro-hydration.**