



THESES

préparée et présentée à

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

pour obtenir le grade de

Docteur

Spécialité : CHIMIE PHYSIQUE

par

Patricia Guevara Level

Sujet de la thèse :

Title in English
Titre en français

Directeur de Thèse: **Pr Didier BÉGUÉ**

Soutenue le XXth, 2016

devant la commission d'examen composée de :

JURY

Dr Christophe RAYNAUD	Université de Montpellier, France	Rapporteur
Dr Johannes GIERSCHNER	IMDEA, Spain	Rapporteur
Dr	A	Examinateur
Dr Brigitte PEPIN-DONAT	CNRS, France	Président

Table of Contents

1 Asphaltènes	7
1.1 Définition	10
1.2 Composition élémentaire	10
1.2.1 Composition en hétéroatomes	10
1.2.1.1 Soufre	10
1.2.1.2 Azote	11
1.2.1.3 Oxygène	12
1.3 Poids moléculaire des asphaltènes	12
1.4 Caractérisation de la structure moléculaire des asphaltènes	12
1.4.1 Analyses par spectroscopie RMN	14
1.4.2 Analyses par diffraction des rayons X	15
1.4.3 Analyses par spectroscopie de Fluorescence Résolue en Temps	15
1.4.4 Analyses par spectroscopie infrarouge	16
1.5 Coexistence des structures moléculaires	16
1.6 Construction de la structure moléculaire des asphaltènes	17
1.7 Modélisation moléculaire des asphaltènes	17
1.8 Problèmes liés aux asphaltènes dans l'industrie pétrolière	18
2 Rappels théoriques : DFT	21
2.1 Les fondements de la DFT	24
2.1.1 Modèle de THOMAS-FERMI	24
2.1.2 Modèle de THOMAS-FERMI-DIRAC	24
2.1.3 Modèle de SLATER	25
2.2 Les théorèmes de HOHENBERG et KOHN	25
2.2.1 Premier théorème : preuve d'existence	26
2.2.2 Second théorème : théorème variationnel	26
2.3 Approche Kohn-Sham	27
2.4 Les différentes classes de fonctionnelle	28
2.4.1 Local Density approximation (LDA)	30
2.4.2 Generalized Gradient Approximation (GGA)	30
2.4.3 Fonctionnelles hybrides	31
2.4.4 Nécessité d'une fonctionnelle « longue portée »	31
2.4.4.1 Mise en évidence des forces de Van der Waals	32
2.4.4.2 Natures des interactions à modéliser	33
2.5 LC-DFT-D hybride : ω BXD	38
2.5.1 B88	38
2.5.2 B97	39

2.5.3	ω B97	39
2.5.4	ω B97X	42
2.5.5	ω B97X-D	42
3	Partie vibrationnelle	45
3.1	Généralités	48
3.2	Séparation des mouvements	50
3.3	Energie cinétique de vibration	52
3.4	Energie potentielle harmonique	53
3.5	Équations de LAGRANGE	54
3.5.1	Espace des coordonnées cartésiennes pondérées par les masses	54
3.5.2	Espace des coordonnées internes : résolution par la méthode de WILSON	55
3.5.3	Espace des coordonnées internes de symétrie	56
3.6	Traitements quantiques de la vibration	56
3.7	La fonction potentielle anharmonique	58
3.8	Représentation matricielle du Hamiltonien	62
3.9	Méthodes Perturbационnelles	64
3.10	Méthodes Variationnelles	66
3.10.0.1	Les méthodes de variation-perturbation	68
4	Corrections à la dispersion	73
4.1	DFT-D2	75
4.1.1	Theorie	75
4.2	DFT-D3	77
4.2.1	Les coefficients de la dispersion	78
4.2.2	Le terme de Trois Corps	79
4.3	La méthode Tkatchenko-Scheffler (DFT-TS)	79
4.3.1	Self-consistent screening dans la méthode Tkatchenko-Scheffler	81
5	Modélisation des interactions intermoléculaires	85
5.1	Bases Théoriques	87
5.2	Interactions intermoléculaires en théorie des perturbations	90
5.3	La Théorie de Perturbations à Symétrie Adaptée (SAPT)	92
5.3.1	L'interaction Electrostatique	94
5.3.2	L'Induction	94
5.3.3	Dispersion	95
5.3.4	Échange	96
5.4	DFT-SAPT	97
5.4.1	L'Energie Électrostatique et d'Échange	97
5.4.2	L'energie d'Induction et d'Échange- Induction	97
5.4.3	L'Energie de Dispersion et d'Échange- Dispersion	97
5.5	Méthodes Basées sur la DFT	98
5.6	CCSD(T)	98
5.7	Dimère du Benzène	99

6	Molecular dynamics study of the nano-aggregation in asphaltene	101
6.1	Introduction	103
6.2	Simulation Method	107
6.3	Results and Discussion	109
6.3.1	C5Pe-derivatives systems	109
6.3.2	Schuler's-derivatives systems	124
6.3.2.1	Effect of the heteroatom	124
6.3.2.2	Effect of chain-end	127
6.3.2.3	Aggregation within heterogeneous boxes	129
6.4	Conclusions	131
7	4	135
7.1	Calculation of vibrational spectra of Tetracene and Pentacene Monomers: Comparison with experimental data	137
7.1.1	Experimental data	137
7.1.2	Computational details: construction of the potential energy surface	137
7.1.3	Geometrical parameters of isolated tetracene molecule optimized using B3LYP and ω B97X-D methods associated with 6-311G and 6-311G** basis set	139
7.1.4	Truncation of the potential energy surface	140
7.1.5	Resolution of the vibrational Schrödinger equation (non-core bands with intensity)	142
7.2	Calculation of vibrational spectra of acene dimers: Comparison with experimental data.	146
7.2.1	Calculated geometry and wavenumbers of dimers and comparison with experimental data	146
7.2.2	Pi-stacking in the acene family	149
7.2.3	The impact of molecule distortion on the vibration modes of acene dimers	151
7.2.4	Identification of unassigned peaks in the experimental spectra of tetracene and pentacene	152
7.3	Normal Vibrational Modes in Crystals	155
7.3.1	Basic theory	155
7.3.2	Results and discussion	155
Appendices		165

List of Figures

1.1	Structure moyenne du modèle continental des asphaltènes	13
1.2	Molecule d'asphaltènes type archipel	14
2.1	Interaction électrostatique entre charges.	33
2.2	Interaction entre dipôles électrostatiques.	35
2.3	Interaction entre une molécule polaire et une seconde polarisable.	35
2.4	Interaction entre deux dipôles instantanés.	36
2.5	Interaction entre une espèce chargée et une molécule polaire.	36
2.6	π -stacking dans le cas d'un dimère de benzène	37
4.1	Les paramètres empiriques C_6 et R_0 pour DFT-D2	76
5.1	Les excitations contribuent à l'énergie de perturbation intermoléculaire en 2^{eme} ordre	91
5.2	Structures du dimère de Benzène	99
6.1	C5Pe molecular models of asphaltenes.	104
6.2	Two different asphaltene structures found in the crude petroleum: (a) PA3 and (b) CA22. R_1 and R_2 are the n -hexyl lateral chains which can have $-COOH$ or $-CH_3$ as chain-ends. X_n ($n=1-3$) can be O, S, or NH atom groups.	105
6.3	Snapshots of the initial configuration and the final frame of the dimers simulations. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.	110
6.4	Snapshots of the initial configuration and the final frame of the dimers simulations for the random disposition case. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.	111
6.5	Snapshots of the initial configuration and the final frame of the trimers simulations. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.	111
6.6	Snapshots of the initial configuration and the final frame of the trimers simulations for the random disposition case. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.	112
6.7	Distances between monomers for “organized” starting point.	113
6.8	Distances between monomers for “random” starting point.	113
6.9	Cosine of angle between aromatic cores in function of distance between dimers monomers for the organized starting point. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.	115

6.10	Cosine of angle between aromatic cores in function of distance between dimers monomers for the random starting point. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.	116
6.11	Distances between monomers for “organized” starting point.	117
6.12	Angle between monomers for “organized” starting point.	120
6.13	Distances between monomers for “random” starting point.	121
6.14	Angle between monomers for “random” starting point.	121
6.15	Normalized Radial Distribution Functions for the trimer systems.	122
6.16	Normalized Radial Distribution Functions for the S–X pairwise interactions, where X can be the oxygen atom of the aromatic core, the oxygen atom of the carboxylic group (COOH) or its hydrogen for the dimer system.	122
6.17	Normalized Radial Distribution Functions for the S–X pairwise interactions, where X can be the oxygen atom of the aromatic core, the oxygen atom of the carboxylic group (COOH) or its hydrogen for the trimer system.	123
6.18	Aggregation mechanism with $-\text{CH}_3$ ended chains: (a) effect of the heteroatom (A13, A23, A33, A43, A53) and (b) effect of the asphaltene type (A53/AAC). Each peak of the RDFs corresponds to a neighbor molecule in the π -stacking structure presented in Figure 6.3.	124
6.19	60 ns-snapshots of 5 molecules of each asphaltene in toluene at 298K and 1 bar conditions: (a) A33 and (b) AAC. Carbon atoms are presented in gray and nitrogen atoms in blue. Aromatic or polar hydrogen atoms are presented in white. Toluene molecules are omitted for clarity.	125
6.20	Aggregation mechanism with $-\text{COOH}$ ended chains: (a) effect of the heteroatom (A11, A21, A31, A41, A51) and (b) effect of the asphaltene type (A51/AAA). Each peak of the RDFs corresponds to a neighbor molecule in the π -stacking structure presented in Figure 6.19.	127
6.21	60 ns-snapshots of 5 molecules of each asphaltene in toluene at 298K and 1 bar conditions: (a) A31 and (b) AAA.	128
6.22	Effect of the mixture of A13 and A23 asphaltene molecules with $-\text{CH}_3$ lateral chain-end. This mixture concerns different ratios between Oxygen and Sulfur in the nano-aggregate.	130
7.1	Representation of a tetracene molecule.	140
7.2	Far-infrared PA spectra of (a) tetracene; (b) pentacene	143
7.3	The deformed dihedral angle C1-C2-C3-C4 in Tetracene molecule	144
7.4	Comparison of theoretical and experimental wavenumbers of tetracene and pentacene	146
7.5	Optimized geometries for naphthalene dimers	148
7.6	Acene dimers in the parallel displaced configuration	151
7.7	Comparison of theoretical intramolecular vibration modes for acenes dimers	152
7.8	Comparison of theoretical intermolecular vibration modes for acenes dimers	153
7.9	Final assignment of the experimental infrared transitions for tetracene and pentacene	154
7.10	Geometrical data for a tetracene molecule	161
7.11	Effect of mode truncature on the first ten mode on wavenumber of tetracene	162

List of Tables

5.1	Energie d'interaction du dimère de Benzène avec CCSD(T) en kcal/mol et DFT-SAPT (PBE0)	100
6.1	Labelling of the systems used in this study, with the chain-ends of the lateral chains. Molecules A11-A53 are of the type PA3 and AAA, AAB and AAC are of the type CA22. NbC is the number of Carbon atoms of the structure, DBE is the Double Bond Equivalent and Class stands for the nomenclature reporting the number of heteroatoms found in the structure by FT-ICR-MS. M_w is the molecular weight of each molecule.	106
6.2	Interaction energies ($< E_i >$) in kcal.mol ⁻¹ calculated for PA3- and CA22-type dimer molecules.	126
7.1	Method and Basis set dependence of C-C bond length (Å) a for tetracene monomer, and mean deviation from experimental data ^b	140
7.2	Calculated vibrational frequencies (cm ⁻¹) of the monomer, dimer and solid-state (PBE tetracene system)	159
7.3	Computed wavenumbers and intensities for distorted tetracene based on the harmonic approximation at the ω B97X-D/6-311G level of calculation. Comparison with symmetric tetracene (D_{2h})	162
7.4	Calculated and Experimental wavenumbers (cm-1)[124] of the vibrational modes of a naphthalene molecule. Values in parenthesis represent the deviation between the calculated and experimental wavenumbers.	163
7.5	Intramolecular vibration modes (cm ⁻¹) and the corresponding intensities in brackets (km/mol) for distorted naphthalene, anthracene and tetracene dimers. Values of corresponding dimer modes of molecules in D_{2h} configuration are written below in italics.	164

Chapter 1

Asphaltènes

Introduction

Les enjeux de la compréhension des propriétés physiques et chimiques d'un système reposent sur l'arrangement des atomes qui constituent sa structure. Dans cette optique, chimistes et physiciens ont développé au cours des années diverses méthodes visant à caractériser les systèmes chimiques, en tenant compte de leur environnement. Ces avancées méthodologiques restent toutefois limitées par les moyens technologiques, et la modélisation des systèmes de grande dimension représente un défi de taille pour les chercheurs. Entrant dans cette catégorie, les asphaltènes demeurent des constituants mal connus du pétrole brut et leur nature exacte comme leur taille restent, aujourd'hui encore, autant de freins à la caractérisation précise, en termes de composition chimique, des huiles brutes. L'étude et la compréhension approfondie de ces systèmes représente pourtant un enjeu scientifique et économique majeur, la présence de ces composés dans toutes les étapes, de la production au raffinage du pétrole, impliquant à l'heure actuelle l'emploi de procédés spécifiques particulièrement onéreux [1]. Plus précisément, les asphaltènes, présents sous forme de clusters ou de nanoparticules au sein des huiles brutes, tendent à augmenter la viscosité du pétrole brut et nécessitent des extractions chimiques successives, certaines étapes requérant de surcroît un traitement à haute température.

Si les informations expérimentales (obtenues par spectrométrie de masse, pyrolyse, etc.) recueillies au fil des années sur les asphaltènes n'ont pas permis, à elles seules, d'élucider leur composition chimique ou leur arrangement, éléments clés qui ouvriraient la voie à une caractérisation approfondie, des études conjointes expérience-théorie ont révélé des éléments intéressants [2]. Notamment, l'étude des huiles lourdes en tant que systèmes biphasiques complexes requiert d'appréhender en premier lieu les phénomènes d'agrégation conduisant à la formation de micelles d'asphaltènes. L'organisation de ces micelles permet ensuite de mettre en lumière l'effet de l'addition de pentane, employé comme solvant dans l'extraction des asphaltènes.

En définitive, bien que la modélisation des asphaltènes sur la base de ces éléments ouvre des pistes à la compréhension des processus d'agrégation, de flocculation et de précipitation de ces systèmes, il est devenu indispensable d'augmenter le niveau de précision des représentations chimiques des asphaltènes pour espérer comprendre leurs propriétés et prédire, sinon prévenir,

les conséquences des phénomènes d'agrégation. Dans cette optique, il convient de revenir à la composition moléculaire de base des asphaltènes. Les asphaltènes sont constitués d'un mélange hétérogène de structures polycondensées et d'hétéroatomes (azote, oxygène, soufre). La présence de ces hétéroéléments induisant la formation de nombreuses interactions, notamment de type van der Waals, ces atomes constituent la clé de voûte de l'assemblage moléculaire [3, 4]. Elucider les interactions intermoléculaires impliquant ces hétéroatomes est le premier enjeu dans la représentation approfondie de la structure des asphaltènes.

En reprenant l'ensemble des études réalisées jusqu'à ce jour sur la structure des asphaltènes, ce chapitre se veut un état des connaissances actuelles de ces systèmes. En pointant du doigt les zones d'ombre qui persistent dans la représentation de ces molécules, nous poserons le cadre de travail dans lequel s'inscrit la présente étude.

1.1 Définition

Historiquement, le premier emploi du terme "asphaltène" est attribué au français Boussingault, qui l'utilise dès 1837 pour caractériser certains des produits de distillation de l'asphalte [5]. Les "asphaltènes" sont alors les solides insolubles dans l'alcool mais solubles dans l'essence de térébenthine, par opposition aux "pétrolènes", constituants volatils solubles dans l'éther.

A l'heure actuelle, le terme "asphaltène" désigne la fraction lourde du pétrole, insoluble dans les n-alcanes (tels que le n-pentane ou le n-heptane) mais soluble dans les solvants aromatiques tels que le toluène, la pyridine ou le benzène. D'un point de vue structural, les asphaltènes sont constitués de noyaux aromatiques condensés, substitués par des groupements aliphatiques et naphténiques et faisant intervenir des hétéroatomes (azote, soufre et oxygène) au sein d'arrangements de type hétérocycles [6]. La proportion globale d'hétéroatomes, comme la fraction précise d'atomes d'azote, de soufre ou d'oxygène, peut être quelque peu variable suivant les procédés d'extraction employés pour isoler les asphaltènes de l'huile brute [7].

1.2 Composition élémentaire

Les méthodes d'analyse principales appliquées pour déterminer la composition élémentaire d'une fraction lourde du pétrole ont été décrites par Barbelet *et al.*[8]. La composition élémentaire des asphaltènes en atomes de carbone et d'hydrogène semble varier faiblement suivant leur origine, et a été évaluée respectivement à $82\% \pm 3\%$ et $8.1\% \pm 0.7\%$ par Speight et Moschopedis [9]. En outre, contrairement aux fractions plus légères du pétrole, au sein desquelles les hydrocarbures sont principalement de nature aliphatique (paraffine cyclique de type mono ou di-naphtène) ou monoaromatiques, les fractions lourdes incluent des structures naphténiques et aromatiques comportant plus de six cycles alkyliques.

1.2.1 Composition en hétéroatomes

1.2.1.1 Soufre

Les dérivés soufrés contenus dans les asphaltènes sont semblables aux espèces que l'on retrouve dans les fractions plus légères du pétrole, mais leur proportion varie sensiblement. En règle

générale, ces dérivés se divisent en cinq grandes classes chimiques : thiols, sulfures, disulfides, sulfoxydes et thiophènes. Thiols, sulfures, sulfoxydes et disulfides peuvent être classés plus précisément suivant leur nature cyclique ou acyclique, soit suivant l'organisation du squelette carboné (alkyle, aryle ou alkyl-aryle). Les dérivés thiophéniques, quant à eux, sont des structures polyaromatiques condensées autour d'un noyau thiophénique (benzo, dibenzo, naphtobenzo-thiophènes, etc.). Au sein des fractions lourdes, la majorité des espèces soufrées sont des dérivés thiophéniques, suivi par des dérivés sulfures (cycliques et acycliques). Des espèces de type sulfoxydes ont également pu être mises en évidence, dans des proportions très variables, leur teneur variant entre 0.3% et 10.3% [10, 11].

1.2.1.2 Azote

Dans les fractions brutes, l'azote est présent en quantité inférieure aux autres hétéroéléments. On distingue néanmoins deux types de composés en fonction de la teneur en azote de l'asphaltène considéré. C'est ainsi que l'on distingue les composés "bases", dont la proportion d'azote est inférieure à l'unité, des composés "neutres".

Les familles basiques qui ont pu être caractérisées présentent des structures dérivées de la quinoléine, contenant entre deux et quatre cycles aromatiques dont les configurations varient (péri ou catacondensées, avec différents degrés d'alkylation). Plus particulièrement, des benzo, dibenzo et tétrahydroquinoléines, ainsi que des azapyrènes ont été mis en évidence.

De même, des structures neutres présentant divers degrés d'alkylation ont été détectées : carbazoles, benzo-carbazole et dibenzo-carbazoles.

La proportion de structures neutres et basiques au sein d'une fraction brute est fortement liée aux paramètres géochimiques du site de provenance du pétrole. Enfin, il est à noter que des structures dérivées de la porphyrine ont été mises en évidence, même si de tels composés se retrouvent en proportions plus anecdotiques. Leur présence, caractérisée au travers du complexe qu'elles ont tendance à former avec les ions nickel ou vanadium, dépend une fois de plus de l'origine du pétrole dont est extraite la fraction lourde. Le pourcentage d'occurrence de ces structures varie communément entre 0.6 et 3.3% [10, 11].

1.2.1.3 Oxygène

Des structures oxygénées se retrouvent également dans les fractions lourdes du pétrole, leur teneur variant entre 0.3% et 4.9% [11]. Différentes familles ont pu être identifiées, notamment des dérivés du phénol, tandis que la caractérisation de groupements carboxyles révèle la présence d'ester, d'acides carboxyliques, de cétones, d'amides ou de sulfoxydes.

1.3 Poids moléculaire des asphaltènes

Un rapide passage en revue de la bibliographie des asphaltènes révèle tout le débat autour de leur poids moléculaire. En effet, la distribution, en termes de poids moléculaire, varie entre 400 et 1500 Da pour les fractions de faible poids moléculaire, et peut atteindre 10^6 Da pour les fractions de haut poids moléculaire [12]. Le phénomène d'agrégation des asphaltènes est une des raisons avancées pour justifier qu'à l'heure actuelle aucune méthode n'ait permis de déterminer avec précision le véritable poids moléculaire de ces espèces. De nombreux procédés ont été employés pour tenter de résoudre ce problème, mais ceux-ci conduisent à des résultats ambigüs, qu'il serait hasardeux de généraliser. A titre d'exemple, des mesures par osmémétrie à tension de vapeur font apparaître des résultats fortement dépendants du type de solvant employé ; en présence d'un solvant apolaire tel que le toluène, les résultats apparaissent surestimés, au contraire de l'utilisation d'un solvant polaire telle que la pyridine ou le nitrobenzène.

Merdrignac et col.[13] rapportent l'évolution du poids moléculaire des asphaltènes et comparer les changements causés par le procédé d'hydroconversion.

1.4 Caractérisation de la structure moléculaire des asphaltènes

La description de la structure des asphaltènes repose à l'heure actuelle sur deux modèles construits sur la base des résultats de diverses caractérisations expérimentales. Le premier modèle, baptisé "modèle continental", représente les asphaltènes comme de larges coeurs constitués de quatre à dix cycles aromatiques condensés, ramifiés par des chaînes alkyles courtes [14]. Ce modèle a été proposé par Zhao *et al.* [15] figure 1.1 après caractérisation des asphaltènes obtenus par fractionnement avec du pentane supercritique, puis remployé par Rogel et Carbognani

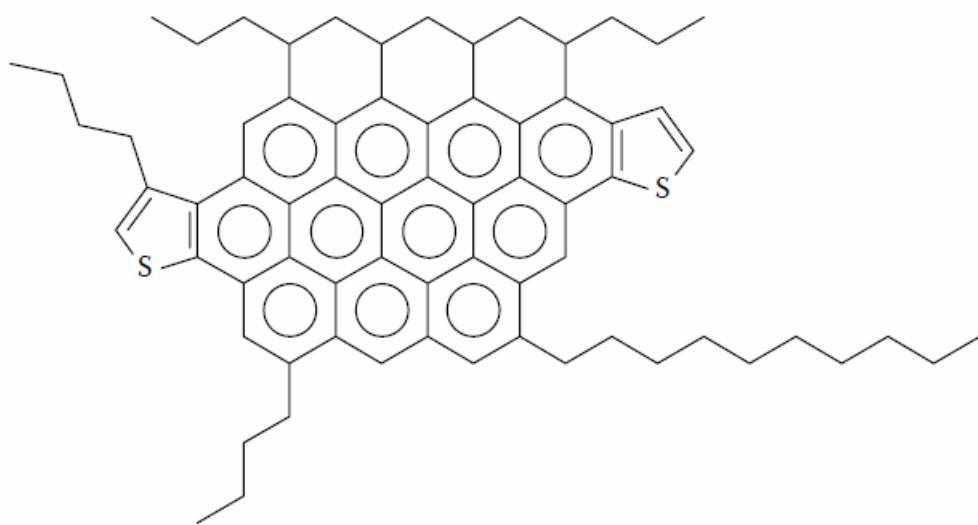


Figure 1.1: Structure moyenne du modèle continental des asphaltènes. Zhao et col. 2001

[16] dans leur travail sur des asphaltènes stables et instables extraits de pétrole provenant du Venezuela. Ce type de représentations est corrélé par des analyses en spectroscopie RMN, diffraction des rayons X et spectroscopie de fluorescence résolue en temps (TRDF).

Le modèle "archipel" présente quant à lui ces systèmes comme un ensemble de régions à caractère aromatique, constituées de deux à trois cycles condensés, reliées par des chaînes carbonées. Ce modèle s'appuie sur des analyses par pyrolyse, oxydation, dégradation thermique et dispersion angulaire neutronique, telles que rapportées par Gawrys *et al.* [17]. Sheremata *et al.* [18] ont notamment employé ce modèle pour la description des asphaltènes, comme le présente la figure reproduite ci-dessous (figure 1.2).

Ces deux modèles, basés sur des structures moléculaires bien distinctes, conduisent néanmoins à des propriétés physico-chimiques différentes. La description des agrégats d'asphaltènes ainsi que leur solubilité dans le pétrole brut seront nettement impactées par l'emploi de l'un ou l'autre de ces modèles. Du fait de la planéité induite par la présence d'un grand nombre de cycles aromatiques condensés, la représentation d'un agrégat d'asphaltènes dans le cadre du modèle continental conduit à un empilement de plans. A l'inverse, les chaînes alkyles du modèle archipel confère une toute autre géométrie à ces systèmes : les asphaltènes peuvent se courber par le biais d'interactions moléculaires pour former des macro-agrégats globulaires susceptibles de piéger les molécules de solvant.

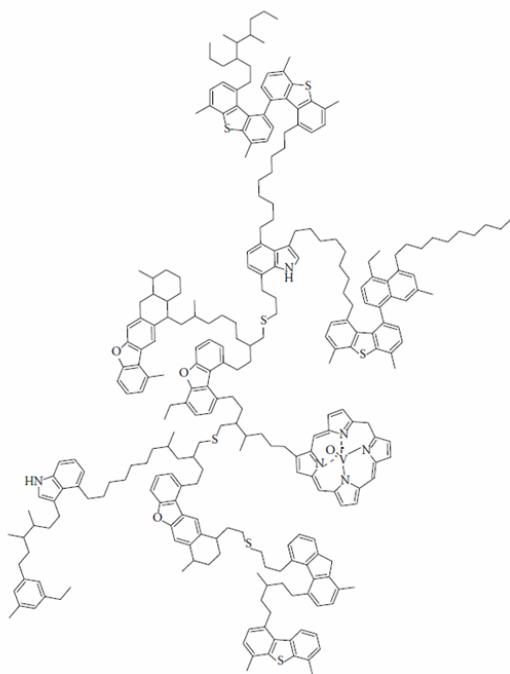


Figure 1.2: Représentation d’asphaltènes à l’aide du modèle archipel, proposé par Sheramata *et al.*

1.4.1 Analyses par spectroscopie RMN

La spectroscopie de résonance magnétique nucléaire est couramment employée pour l’analyse structurale et la caractérisation de matériaux et de substances organiques. Cette technique représente donc une méthode d’analyse de choix pour l’étude des asphaltènes et a, à ce titre, permis d’acquérir les premiers résultats visant à résoudre leur composition chimique. Dès 1982, les travaux de Murphy *et al.* [19] révèlent que la structure primaire des asphaltènes se compose de multiples cycles aromatiques condensés (l’indice de condensation relevé sur ces systèmes étant de 3 au moins) ainsi que d’une grande variété de chaînes aliphatiques, longues et ramifiées. Par la suite, les travaux de Yen *et al.* en RMN 1H ont donné une première idée de la proportion d’hydrocarbures saturés en regard de la fraction d’hydrocarbures aromatiques au sein de molécules d’asphaltènes [20]. Enfin, plus récemment, Durand *et al.* ont employé les résultats de leurs analyses par RMN ^{13}C pour évaluer les indices de substitution et de condensation d’échantillons d’asphaltènes de différentes origines [21]. Leur étude tend à démontrer qu’au sein d’agrégats d’asphaltènes cohabitent les deux modèles structuraux évoqués ci-dessus, à savoir le modèle continental et le modèle archipel.

1.4.2 Analyses par diffraction des rayons X

Une étude du caractère aromatique et des paramètres cristallins des asphaltènes, des résines et des gilsonites issus du pétrole a été présentée par Yen *et al.* sur la base d'analyses par diffraction des rayons X [22]. Leurs résultats, confortés par les travaux réalisés par Shirokoff *et al.* sur quatre échantillons d'asphaltènes issus d'huiles brutes provenant d'Arabie Saoudite [23], représentent les asphaltènes comme des feuillets de cycles aromatiques condensés portant des ramifications de natures naphténiques et aromatiques. Toutefois, Altget et Boduszynky rappellent judicieusement dans leur étude [24] qu'il convient de prendre avec précaution les résultats provenant d'analyses en diffraction, lorsque les données géométriques fournies par l'analyse servent de base pour construire un modèle structural du système étudié. Leurs travaux semblent en outre indiquer que la détermination du caractère aromatique des asphaltènes est, par ce biais, inappropriée sinon arbitraire.

1.4.3 Analyses par spectroscopie de Fluorescence Résolue en Temps

Cette méthode d'analyse a été employée à plusieurs reprises pour tenter d'élucider plus précisément la taille et la structure de molécules d'asphaltènes faiblement agrégées. Les premiers travaux en ce sens, réalisés par Groenzin et Mullins [25], tendent à montrer que les molécules d'asphaltènes ne seraient constituées que d'un seul groupe chromosphère. En effet, le faible poids moléculaire obtenu pour ces systèmes, comme la coloration des molécules, sont autant d'éléments recueillis incompatibles avec la représentation des asphaltènes sous forme de larges structures pseudo-polymériques constitués de deux ou trois cycles aromatiques reliés entre eux par de longues chaînes aliphatiques. En définitive, ces premières analyses semblent conforter la représentation des asphaltènes par le modèle continental.

Par la suite, Souza *et al.* ont rapporté l'agrégation persistante des asphaltènes à la faible concentration de 0.8 g/L, soit sous le seuil de concentration critique de nanoagrégation [26]. Leurs travaux corroborent par ailleurs les résultats de Groenzin *et al.* quant à la structure des molécules d'asphaltènes en cela qu'ils concluent à une structure primaire constituée par un anneau polyaromatique de quatre cycles ou plus.

1.4.4 Analyses par spectroscopie infrarouge

La caractérisation d'un système chimique en termes de composition passe nécessairement par des analyses en spectroscopie infrarouge, laquelle permet l'identification des groupements fonctionnels de ce système. Toutefois, l'étude des fractions lourdes du pétrole, la diversité chimique et la taille des espèces présentes rend complexe l'utilisation de la spectroscopie IR. C'est ce que soulignent Yuan *et al.*, qui relèvent l'existence d'un faible nombre d'analyses employant l'infrarouge moyen pour la caractérisation physique et chimique des huiles lourdes [27]. Parmi les travaux notables, l'étude couplée en spectroscopies IR et UV-visible menée par El-Bassoussi *et al.* sur deux échantillons d'asphaltènes provenant d'Egypte a permis de classifier les espèces présentes en mono, di et polyaromatiques. En accord avec les deux modèles de représentation des asphaltènes, les espèces prédominantes sont les espèces di ou polyaromatiques [28]. En 2007, Rodrigues Coelho *et al.* [29] démontrent l'existence d'une corrélation linéaire entre les intensités des bandes infrarouges symétriques et antisymétriques associées aux atomes d'hydrogène aromatiques de type arènes méthyl-substitués dans les régions $2900\text{-}3100\text{ }cm^{-1}$ et $700\text{-}900\text{ }cm^{-1}$. Enfin, Laxalde *et al.* [30] rapportent une interprétation ajustée sur les structures modèles des asphaltènes en repérant les vibrations de stretching des liaisons C-H (hors du plan) et C=C.

1.5 Coexistence des structures moléculaires

La compilation de l'ensemble de ces résultats expérimentaux tend à suggérer que les deux modèles structuraux des asphaltènes peuvent coexister au sein des fractions lourdes. Cette idée est appuyée par les résultats d'Acevedo *et al.* [31, 32] qui ont montré que les asphaltènes pouvaient se diviser en deux fractions : une fraction insoluble dans le toluène, baptisée A1, qui serait rigide et plane, conformément au modèle continental, et une fraction soluble dans le toluène grâce à de nombreuses interactions intermoléculaires, baptisée A2 et qui correspondrait au modèle archipel. Sur la base de ces éléments, Acevedo et son équipe ont mis en place une nouvelle représentation des agrégats d'asphaltènes. Le système colloïdal consisterait en une fraction d'asphaltènes A1, entourés par une couche d'asphaltènes A2, ces derniers aidant à la solubilisation de l'ensemble. En définitive, une revue bibliographique des analyses menées à ce jour sur les asphaltènes montre qu'aucun modèle structural définitif n'a pu être validé pour ce type de systèmes. De nombreux travaux cherchent encore à clarifier l'architecture molécu-

laire des asphaltènes, à deux échelles distinctes. En premier lieu, l'étude de la microstructure correspond à des systèmes de poids moléculaires variant entre 500 et 10 000 Da, soit à des nano-agrégats d'asphaltènes. En second lieu, les travaux sur la macrostructure de ces entités correspondent à l'étude d'assemblages de ces nano-agrégats d'asphaltènes et sont intimement liés au milieu.

1.6 Construction de la structure moléculaire des asphaltènes

La construction d'un modèle moléculaire utilisable en simulation se fait sur la bases des données expérimentales recueillies, suivant deux méthodes principales. La première méthode repose sur une corrélation des résultats empiriques utilisés pour déduire des informations structurales. Dans le cas des asphaltènes, l'existence de nombreux cycles aromatiques et l'ensemble des données obtenues concernant la longueur des chaînes aliphatiques ont permis à Takanohashi *et al.* d'étudier trois structures distinctes d'asphaltènes générées par la méthode de Sato par le biais de simulation en dynamique moléculaire [33]. La seconde méthode consiste à identifier des caractéristiques structurales propres au système ainsi que sa composition moléculaire en suivant un processus stochastique. Le modèle structural se construit par l'addition des données expérimentales et/ou par recouplement de ces résultats avec une base de données existantes. Cette méthode a été poursuivie aussi bien par Elyashberg *et al.* [34], qui ont développé un algorithme permettant la construction d'une structure sur la base de données RMN, que par Todeschini *et al.* [35] qui ont quant à eux employé des propriétés physiques telles que l'hydrophobie, le point de fusion et le point d'ébullition. La construction d'un modèle moléculaire des asphaltènes peut en dernier lieu se faire initialement de façon plus intuitive, les informations expérimentales servant ensuite à optimiser le modèle de départ [36, 37, 38].

1.7 Modélisation moléculaire des asphaltènes

La chimie computationnelle a révolutionné notre façon d'appréhender la structure et la réactivité des molécules, de sorte que la simulation est désormais une des clés de voûte de l'avancée scientifique. Comme nous l'avons souligné dans les paragraphes précédents, la nature et la complexité des asphaltènes rendent impossible une leur caractérisation exhaustive par le seul biais

de données expérimentales. Dans ce cadre, l'intérêt des simulations est incontestable et nombreux sont les travaux qui se sont déjà penchés sur la configuration moléculaire, les interactions intermoléculaires, la stabilité ou les phénomènes d'agrégation de ces systèmes. Concernant ce dernier point, la modélisation moléculaire a permis de justifier que l'agrégation des asphaltènes représentait la conformation la plus stable, sur la base d'une double approche structurale et thermodynamique révélant les interactions avec les résines, présentes au sein des fractions lourdes, et les solvant [39]. Selon ces travaux, l'interaction responsable de la formation comme de la stabilité des micelles formés d'asphaltènes et de résines serait la force d'attraction qui s'exerce entre les plans aromatiques. D'autres études théoriques se fondent sur les modèles structuraux proposés pour les asphaltènes. Les orbitales moléculaires de dimères d'asphaltènes, envisagés successivement par les modèles continental et archipel, ont été étudiés dans une approche semi-empirique ZINDO après optimisation des structures en DFT (théorie de la fonctionnelle de la densité). Du point de vue de la stabilité, la configuration continentale, représentées dans les simulations par des dimères empilés pour symboliser les plans d'asphaltènes propres à ce modèle, s'avère globalement plus stable que la configuration archipel [40].

1.8 Problèmes liés aux asphaltènes dans l'industrie pétrolière

Comme évoqué en introduction du présent chapitre, les problèmes techniques induits par la présence d'asphaltènes au sein des fractions lourdes du pétrole se posent à tous les stades de la production pétrolière, de l'extraction au raffinage. S'agissant de l'étape d'extraction, les asphaltènes, possédant des propriétés de sédimentation, viennent réduire la perméabilité de la roche et limiter le rendement d'extraction, lorsqu'ils ne provoquent pas le blocage de l'ensemble du pipeline et l'arrêt du procédé [41]. Par ailleurs, les asphaltènes ont tendance à précipiter lors du mélange de deux pétroles d'origine différente ; ce phénomène physique augmente la viscosité du fluide obtenu par mélange et le rend de fait difficile à manipuler et transporter. Dans les phases de raffinage, et plus spécifiquement au cours des procédés d'hydrotraitement, la forte teneur en hétéroéléments des asphaltènes est susceptible de perturber l'action des catalyseurs, sinon de les désactiver. De manière générale, les conditions physiques nécessaires aux différents stades de production pétrolière (hautes températures, pressions élevées) induisent des modifications chimiques structurales susceptibles de rendre le système instable. Dans ces

conditions, les asphaltènes floquent et précipitent ; leur taille augmente et ces systèmes ont alors tendance à adhérer à la surface des pipelines [42]. Enfin, les asphaltènes contribuent à la formation d'émulsions huile/eau stables qui affectent le procédé de séparation pétrole/eau dans la phase de récupération assistée du pétrole par l'eau [43]. L'ensemble des problèmes induits par la présence des asphaltènes ont un impact économique majeur sur le procédé. Ces systèmes obligent à requérir à un traitement spécifique et onéreux pour le traitement des fractions lourdes du pétrole brut.

Chapter 2

Rappels théoriques : Méthodes de la fonctionnelle de la densité (DFT)

Introduction

La chimie théorique est une science relativement récente, puisque ce n'est qu'en 1933 que le physicien autrichien Erwin Schrödinger a reçu le prix NOBEL de physique, en commun avec Paul DIRAC, pour ses travaux représentant, depuis, les fondements de la chimie quantique. En effet, l'équation de Schrödinger nous prouve que la connaissance de la fonction d'onde du système donne accès, explicitement ou non, à toutes les valeurs caractéristiques du système chimique étudié. Dans l'article *La situation actuelle en mécanique quantique* [44] paru en 1935, la métaphore du “chat de Schrödinger” a largement contribué à la vulgarisation de cette science auprès du public scientifique. En effet, il y symbolise, à l'aide d'un exemple macroscopique régit par les lois de la physique classique, la philosophie de la mécanique quantique dévolue à l'étude des systèmes microscopiques. Le plus grand défi à relever ici était de faire comprendre qu'une “mesure quantique” est une combinaison linéaire d'une somme d'états de probabilités non-nulles et non une valeur unique. L'expérience fictive consiste à placer un chat dans une boîte contenant un flacon de poison ainsi qu'une source radioactive. Lorsqu'un compteur Geiger détecte un certain seuil de radiation, un mécanisme vient briser le flacon et libère ainsi des vapeurs mortelles. Dans un raisonnement quantique, le chat est donc à la fois vivant et mort dans la boîte tant qu'elle reste close, avec une probabilité de vie de plus en plus faible au cours du temps.

En pratique, le nombre considérable de calculs à réaliser dans le cadre de la chimie théorique lie intrinsèquement cette science au développement de l'informatique, tout aussi récent. Il s'agit même de l'une des plus grandes limites encore rencontrée de nos jours. Même si la loi de MOORE [45]¹, empirique de son état, a commencé à perdre en véracité dès lors que la fréquence des processeurs (CPU) engendrait une déperdition de chaleur trop importante et commercialement non-maîtrisable ($\gtrsim 5$ Ghz), l'apparition des processeurs multicoeurs s'est vu être la solution la plus efficace. Après l'adaptation nécessaire des codes séquentiels en version parallèle, la démocratisation relative des clusters de calcul² permet à l'heure actuel

¹La loi de MOORE (1965), renommée première loi de MOORE compte-tenu de l'ajustement ultérieur, énonce que la complexité des semi-conducteurs (et donc de la puissance de calcul) suit une loi exponentielle au cours du temps.

²Un cluster de calcul, ou grappe de serveurs, est un ensemble de serveurs esclaves, les nœuds, contrôlés par un ou plusieurs serveurs maîtres, le(s) frontal(aux). Ce groupe de serveurs indépendants, fonctionnant comme un seul et même système, permet d'optimiser les ressources (processeur, mémoire vive, stockage...) et une

le développement de codes hautement parallélisés (> 1000 processeurs). Ceci permet alors de faire moins d'approximations théoriques et donc de tendre vers des valeurs de plus en plus exactes, tout en gardant des temps de calcul raisonnables. Un engouement pour les clusters de cartes graphiques (GPU) est aussi à noter dans les domaines scientifiques où, tel que la dynamique moléculaire, les calculs sont extrêmement fragmentables et peu interdépendants.

Dans ce chapitre, nous rappellerons dans un premier temps les fondements de la théorie de la fonctionnelle de la densité, notée DFT, par le biais de l'évolution des différents modèles qui ont été proposés. Nous verrons ensuite comment les théorèmes de Hohenberg et Kohn prouvent que la seule connaissance de la densité électronique permet de résoudre l'équation de Schrödinger dans le cadre de la DFT. La fonctionnelle universelle $F_{HK}[\rho]$, qui permettrait une résolution exacte du problème, restant inconnue, nous aborderons dans une troisième partie l'approche KS qui contourne ce problème et légitime certaines approximations. Ces dernières donnant naissance à différents types de fonctionnelle, elles seront succinctement présentées dans la pénultième partie de ce chapitre. Finalement, le développement de la méthode de calcul utilisée dans ce travail de thèse, sera présentée étape par étape en fin de chapitre.

meilleure répartition des tâches sur les différents noeuds.

2.1 Les fondements de la DFT

Contrairement aux méthodes HARTREE-FOCK, noté HF, et *a fortiori* post-HF qui décrivent le système électronique par une fonction d'onde $\Psi_{(\vec{r})}$, la théorie de la fonctionnelle de la densité le décrit par la densité électronique, notée $\rho_{(\vec{r})}$, qui est liée à la fonction d'onde $\Psi_{(\vec{r})}$ par la relation suivante :

$$\begin{aligned}\rho_{(\vec{r})} &= \int \Psi_{(\vec{r})}^* \Psi_{(\vec{r})} \\ &= \int |\Psi_{(\vec{r})}|^2\end{aligned}\tag{2.1}$$

avec \vec{r} : ensemble des coordonnées électroniques.

L'énergie de l'état fondamental est ainsi une fonctionnelle de la densité électronique, c'est-à-dire que $E_0 = E_{(\rho)}$.

2.1.1 Modèle de Thomas-Fermi

Le terme d'énergie cinétique a été exprimé comme une fonctionnelle de la densité pour la première fois en 1927 par THOMAS et FERMI :

$$\hat{T}_{TF}[\rho] = \frac{3}{10}(3\pi)^{2/3} \int \rho_{(\vec{r})}^{5/3} d\vec{r}\tag{2.2}$$

Celle fonctionnelle est alors combinée aux expressions classiques des interactions électrons-noyaux et électrons-électrons, exprimées elles aussi en fonction de la densité électronique :

$$E_{TH}[\rho] = T_{TH}[\rho] + V_{Ne}[\rho] + V_{ee}[\rho]\tag{2.3}$$

2.1.2 Modèle de Thomas-Fermi-Dirac

Le terme d'échange, résultant du principe d'exclusion de PAULI, a ensuite été ajouté par Dirac en 1930 afin d'affiner le modèle :

$$\begin{aligned} K[\rho] = E_x[\rho] &= \int \rho_{\vec{r}} \epsilon_x[\rho].d\vec{r} \\ &= -\frac{3}{4} \left(\frac{3}{\pi}\right)^{1/3} \int \rho_{(\vec{r})}^{4/3}.d\vec{r} \end{aligned} \quad (2.4)$$

avec $\epsilon_X[\rho]$: énergie d'échange par électron.

Le modèle de THOMAS-FERMI-DIRAC est défini par la combinaison de cette expression avec l'équation 2.2 et le potentiel d'interaction électrons-noyaux $V_{Ne}[\rho]$. Notons que la corrélation électronique n'est toujours pas prise en compte dans ce modèle.

2.1.3 Modèle de Slater

Partant d'une approche basée sur la méthode HF, SLATER proposa en 1951 de substituer le terme d'énergie d'échange par une fonctionnelle de la densité issue de l'énergie d'échange de Dirac. Ce terme d'échange dans le formalisme HF peut alors être généralisé en introduisant le paramètre α :

$$E_x[\rho] = -\frac{9\alpha}{8} \left(\frac{3}{\pi}\right)^{1/3} \int \rho_{(\vec{r})}^{4/3}.d\vec{r} \quad (2.5)$$

Des analyses empiriques basées sur différents types de systèmes chimiques ont conduit à une valeur de $3/4$ pour α , offrant une meilleure précision que la valeur originelle de l'expression de Dirac ($2/3$).

2.2 Les théorèmes de Hohenberg et Kohn

Tous ces modèles, qui constituent les fondements de la DFT, ne démontrent pas formellement que seul la connaissance de la densité est importante pour atteindre la valeur de l'énergie totale d'un système. C'est ainsi que HOHENBERG et KOHN eurent l'idée en 1965 de démontrer, par le biais de deux théorèmes, que l'équation de Schrödinger pouvait être résolue de façon exacte, dans le cadre de l'approximation de BORN-OPPENHEIMER uniquement grâce à la densité électronique.

2.2.1 Premier théorème : preuve d'existence

Ce premier théorème énonce que l'ensemble des propriétés du système, notamment l'énergie, peuvent être calculées à partir de la seule densité électronique de l'état fondamental. Elles peuvent donc être décrites comme une fonctionnelle de la densité électronique, et l'énergie totale s'écrit alors :

$$E[\rho] = F_{HK}[\rho] + \int \rho(\vec{r}) \nu_{ext} d\vec{r} \quad (2.6)$$

où :

$$F_{HK}[\rho] = T_e[\rho] + V_{ee}[\rho] \quad (2.7)$$

$$\nu_{ext} = V_{Ne}[\rho]$$

Notons que la fonctionnelle universelle $F_{HK}[\rho]$, qui regroupe les termes d'énergie cinétique des électrons et celui d'énergie potentielle d'interaction électron-électron, n'est pas liée au potentiel externe ν_{ext} . L'énergie de l'état fondamental est *a priori* accessible de manière exacte car cette fonctionnelle ne repose sur aucune approximation.

2.2.2 Second théorème : théorème variationnel

Basé sur l'équation 2.6, HOHENBERG et KOHN ont ensuite construit un principe variationnel pour déterminer la densité électronique de l'état fondamental :

$$E[\rho] \geq E[\rho_0] \quad (2.8)$$

avec ρ_0 : densité électronique de l'état fondamental,

ρ : densité électronique quelconque.

Dans cette équation, à une densité d'essai ρ correspond une seule énergie potentielle $\int \rho(\vec{r}) \nu_{ext} d\vec{r}$ et une seule fonction d'onde Ψ_ρ . La méthode de double minimisation, *i.e.* sous contrainte de LEVY, permet de différencier la fonction d'onde Ψ_{ρ_0} , correspondant à l'état fondamental, parmi le jeu infini des fonctions d'ondes Ψ_ρ donnant la même densité. Ainsi, nous pouvons déterminer, parmi toutes les densités, celle qui minimisera l'énergie par la relation suivante :

$$E[\rho_0] = \min_{\rho} \left(\min_{\Psi \rightarrow \rho} (F[\rho] + \int \rho_{\vec{r}} \nu_{\vec{r}} . d\vec{r}) \right) \quad (2.9)$$

Si les théorèmes de HOHENBERG et KOHN démontrent une correspondance unique entre une densité $\rho_{\vec{r}}$ et la fonction d'onde Ψ du système, la fonctionnelle universelle $F_{HK}[\rho]$ reste cependant inconnue.

2.3 Approche Kohn-Sham

Afin de contourner ce problème, KOHN et SHAM substituèrent au Hamiltonien réel, décrivant un système de n particules en interaction, un Hamiltonien de référence décrivant un système de n particules sans interaction mais ayant la même densité que le système réel. Le problème est ainsi réduit à la résolution de n équations monoélectronique couplées, analogues aux équations de HF. L'opérateur monoélectronique de Kohn-Sham \hat{K}_{KS} s'exprime ainsi :

$$\hat{H}_{KS} = -\frac{1}{2} \nabla^2 + \nu_H[\rho] + \nu_{xc}[\rho] + \nu_{ext}[\rho] \quad (2.10)$$

où :

$$\nu_H[\rho] = \int \frac{\rho(\vec{r}) - \rho(\vec{r}')}{|\vec{r} - \vec{r}'|} . d\vec{r}' \quad (2.11)$$

$$\nu_{xc}[\rho] = \frac{\partial E_{xc}[\rho(\vec{r})]}{\partial \rho(\vec{r})} \quad (2.12)$$

$\nu_H[\rho]$ et $\nu_{xc}[\rho]$ étant respectivement le potentiel de HARTREE et le potentiel d'échange et de corrélation, dans lequel $E_{xc}[\rho(\vec{r})]$ est l'énergie d'échange et de corrélation.

En définissant un potentiel fictif $\nu_{eff}(\vec{r})$ pouvant être appliqué à des systèmes sans interaction de densité ρ :

$$\nu_{eff}(\vec{r}) = \nu_H[\rho] + \nu_{xc}[\rho] + \nu_{ext}[\rho] \quad (2.13)$$

nous introduisons alors un jeu d'orbitales $\psi_{(\vec{r})}$, appelées orbitales de KOHN-SHAM, et nous obtenons un jeu d'équations aux valeurs propres :

$$\hat{H}_{KS} \psi_{i(\vec{r})} = \epsilon_i \psi_{i(\vec{r})} \quad (2.14)$$

Comme dans le cas de la méthode HF, l'énergie du système peut être minimisée en résolvant ce jeu d'équations de façon auto-cohérente grâce à l'utilisation des orbitales de KOHN-SHAM. L'énergie du système est alors donnée par :

$$E_{KS}^{tot}[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho] + \int V_{ext}(\vec{r})\rho(\vec{r}).d\vec{r} \quad (2.15)$$

avec $T_s[\rho]$: énergie cinétique des électrons sans interaction,

$J[\rho]$: énergie d'interaction coulombienne entre les électrons,

$E_{xc}[\rho]$: énergie d'échange et de corrélation,

$\int V_{ext}(\vec{r})\rho(\vec{r}).d\vec{r}$: énergie d'interaction avec le potentiel externe.

D'après les théorèmes de HOHENBERG et KOHN, $E_{KS}^{tot}[\rho]$ doit être égale à l'énergie totale du système réel $E_{reel}^{tot}[\rho]$, qui peut être décrite comme suit :

$$E_{reel}^{tot}[\rho] = T[\rho] + V_{ee}[\rho] + \int V_{ext}(\vec{r})\rho(\vec{r}).d\vec{r} \quad (2.16)$$

Le terme d'échange corrélation peut ainsi être explicité comme étant la somme de la correction à l'énergie cinétique due à l'interaction entre électrons ($T[\rho] - T_s[\rho]$) et les corrections non classiques à la répulsion électron-électron ($V_{ee}[\rho] - J[\rho]$) :

$$E_{xc}[\rho] = T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho] \quad (2.17)$$

La théorie de la fonctionnelle de la densité de KOHN-SHAM a connu un grand succès parmi les méthodes de calcul appliquées aux grands systèmes dû à son ratio coût calculatoire/performance très intéressant. C'est pour cet avantage indéniable qu'elle sert de base à de nombreuses évolutions de la DFT.

2.4 Les différentes classes de fonctionnelle

Formellement, la DFT est donc une méthode exacte, dans la limite de la connaissance de la fonctionnelle universelle $F_{HK}[\rho]$ ou de sa fonctionnelle exacte d'échange et de corrélation $F_{xc}[\rho]$. Malheureusement, la forme exacte de l'énergie d'échange et de corrélation est inconnue, si bien qu'il est nécessaire de faire des approximations. Dans la pratique, l'énergie d'échange et de

corrélation $E_{xc}[\rho]$ est calculée à l'aide de fonctionnelles d'échange et de corrélation, définies comme suit :

$$\int F_{xc}[\rho(\vec{r})].d\vec{r} = E_{xc}[\rho(\vec{r})] \quad (2.18)$$

L'énergie d'échange et de corrélation est généralement séparée en deux termes distincts, l'un d'échange $E_x[\rho]$ et l'autre de corrélation $E_c[\rho]$:

$$E_{xc}[\rho(\vec{r})] = E_x[\rho(\vec{r})] + E_c[\rho(\vec{r})] \quad (2.19)$$

Plusieurs fonctionnelles ont donc été développées pour traiter chacune de ces contributions, de façon simultanée ou indépendante. Nous allons ici donner un bref aperçu – non exhaustif – des différentes familles de fonctionnelles, en suivant le critère de classification donné par PERDEW et couramment dit “échelle de Jacob de PERDEW”. Il a proposé de classer les fonctionnelles en fonction du degré d'information non local contenu dans leur forme analytique. Au premier échelon se trouvent les fonctionnelles qui dépendent uniquement de la densité électronique, dites fonctionnelles LDA (“Local Density Approximation”). Viennent ensuite les fonctionnelles corrigées par gradient GGA (“Generalized Gradient Approximation”) dans lesquelles la non-localité est introduite grâce à leur dépendance par rapport au gradient de la densité. A ce même niveau se trouvent également les fonctionnelles de type méta-GGA, dépendant aussi de l'énergie cinétique (calculée à partir des orbitales moléculaires remplies). Il faut noter que, d'un point de vue mathématique, toutes ces familles de fonctionnelles sont strictement locales. Pour parvenir à des fonctionnelles véritablement non locales il faut encore monter d'un cran dans l'échelle de PERDEW, jusqu'aux fonctionnelles dites hybrides, où la présence d'un pourcentage (variable) d'échange HF calculé en utilisant les orbitales KS, permet d'introduire un véritable terme non local. Plus récemment une autre famille de fonctionnelles hybrides a été développée (dite à longue portée ou encore à séparation de portée) dans laquelle le pourcentage d'échange calculé de façon HF n'est pas constant mais dépend de la distance inter-électronique. Enfin, des fonctionnelles non locales montrant une dépendance explicite des orbitales KS occupées et vacantes représentent le dernier niveau dans l'échelle de PERDEW.

2.4.1 Local Density approximation (LDA)

L'approximation locale de la densité est l'approximation la plus grossière, dans laquelle l'énergie d'échange et de corrélation E_{xc} n'est fonction que de la seule densité électronique :

$$E_{xc}^{LDA}[\rho(\vec{r})] = \int \rho(\vec{r}) \epsilon_{xc}[\rho(\vec{r})].d\vec{r} \quad (2.20)$$

où la valeur de ϵ_{xc} à une position \vec{r} est calculée exclusivement à partir de la valeur de la densité électronique ρ à cette position. En pratique, ϵ_{xc} décrit l'énergie d'échange et de corrélation par particule pour un gaz uniforme d'électrons de densité ρ . Le potentiel d'échange et de corrélation correspondant est alors :

$$\nu_{xc(\vec{r})}^{LDA} = \epsilon_{xc}[\rho(\vec{r})] + \rho(\vec{r}) \frac{\partial \epsilon_{xc}[\rho(\vec{r})]}{\partial \rho} \quad (2.21)$$

Malgré le fait que les résultats obtenus soient généralement en bon accord avec les résultats expérimentaux, notamment au niveau de la géométrie et de la structure électronique, cette approximation reste une approximation locale, c'est-à-dire dans laquelle on ne tient pas compte de l'inhomogénéité de la densité électronique.

2.4.2 Generalized Gradient Approximation (GGA)

L'idée directrice de l'approximation du gradient généralisé est donc de mieux tenir compte de l'inhomogénéité de la densité du système, en introduisant une dépendance de la densité ρ à son gradient $\nabla\rho$. L'expression générale des fonctionnelles de type GGA est la suivante :

$$E_{xc}^{GGA}[\rho(\vec{r})] = A_x \int \rho_{(\vec{r})}^{4/3} E^{GGA}(s).d\vec{r}^3 \quad (2.22)$$

où s , gradient de la densité réduite, est tel que :

$$s = \frac{|\nabla \rho(\vec{r})|}{2k_F \rho(\vec{r})} \quad (2.23)$$

avec $k_F = (3\pi^2 \rho(\vec{r}))^{1/3}$. Ainsi, on fait apparaître avec s un terme quasi-local, dépendant non seulement de la densité électronique mais également de son gradient au voisinage de \vec{r} .

Un exemple de fonctionnelle GGA est celle de PERDEW, BURKE et ERNZHERHOF, notée PBE [46].

2.4.3 Fonctionnelles hybrides

La dernière grande famille de fonctionnelles est celle des fonctionnelles hybrides. L'idée consiste à introduire une fraction d'échange calculée de façon exacte (telle qu'utilisée dans la méthode HF) dans une fonctionnelle d'échange de type GGA. L'expression de E_{xc} devient alors :

$$E_{xc}^{hybride}[\rho(\vec{r})] = (1 - \alpha)E_{xc}^{GGA}[\rho(\vec{r})] + \alpha E_{xc}^{HF}[\rho(\vec{r})] \quad (2.24)$$

où le coefficient de la combinaison α donne le rapport HF/DFT. PBE0 [47] Celle-ci présente 25% d'échange HF [48] dans une fonctionnelle GGA de type PBE [46] :

$$E_{xc}^{PBE0} = E_{xc}^{PBE} + \frac{1}{4}(E_x^{HF} - E_x^{PBE}) \quad (2.25)$$

Elle a l'avantage d'être non paramétrée (car le pourcentage d'HF inclus n'est pas empirique mais basé sur des arguments de théorie perturbationnelle), et de fournir des résultats très précis, que ce soit au niveau du calcul des structures moléculaires, des structures électroniques ou encore des propriétés spectroscopiques [47].

Nous mentionnerons également la fonctionnelle hybride la plus utilisée pour traiter des systèmes moléculaires, à savoir la fonctionnelle B3LYP [49]. Comme son nom l'indique, elle inclus trois paramètres et est basée sur les fonctionnelles GGA d'échange et de corrélation de BECKE (B) [50] et LEE, YANG et PARR (LYP) [51], suivant l'expression :

$$E_{xc}^{B3LYP} = (1 - a)E_x^{LSDA} + aE_x^{HF} + b\Delta E_x^B + (1 - c)E_c^{LSDA} + cE_c^{LYP} \quad (2.26)$$

avec a , b et c fixés respectivement à 0,20 , 0,72 et 0,81.

2.4.4 Nécessité d'une fonctionnelle « longue portée »

Bien que calculatoirement très pratiques, les méthodes basées sur la fonctionnelle de la densité connaissent des limites évidentes lorsqu'il s'agit de traduire correctement les interactions à longue distance. Ces interactions se révèlent pourtant être à l'origine de nombreuses propriétés chimiques et physiques des matériaux. En effet, elles sont par exemple responsables de la cohésion des cristaux liquides et moléculaires, des propriétés d'adhésion et de physisorption sur des surfaces, de la spécificité de site de l'ADN mais aussi de la cohésion dans les

systèmes lamellaires, et plus particulièrement dans le cas des matériaux graphitiques qui nous intéresse présentement. La compréhension et la bonne modélisation de ce type d’interactions est donc naturellement devenu un axe important dans le domaine de la physique et de la chimie quantique.

2.4.4.1 Mise en évidence des forces de Van der Waals

L’existence de phases naturellement condensées comme les liquides et les solides moléculaires a permis de démontrer de façon empirique la présence de forces attractives entre les molécules. Avec un raisonnement analogue, des forces répulsives à faibles distances ont été mises en avant grâce aux propriétés d’incompressibilité infinie de ces mêmes systèmes chimiques.

Johannes Diderik Van der Waals fut le premier en 1873 à intégrer cette notion de forces attractives et répulsives dans le modèle du “gaz parfait” (2.27) en introduisant comme suit deux termes correctifs dans sa loi des “gaz réels” (2.28) :

$$PV = nRT \quad (2.27)$$

$$\left(P + \frac{n^2 a}{V^2} \right) (V - nb) = nRT \quad (2.28)$$

avec P : pression (Pa),
 V : volume du gaz (m^3),
 n : quantité de matière de gaz (mol),
 R : constante des gaz parfaits,
 T : température (K),
 a et b : constantes ajustables positives.

Cette équation d’état est beaucoup plus proche des valeurs expérimentales, notamment aux fortes pressions. En effet, le volume d’exclusion nb permet de tenir compte du volume propre occupé par les molécules. Nous rappelerons que dans le modèle des gaz parfaits, dans lequel les molécules sont assimilées à des points ponctuels, ce volume est négligé. Notons que la pression est aussi corrigée par le terme $\frac{n^2 a}{V^2}$ afin de tenir compte de la diminution de pression au voisinage des parois, du fait de l’existence d’interactions attractives entre les molécules.

L’écriture de la loi des “gaz réels” par Van der Waals a réellement initié la volonté de

comprendre la nature de ces interactions et a naturellement conduit à de nombreuses études concernant les différents systèmes chimiques. Des efforts sont particulièrement portés sur la modélisation de forces de cohésions universelles, et non pas liées à une seule problématique.

2.4.4.2 Natures des interactions à modéliser

Avant de pouvoir établir une théorie, il faut avant tout comprendre la nature du phénomène. Les atomes et molécules étant composés de particules naturellement chargées, *i.e.* de noyaux et d'électrons, ils interagissent donc par des forces de COULOMB. Celles-ci se décomposent en plusieurs éléments selon leur nature : l'électrostatique, l'induction (ou polarisation), la dispersion et l'échange. L'effet du terme électrostatique représente la répulsion ou l'attraction entre des charges respectivement de même signe ou de signes opposés (figure 2.1). Son énergie $E_{Coulomb}$ a été définie par Charles-Augustin COULOMB en 1785 comme suit :

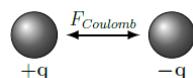


Figure 2.1: Interaction électrostatique entre charges.

$$E_{Coulomb} = \frac{1}{4\pi\epsilon_0} * \frac{q_1 q_2}{r} \quad (2.29)$$

avec q_1 et q_2 : charges respectives des particules 1 et 2,

r : distance entre les particules.

L'induction est quant à elle due à la déformation de la densité électronique d'un atome ou d'une molécule par l'effet du champ électrique d'une molécule voisine. Ces deux contributions sont très bien définies en physique classique, contrairement à la dispersion et l'échange, qui sont liés à des effets quantiques. En effet, ce sont les phénomènes de fluctuations quantiques de la distribution des charges qui sont à l'origine du terme de dispersion, et la contribution d'échange est dûe au principe d'exclusion de PAULI qui impose que deux électrons ne peuvent pas posséder le même état quantique au même moment. Dans le cas d'une paire d'atomes en interaction ayant leurs couches électroniques partiellement occupées, cette contribution est positive et engendre une liaison chimique liante forte. À l'inverse, lorsqu'il s'agit de systèmes électroniques à couche fermée, elle devient un terme de répulsion à courte distance responsable

du phénomène d'exclusion stérique. Ce phénomène étant inclu dans la loi des "gaz réels" de Van der Waals (équation 2.28) sous la forme du volume d'exclusion nb .

D'une manière plus générale, c'est-à-dire en dépassant le cadre des gaz, les interactions de Van der Waals sont générées par les fluctuations de distributions de charge des atomes et molécules et conduisent aux équations suivantes, où l'énergie est exprimée en Joules :

$$E_{Keesom} = -\frac{1}{r^6} \left(\frac{\mu_1^2 \mu_2^2}{3(4\pi\epsilon_0\epsilon)^2 k_B T} \right) \quad (2.30)$$

$$E_{Debye} = -\frac{1}{r^6} \left(\frac{\mu_1^2 \alpha_2 + \mu_2^2 \alpha_1}{(4\pi\epsilon_0\epsilon)^2} \right) \quad (2.31)$$

$$E_{London} = -\frac{1}{r^6} \left(\frac{3 h \nu \alpha_1 \alpha_2}{4 (4\pi\epsilon_0)^2} \right) \quad (2.32)$$

avec μ_1 et μ_2 : moments dipolaires respectifs des particules 1 et 2,
 α_1 et α_2 : polarisabilités respectives des particules 1 et 2,
 ϵ_0 : permittivité diélectrique du vide,
 ϵ : permittivité diélectrique du milieu,
 k_B : constante de BOLTZMANN, égale à la constante des gaz parfaits R divisée par le nombre d'AVOGADRO N_A ,
 T : température,
 r : distance entre les particules.

L'énergie de KEESEM (effet d'orientation) représente l'énergie d'interaction entre deux dipôles électrostatiques, c'est-à-dire deux molécules ayant un moment dipolaire³ μ non nul (figure 2.2). Il s'agit concrètement de l'attraction mutuelle de deux dipôles permanents qui est d'autant plus forte que les moments dipolaires sont élevés (grande charge et petite taille de la molécule) et que la température est basse.

Celle de DEBYE (effet d'induction) résulte de la déformation du nuage électronique d'une molécule, d'un atome ou d'un ion, par action du champ électrique engendré par le moment dipolaire d'une molécule voisine (figure 2.3). Il en résulte ainsi un moment dipolaire induit. Elle est souvent nommée interaction dipôle permanent-dipôle induit et fait intervenir le moment

³Le moment dipolaire d'une molécule résulte d'une répartition hétéroclite de charges électriques, telle que le barycentre des charges positives (noyaux) ne coïncide pas avec celui des charges négatives (électrons), les électrons étant en effet attirés par l'atome le plus électronégatif de la liaison.

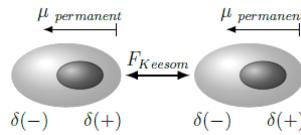


Figure 2.2: Interaction entre dipôles électrostatiques.

dipolaire μ et la polarisabilité⁴ α des molécules concernées. Le nominateur $\mu_1^2 \alpha_2 + \mu_2^2 \alpha_1$ décrit l’interaction lorsque les deux molécules sont polaires (cas A) mais s’écrit naturellement $\mu_1^2 \alpha_2$ quand la seconde molécule est apolaire (cas B).

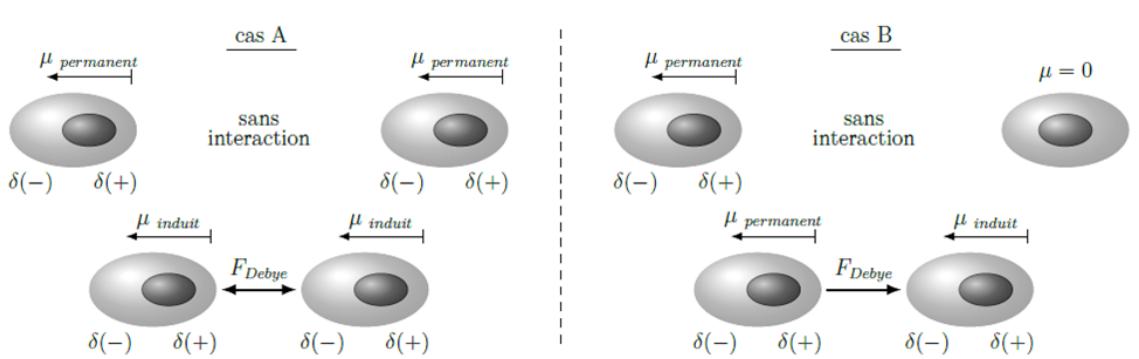


Figure 2.3: Interaction entre une molécule polaire et une seconde polarisable.

Finalement, l’énergie de LONDON (effet de dispersion), qui est la plus importante en terme de grandeur, représente l’interaction entre deux dipôles instantanés (figure 2.4). Par définition, une molécule apolaire possède un moment dipolaire moyen nul mais la combinaison des mouvements des noyaux et des électrons fait qu’il existe malgré tout un moment dipolaire instantané. Cette interaction est d’autant plus forte que les deux molécules sont facilement polarisables, donc d’autant plus forte que leur taille est importante. Les forces de LONDON étant présentes entre toutes les particules, quelle que soit leur nature, ce sont principalement elles qui permettent la cohésion de la matière dans l’univers.

Notons que les énergies de Van der Waals varient en fonction de l’inverse de la distance à l’ordre 6.

Il existe toutefois d’autres types d’interactions électrostatiques que celles décrites par le modèle des forces de VdW. En effet, similaire dans l’esprit à l’énergie de KEESOM, l’énergie d’interaction entre un ion et un dipôle permanent est donné par la formule suivante :

⁴La polarisabilité est la capacité du nuage électronique à se déformer sous l’action d’un champ électrique. Le barycentre des charges négatives (électrons) étant légèrement décalé par rapport à celui des charges positives (noyaux) sous l’effet du champ, un moment électronique induit \vec{m}_e apparaît, engendrant la notion de polarisabilité α .

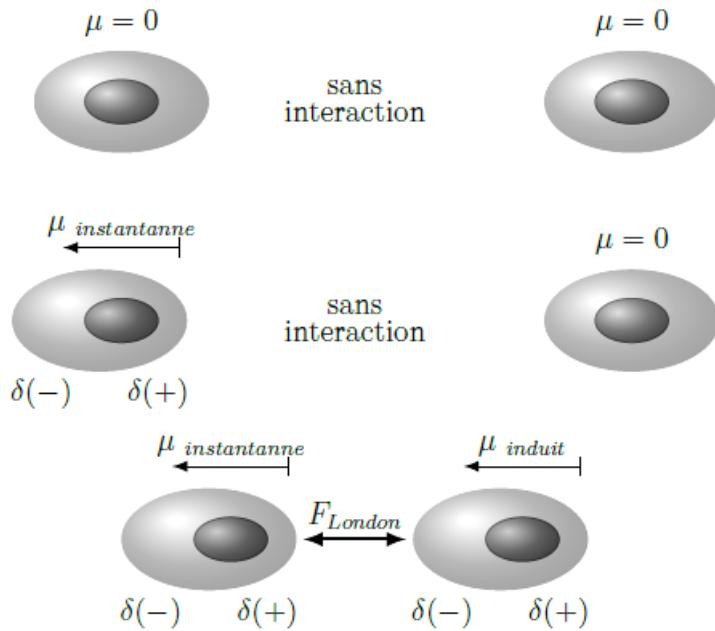


Figure 2.4: Interaction entre deux dipôles instantanés.

$$E_{ion-dipôle} = -\frac{1}{r^2} \frac{\mu_1 q_2}{4\pi\epsilon_0\epsilon} \quad (2.33)$$

Il s'agit là encore de l'interaction positive entre une espèce chargée, anion ou cation, et une molécule possédant un moment dipolaire non-nul (2.5). Ce phénomène est notamment à l'origine de la dissolution des espèces ioniques (ex : NaCl) dans un solvant polaire, de l'étape de solvatation qui suit, puis de la bonne dispersion des charges en solution.

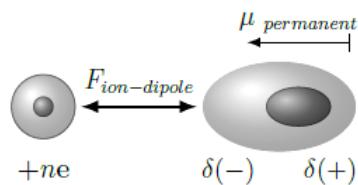


Figure 2.5: Interaction entre une espèce chargée et une molécule polaire.

De plus, notons aussi la spécificité de la liaison hydrogène qui est une liaison chimique non covalente, de type dipôle-dipôle. Lorsqu'un hétéroatome possédant au moins une paire libre est suffisamment électronégatif (ex : O, N, F), il vient se positionner aux abords d'un hydrogène acide porté par un autre atome fortement électronégatif afin d'en stabiliser la charge partielle δ(+) ainsi créée.

Bien que de la même famille que les forces de Van der Waals, *i.e.* électrostatique, les liaisons

hydrogènes s'en distinguent par une intensité environ dix fois supérieure. Elles restent toutefois une vingtaine de fois plus faibles qu'une liaison covalente. La distance moyenne entre les deux hétéroatomes est de l'ordre de 2,5 Å.

Dans le cadre de cette thèse, ce sont principalement les interactions entre systèmes conjugués, *i.e.* interactions $\pi - \pi$, qu'il sera nécessaire de traduire. Comme nous pouvons le voir dans le cas simple d'un dimère de benzène représenté en figure 2.6, l'interaction positive se fait entre les liaisons σ et les liaisons π alors que les nuages électroniques des liaisons π se repoussent naturellement, dû à leurs charges négatives.

Notons que nous retrouvons ce phénomène sous des aspects intramoléculaires, comme par exemple dans le cas de l'hyperconjugaison $\sigma - \pi$ qui vient stabiliser certaines conformations du toluène.

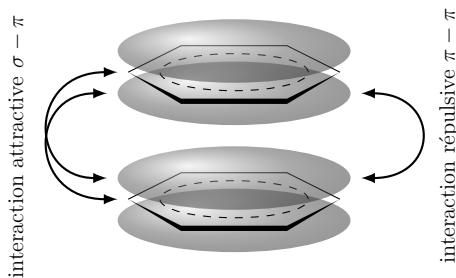


Figure 2.6: π -stacking dans le cas d'un dimère de benzène

Même si la théorie de la fonctionnelle de la densité connaît un large succès tant elle arrive désormais bien à traduire les phénomènes de liaisons chimiques, les structures géométriques et même la cohésion des solides moléculaires et cristallins, il reste toujours l'obstacle des systèmes chimiques où les forces de Van der Waals sont prédominantes. En effet, les effets de corrélation électronique des forces de dispersion étant purement non-locaux, l'approximation locale ou non-locale qui fait le fondement de la DFT restera problématique. Se pose alors la question de savoir comment modéliser ces types d'interaction de façon idiomatique. Nous allons voir que l'élaboration d'une fonctionnelle hybride à longue portée est capable de répondre à cette problématique.

2.5 Construction d'une LC-DFT-D hybride : cas de la ωBXD

Les DFT hybrides avec correction à longue portée basées sur la théorie KOHN-SHAM ont naturellement rencontré un grand engouement puisque la précision apportée n'accroît pas le coût calculatoire par rapport aux DFT hybrides.

2.5.1 B88

Comme nous l'avons vu dans le cadre des approximations de la fonctionnelle de la densité, notées DFAs⁵, la décroissance en exponentielle du potentiel d'échange-corrélation, au lieu d'être en $1/r$, engendre une mauvaise représentation des interactions à longue distance. Cette erreur, nommée erreur d'auto-interaction (SIE, pour “self-interaction error”), est liée au fait que ces approximations, basées sur la densité de spin locale (LSDA, pour “local spin density approximation”), décrivent mal l'état fondamental qui devrait être, dans le cadre de la DFT pure, strictement sans auto-interaction. C'est pourquoi, afin d'introduire un effet non-local de l'échange-corrélation dans le modèle KS-DFT (partie 2.3), BECKE proposa en 1988 d'incorporer dans sa fonctionnelle d'échange B88 [50] une petite part d'échange exact HARTREE-FOCK.

Dans le cadre général des DFAs, l'énergie d'échange-corrélation s'écrit donc :

$$E_{xc} = c_x E_x^{HF} + E_{xc}^{DFA} \quad (2.34)$$

où c_x prend généralement des valeurs comprises entre 0,2 et 0,25 [49] pour les données thermodynamiques et entre 0,4 et 0,6 [52] pour les études cinétiques.

Basée sur ce modèle, la désormais bien connue DFT hybride B3LYP [49] (équation 2.26) donne des résultats comparables à ceux obtenus à partir de la théorie perturbative MØLLER-PLESSET à l'ordre 2 [53], noté MP2, souvent utilisée comme référence, dans le cadre de systèmes fortement liés. Depuis, de nombreuses recherches ont porté sur l'amélioration constante de ce potentiel d'échange-corrélation $E_{xc}[\rho]$.

⁵“Density Functional Approximations”

2.5.2 B97

Une avancée significative a de nouveau été faite par BECKE en 1997 dans le domaine des KS-DFT. Par une méthode similaire à la combinaison linéaire d'orbitales atomiques, notée LCAO⁶, il a proposé un modèle mathématique basé sur l'approximation de densité de spin local (LSDA), sa première dérivée et une petite fraction d'échange HF pour décrire le potentiel d'échange-corrélation $E_{xc}[\rho]$. Une optimisation systématique des coefficients linéaires à partir d'un jeu classique de données expérimentales a conduit à l'apparition de la méthode B97 [54]. La base de données contient notamment des valeurs relatives à l'interaction entre systèmes conjugués.

Cette méthodologie a été reprise par F. A. HAMPRECHT et al, P. J. WILSON et al et T. W. KEAL et al pour respectivement conduire à la B97-1 [55] (1998), la B97-2 [56] (2001) et la B97-3 [57] (2005). Il s'agissait alors de réoptimisations des coefficients linéaires par rapport à d'autres bases de données expérimentales plus complètes.

Mais cette reparamétrisation empirique du terme d'échange-corrélation ne résout pas le problème de sa non-décroissance en $1/r$. La prise en compte totale du terme d'échange HF E_x^{HF} ($c_x=1$ dans l'équation 2.34) pourrait résoudre ce problème mais cela serait incompatible avec le terme de corrélation DFA E_c^{DFA} . En effet, il existerait alors une mauvaise compensation des erreurs respectives.

2.5.3 ω B97

L'idée de séparer le traitement des interactions courtes (SR, pour "short range") et longues portées (LR, pour "long range") s'est alors présentée comme le choix le plus évident, aussi bien au niveau de la compréhension des phénomènes que mathématiquement parlant. Nous pouvons ainsi traiter séparément à l'aide d'une fonction erreur (*erf*) les interactions à courtes distances par une fonctionnelle de la densité et celles longue distance par une fonction d'onde. Ce principe conduit naturellement à l'élaboration d'une fonctionnelle hybride à séparation de portée. L'introduction de la fonction erreur, avec un paramètre libre, permet de contrôler le rayon d'action des interactions de courte-portée.

La première proposition faite par IIKURA et al [58] a été de traiter la partie d'échange LR par la théorie HF alors que la partie SR est approximée par une DFA; le terme de corrélation est quant à lui le même que celui de COULOMB, quelle que soit la distance :

⁶"Linear Combination of Atomic Orbitals".

$$E_{xc}^{LC-DFA} = E_x^{LR-HF} + E_x^{SR-DFA} + E_c^{DFA} \quad (2.35)$$

Ce schéma de séparation de portée a l'avantage de conduire à des temps de calcul très proches des DFT hybrides, mais il reste à développer une fonctionnelle d'échange SR précise et une fonctionnelle de corrélation qui soit entièrement compatible entre elles.

Le type d'opérateur de coupure le plus utilisé dans le cadre des LC-DFT hybrides est la fonction d'erreur standard (erf) :

$$\frac{1}{r} = \frac{erf(\omega r_{12})}{r_{12}} + \frac{erfc(\omega r_{12})}{r_{12}} \quad (2.36)$$

avec $\frac{erf(\omega r_{12})}{r_{12}}$: interaction de courte portée,
 $\frac{erfc(\omega r_{12})}{r_{12}}$: interaction complémentaire,
 r_{12} : distance entre les particules 1 et 2,
 ω : paramètre contrôlant la séparation.

Notons que l'introduction du paramètre ω , qui s'exprime comme l'inverse d'une distance, permet de donner un sens physique à cette valeur, en cela qu'il est étroitement lié à une longueur caractéristique de la séparation. Naturellement, il existe différents types de fonctions erreur (erf) afin de faciliter son intégration mathématique dans les codes de calculs. Dans le cas de la ω B97 [59] et, par conséquent, des fonctionnelles ω B97X et ω B97X-D, c'est la fonction $erf/erfc$ qui a été choisie par Jeng-Da CHAI et Martin HEAD-GORDON dans leurs travaux.

Le choix des auteurs s'est porté sur un terme d'échange exact HF longue portée E_x^{LR-HF} , calculé à partir des spin-orbitales occupées $\phi_{i\sigma}(r)$, et une forme analytique du terme d'échange E_x^{SR-DFA} obtenue par l'intégration du carré de la matrice densité LSDA :

$$E_x^{LR-HF} = -\frac{1}{2} \sum_{\sigma} \sum_{ij}^{occ.} \iint \phi_{i\sigma}^*(r_1) \phi_{j\sigma}^*(r_1) \frac{erf(\omega r_{12})}{r_{12}} \phi_{i\sigma}(r_2) \phi_{j\sigma}(r_2) dr_1 dr_2, \quad (2.37)$$

$$E_x^{SR-LSDA} = \sum_{\sigma} \int \underbrace{-\frac{3}{2} \left(\frac{3}{4\pi} \right)^{1/3} \rho_{\sigma}^{4/3}(r) F(a_{\sigma})}_{e_{x\sigma}^{SR-LSDA}(\rho_{\sigma}).dr} . \quad (2.38)$$

où :

$$k_{F\sigma} = (6\pi^2 \rho_{sigma}(r))^{1/3},$$

$$F(a_\sigma) = 1 - \frac{8}{3}a_\sigma \left[\sqrt{\pi} \operatorname{erf} \left(\frac{1}{2a_\sigma} \right) - 3a_\sigma + 4a_\sigma^3 + (2a_\sigma - 4a_\sigma^3) \exp \left(-\frac{1}{4a_\sigma^2} \right) \right],$$

$$a_\sigma = \frac{\omega}{2k_{F\sigma}}.$$

avec $k_{F\sigma}$: vecteur d'onde local de Fermi,
 $F(a_\sigma)$: fonction d'atténuation,
 a_σ : paramètre de contrôle (sans unité) de la fonction
d'atténuation $F(a_\sigma)$.

En retenant une fonctionnelle de corrélation basée elle aussi sur la LSDA E_c^{LSDA} , la plus simple des DFT hybrides à correction de longue portée (RSHX-LDA) s'écrit :

$$E_{xc}^{RSHXLDA} = E_x^{LR-HF} + E_x^{SR-LSDA} + E_c^{LSDA} \quad (2.39)$$

La fonctionnelle ω B97[59] s'écrit alors :

$$E_{xc}^{\omega B97} = E_x^{LR-HF} + E_x^{SR-B97} + E_c^{B97} \quad (2.40)$$

Il est à noter que celle-ci ne possède pas d'échange HARTREE-FOCK à courte portée (SR), comme la plupart des fonctionnelles hybrides à correction de portée.

Malgré plusieurs études visant à optimiser la valeur du paramètre ω , la précision calculatoire reste insuffisante en terme de thermochimie. En effet, nous l'avons déjà vu, une valeur trop grande pour ω tendrait vers une incompatibilité entre le terme d'échange non-local E_x^{LR-HF} et le terme local de corrélation E_c^{LSDA} . De plus, nous pouvons aisément comprendre, d'après l'équation 2.36, que plus ω est petit, plus la contribution du terme d'échange SR $E_x^{SR-LSDA}$ sera importante. L'utilisation d'une trop faible valeur reviendrait alors à traiter le problème dans un cadre très proche de la LDA classique qui, comme nous l'avons vu dans la partie 2.4.1, est incapable de traduire correctement le terme d'échange à courte portée.

2.5.4 ωB97X

Afin d'y remédier, une partie d'échange SR HF E_x^{SR-HF} , est ajoutée à $E_x^{SR-LSDA}$ dans une proportion d'environ 16%, de la même manière que BECKE dans la fonctionnelle B88. Ceci à l'avantage de ne pas perturber la partie LR qui est dorénavant correcte. Ainsi, la nouvelle fonctionnelle comporte désormais un paramètre c_x contrôlant la proportion d'échange exact HF à courte distance, comme nous pouvons le voir dans son expression :

$$E_{xc}^{LC-DFA} = E_x^{LR-HF} + c_x E_x^{SR-HF} + E_x^{SR-DFA} + E_c^{DFA} \quad (2.41)$$

où :

$$E_x^{SR-HF} = -\frac{1}{2} \sum_{\sigma} \sum_{ij}^{occ.} \iint \phi_{i\sigma}^*(r_1) \phi_{j\sigma}^*(r_1) \frac{erfc(\omega r_{12})}{r_{12}} \phi_{i\sigma}(r_2) \phi_{j\sigma}(r_2). dr_1. dr_2, \quad (2.42)$$

C'est ainsi que la fonctionnelle ωB7X [59] se décompose de la façon suivante :

$$E_{xc}^{\omega B97X} = E_x^{LR-HF} + c_x E_x^{SR-HF} + E_x^{SR-B97} + E_c^{B97} \quad (2.43)$$

La valeur de ω , comme les valeurs des coefficients de développements linéaires et de développements à l'ordre m des fonctionnelles ωB97 et ωB97X ont été déterminées par la méthode des moindres carrés appliquée à une base de données composées de 412 valeurs précises, expérimentales et théoriques.

Malgré toutes ces optimisations conduisant à une bien meilleure représentation des systèmes en interaction, ces fonctionnelles connaissent encore des lacunes quant à la traduction des interactions de dispersion entre atomes, ie les forces de LONDON. Comme nous allons le voir dans le cas de la fonctionnelle $\omega\text{B97X-D}$, ceci peut être corrigé par une prise en compte empirique des effets de dispersion.

2.5.5 $\omega\text{B97X-D}$

Cette dernière correction pourrait naturellement passer par le calcul idiomatique de l'énergie de dispersion entre chaque atome, mais cela occasionnerait alors un coût calculatoire prohibitif. C'est pourquoi Jeng-Da CHAI et Martin HEAD-GORDON ont fait le choix d'appliquer cette correction de façon empirique par l'ajout d'un terme E_{disp} à la fonctionnelle KS-DFT, ici la

ω B97X. L'expression de l'énergie de la fonctionnelle ω B97X-D [59] ainsi créée devient alors :

$$E_{DFT-D} = E_{\omega B97X} + E_{disp} \quad (2.44)$$

L'énergie de dispersion E_{disp} est définie par rapport à une fonction d'amortissement f_{damp} :

$$E_{disp} = - \sum_{i=1}^{N_{at}-1} \sum_{j=i+1}^{N_{at}} \frac{C_6^{ij}}{R_{ij}^6} f_{damp}(R_{ij}) \quad (2.45)$$

où :

$$f_{damp}(R_{ij}) = \frac{1}{1 + a(\frac{R_{ij}}{R_r})^{-12}} \quad (2.46)$$

Une nouvelle fois, la partie empirique a été paramétrée par rapport à la même base de données que pour les fonctionnelles ω B97 et ω BX97.

Conclusion

En résumé, l'apport de la fonction erreur (*erf*) permet de mieux gérer les contributions d'échange-corrélation selon la distance d'interaction. Les DFT hybrides ω B97 et ω B97X prennent ainsi en compte la totalité de l'échange exact à longue distance et utilisent la méthode des gradients généralisés à faible distance, alors que la corrélation électronique reste basée sur celle initialement développée par BECKE dans la fonctionnelle B97. Ceci a pour effet de supprimer le problème d'auto-interaction de la fonctionnelle d'échange à longue distance.

Les travaux de Jeng-Da CHAI et Martin HEAD-GORDON ont finalement conduit à la fonctionnelle ω B9X-D, de type LC-DFT-D hybride, où la totalité de l'échange exact HF est pris en compte à longue distance, en même temps qu'une petite partie – environ 22 % – de l'échange exact HF est introduite à courte distance pour compléter une fonctionnelle d'échange B97 modifiée ; une correction empirique de la dispersion est finalement appliquée.

Comme toutes les fonctionnelles LC-DFT, le problème de l'auto-interaction est corrigé à longue distance mais reste quelque peu présent à courte distance. Les effets de corrélation à longue distance sont quant à eux purement et simplement traités par la correction empirique de dispersion.

Cette fonctionnelle est, d'après les tests des auteurs, définitivement plus adaptée à l'étude de systèmes chimiques où les interactions non-covalentes sont importantes.

Chapter 3

Partie vibrationnelle

Introduction

La littérature consacrée aux applications des spectrométries vibrationnelles dans le domaine de la caractérisation des constituants présent dans les pétroles est relativement peu abondante même si la spectrométrie infrarouge est devenue une technique d'analyse de "routine" dans de très nombreux laboratoires, académiques comme industriels. La raison de cette faible abondance réside essentiellement dans le fait de la complexité des mélanges qui constituent un pétrole, un asphaltène ... Pourtant les champs de cette technique d'applications se sont considérablement développés depuis l'apparition sur le marché de spectrophotomètres à transformée de FOURIER. Après avoir vu sa position privilégiée menacée par d'autres méthodes, telles que la spectroscopie RMN¹ ou la spectrométrie de masse, la spectrométrie infrarouge à transformée de FOURIER (IRTF) a connu de nouvelles avancées technologiques, telle que la spectrométrie photoacoustique que nous avons employés dans ce travail, qui lui confèrent à l'heure actuelle une précision d'analyse permettant d'atteindre des informations détaillées sur :

- la structure chimique de molécules, de macromolécules: identification de l'unité de base, des ramifications ; analyses des extrémités de chaînes ; identification des défauts, d'éventuelles impuretés...
- les interactions intra et inter-moléculaires, la conformation des chaînes, l'orientation des molécules et des macromolécules, les auto-associations éventuelles, *etc...*

Cette partie a pour but essentiel de définir le vocabulaire et les notions fondamentales que nous emploierons par la suite, le développement détaillé du traitement classique de la vibration se trouvant dans de nombreux ouvrages et repris dans de nombreuses thèses. Parallèlement à ces développements expérimentaux, la modélisation en spectroscopie vibrationnelle a connu ces dernières décennies de très grandes mutations. Ce chapitre vise aussi à rappeler les fondements essentiels de la résolution de l'équation vibrationnelle de Schrödinger, qui sous-tendent les principales méthodes – que nous développerons – actuellement disponibles pour tenter de répondre aux problématiques posées par les expérimentateurs.

En particulier, le domaine de la pétrochimie qui nous intéresse dans ce travail, et précisément la thématique visant à l'élucidation de la composition de ces mélanges complexes et composés de

¹"Nuclear Magnetic Resonance Spectroscopy".

milliers de molécules différentes non encore identifiées, est un domaine au champ d’investigation large et pour lequel les attentes sont grandes en termes de caractérisation comme en terme de compréhension des processus d’interaction et d’agrégation de ces molécules. Dans ce domaine encore, la spectrométrie IRTF est certainement un des outil le plus efficace pour élucider les mécanismes impliqués, mais les données résultantes sont complexes et de fait difficilement interprétables et justifiables sans un support théorique adapté.

L’Équipe de Chimie Physique (ECP) de l’IPREM est depuis très longtemps spécialisée dans les développements méthodologiques et logiciels dans la double hypothèse des anharmonicités électriques et mécaniques. Pour ces compétences, les expérimentateurs font généralement appel à la modélisation, dans le but d’interpréter leurs données spectrales. Le problème qui nous a été posé dans ce travail a cependant constitué un challenge qui nous a contraint à adapter nos méthodes de calculs pour proposer, *in fine*, une méthode de type variation-perturbation adaptée et simplifiée permettant une interprétation des données dans des gammes spectrales non usuelles (très bas nombres d’ondes) et sur un ensemble de molécules suffisamment large et représentatif de quelques familles moléculaires suspectées comme étant présentent dans les asphaltènes.

3.1 Généralités

L'identification des principaux composés – ou, pour le moins, des familles de composés – chimiques fait l'objet, depuis de nombreuses années, de recherches intensives.

En soit, la thématique liée à l'identification et à la caractérisation des molécules au sein d'un milieu chimique donné n'a rien de novatrice. En effet, depuis de nombreuses décennies, les chimistes de tous domaines recherchent ce “graal” avec plus ou moins de succès. Parmi les techniques expérimentales les plus utilisées pour répondre au problème, la spectroscopie vibrationnelle est certainement celle qui a permis le plus grand nombre de progrès dans des domaines aussi variés que la biochimie, l'agroalimentaire, la chimie interstellaire ou encore la chimie des matériaux. Le point commun à l'ensemble de ces études est qu'elles sont toutes basées sur une connaissance *a minima* des molécules constituant le milieu étudié. La complexité des problèmes de suivi et de devenir d'un ensemble de molécules mal défini et ayant évolué dans des conditions extrêmes sur une échelle de temps démesurée (en terme de réactivité chimique) rend toutefois l'utilisation de cette technique plus délicate et plus hasardeuse, si bien qu'il est indispensable de faire appel à des techniques complémentaires ou de recourir au soutien de la modélisation prédictive. Il est en effet indéniable que les progrès conjoints des techniques de modélisation et des moyens informatiques font aujourd'hui de cet outil un support indispensable et performant à l'identification de systèmes moléculaires de plus en plus variés.

Le développement de ces modélisations en spectroscopie vibrationnelle fait état depuis une vingtaine d'années de progrès fulgurants. Les techniques mathématiques développées par les générations précédentes sont désormais largement éprouvées et mises en application au service des expérimentateurs. Il ne se passe plus une seule année sans que les limites calculatoires et les précisions atteintes par ces simulations ne soient repoussées, grâce au développement de méthodes adaptées et développées dans le cadre d'hypothèses mathématiques précises et contrôlées.

Le travail présenté dans ce chapitre s'inscrit donc dans le cadre de ces développements mathématiques au service de l'identification de systèmes chimiques complexes. Les calculs que nous développons sont réalisés dans le cadre de la Résolution de l'Equation vibrationnelle de

Schödinger (RES), dans la double hypothèse des approximations anharmoniques électriques et mécaniques permettant d'accéder à la détermination des intensités et des fréquences de tous les modes de vibrations intrinsèques à un système chimique donné, dans un environnement donné. Il est encore communément admis qu'un calcul mené dans une hypothèse plus simple, dite harmonique, suffit aux identifications. En réalité, les raisons fondamentales qui poussent les chercheurs à préférer l'approximation harmonique sont aussi bien liées au problème de coût calculatoire autant qu'au manque d'implémentation d'approches de type anharmoniques dans les grands codes de calculs commerciaux. Dans la stratégie que nous développons, nos calculs se distinguent des études menées dans le domaine en cela qu'ils se placent précisément dans l'hypothèse anharmonique. Néanmoins, il est important de savoir qu'un calcul réalisé dans l'hypothèse harmonique engendre une erreur que le modélisateur à pour habitude de "contrôler" par un facteur correctif adapté qu'il applique à ses résultats selon les conditions de calculs utilisées lors de la RES. Malheureusement, cette technique de calcul, utilisée depuis une cinquantaine d'années, et qui a fait de la modélisation en spectroscopie vibrationnelle une méthode quelque peu empirique dans l'esprit des expérimentateurs, est toujours assujettie à un doute quant à l'identification précise des vibrateurs, car il n'est fondamentalement pas concevable que l'erreur commise sur chaque mode soit la même pour tous et que la correction proposée soit universellement applicable à tous les système étudiés quels que soient les milieux dans lesquels ils se trouvent. De plus, les calculs développés dans cette hypothèse ne permettent pas d'identifier d'autres vibrateurs que les modes fondamentaux puisqu'aucun couplage entre modes n'est pris en compte.

En résumé, toute modélisation en spectroscopie vibrationnelle, qu'elle soit développée dans l'hypothèse harmonique ou anharmonique, est directement dépendante de la qualité de la fonction d'onde moléculaire électronique, donc de la prise en compte de la corrélation électronique. S'il est aujourd'hui commun de réaliser la REVS pour des systèmes de petite taille (3, 4 atomes), ce critère devient toutefois pratiquement rédhibitoire lorsqu'il s'agit de résoudre ces mêmes problèmes dans l'hypothèse anharmonique sur des systèmes de taille plus importante.

Ce chapitre constituera avant tout une occasion pour moi de recenser les difficultés inhérentes au développement méthodologique vibrationnel et de montrer les pistes avancées pour l'étude des systèmes moléculaires dont la taille excède la vingtaine d'atomes, taille minimale nécessaire à la caractérisation des motifs/familles de base présentes dans les asphaltènes. Ces activités s'inscrivent dans le prolongement et le complément des actions antérieures menées au

sein de l'ECP, notamment pour le développement de méthodes de variation-perturbation et de calcul des intensités IR de petits systèmes moléculaires [60].

3.2 Séparation des mouvements rotationnels et vibrationnels

Cette partie a pour but essentiel de définir le vocabulaire et les notions fondamentales que nous emploierons par la suite, le développement détaillé du traitement classique de la vibration se trouvant dans de nombreux ouvrages [61, 62, 62] et repris dans de nombreuses thèses [63, 64].

Dans une approximation d'ordre 0 supplémentaire à celle de BORN- OPPENHEIMER, les mouvements nucléaires peuvent être séparés en deux classes : les mouvements de rotation et les mouvements de vibration. Pour ce faire, il est nécessaire d'expliciter l'expression de l'énergie cinétique des noyaux au sens classique et de repérer la molécule dans un référentiel respectant les conditions d'ECKART [65] Considérons dans cet espace un repère mobile $oxyz$, lié à la molécule, et un repère fixe $OXYZ$, définissant les mouvements de translation et de rotation du repère mobile. Le mouvement du trièdre mobile par rapport au trièdre fixe est défini par la distance R et la vitesse angulaire instantanée α . Le mouvement de la molécule est défini par le trièdre mobile représentant à chaque instant la position \vec{r}_α des noyaux α par rapport à leur position d'équilibre \vec{a}_α . Soit :

$$\vec{\rho}_\alpha = \vec{r}_\alpha - \vec{a}_\alpha \quad (3.1)$$

La vitesse \vec{v}_α du α^{ieme} noyau est donc : $\vec{v}_\alpha = \vec{r}'_\alpha = \vec{\rho}'_\alpha$ puisque, par définition, \vec{a}_α est constant dans le temps.

Ainsi, dans le repère fixe, la vitesse de ce noyau s'écrit :

$$\vec{V}_\alpha = \vec{R} + (\vec{\omega} \wedge \vec{r}_\alpha) + \vec{v}_\alpha \quad (3.2)$$

On peut facilement déduire de cette expression l'énergie cinétique totale des noyaux :

$$\begin{aligned}
2T = & \dot{R}^2 \sum_{\alpha} m_{\alpha} (\vec{\omega} \wedge \vec{r}_{\alpha})^2 + \sum_{\alpha} m_{\alpha} v_{\alpha}^2 \\
& + 2 \dot{R} \sum_{\alpha} m_{\alpha} \vec{v}_{\alpha} \\
& + 2 \left(\vec{R} \wedge \vec{\omega} \right) \sum_{\alpha} m_{\alpha} \vec{r}_{\alpha} \\
& + 2 \vec{\omega} \sum_{\alpha} m_{\alpha} (\vec{r}_{\alpha} \wedge \vec{v}_{\alpha})
\end{aligned} \tag{3.3}$$

Si nous supposons que les noyaux ne possèdent aucun mouvement de translation dans le système mobile et que l'origine de ce dernier correspond au centre de gravité de la molécule, alors :

$$\sum_{\alpha} m_{\alpha} \vec{v}_{\alpha} = 0 \tag{3.4}$$

et

$$\sum_{\alpha} m_{\alpha} \vec{r}_{\alpha} = 0 \tag{3.5}$$

Si, de plus, nous considérons que dans le trièdre mobile la molécule ne possède aucun mouvement de rotation, nous pouvons écrire :

$$\sum_{\alpha} m_{\alpha} (\vec{a}_{\alpha} \wedge \vec{v}_{\alpha}) = 0 \tag{3.6}$$

et donc

$$\sum_{\alpha} m_{\alpha} (\vec{r}_{\alpha} \wedge \vec{v}_{\alpha}) = \sum_{\alpha} m_{\alpha} (\vec{p}_{\alpha} \wedge \vec{v}_{\alpha}) \tag{3.7}$$

Les deux conditions ci-dessus portent le nom de conditions d'ECKART et simplifient l'expression 3.3 :

$$2T = \dot{R}^2 \sum_{\alpha} m_{\alpha} + \sum_{\alpha} m_{\alpha} (\vec{\omega} \wedge \vec{r}_{\alpha})^2 + \sum_{\alpha} m_{\alpha} v_{\alpha}^2 + 2 \vec{\omega} \sum_{\alpha} m_{\alpha} (\vec{r}_{\alpha} \wedge \vec{v}_{\alpha}) \tag{3.8}$$

Le premier terme correspond à l'énergie cinétique de translation de la molécule. Il ne contribue pas à la quantification de l'énergie. Le second terme correspond à l'énergie cinétique de rotation. Le troisième terme correspond à l'énergie de vibration moléculaire. Le dernier terme est appelé terme de CORIOLIS. Il est relatif à l'interaction entre la rotation et la vibration, et

peut être négligé si nous considérons que les mouvements vibrationnels sont de faible amplitude : $\vec{r}_\alpha \approx \vec{a}_\alpha$. Cette approximation appelée condition de CASIMIR [66] est généralement vérifiée pour les vibrations d'elongation, par opposition aux modes très mous de torsion, qui restent souvent mal traduits dans ce cadre. En conséquence, l'énergie cinétique des noyaux peut s'écrire, en première approximation, comme la somme d'un terme rotationnel et d'un terme vibrationnel :

$$2T_n = 2T_R + 2T_V \text{ en supposant } 2T_{VR} = 0 \quad (3.9)$$

D'un point de vue quantique, les considérations ci-dessus conduisent à séparer les mouvements rotationnels et vibrationnels de l'équation nucléaire en deux équations distinctes :

$$\psi_{R_\alpha}^n = \psi_{R_\alpha}^R \bullet \psi_{R_\alpha}^V \quad (3.10)$$

$$E_n = E_V + E_R \quad (3.11)$$

avec $\psi_{R_\alpha}^R$: fonction d'état rotationnelle,

$\psi_{R_\alpha}^V$: fonction d'état vibrationnelle,

E_R : énergie correspondant à la fonction d'état rotationnelle,

E_V : énergie correspondant à la fonction d'état vibrationnelle.

On obtient ainsi l'équation de Schrödinger décrivant les mouvements vibrationnels :

$$(\hat{T}_V + \hat{V}_V) \psi_{R_\alpha}^V = E_V \psi_{R_\alpha}^V \quad (3.12)$$

et l'équation de Schrödinger décrivant les mouvements rotationnels dans l'hypothèse où les liaisons interatomiques ne s'allongent pas pendant la rotation (hypothèse du rotateur rigide):

$$T_R \psi_{R_\alpha}^R = E_R \psi_{R_\alpha}^R \quad (3.13)$$

3.3 Energie cinétique de vibration

L'énergie cinétique de vibration d'une molécule composée de n atomes dans le repère d'ECKART s'écrit :

$$2T_V = \sum_{\alpha}^n m_{\alpha} (\dot{x}_{\alpha}^2 + \dot{y}_{\alpha}^2 + \dot{z}_{\alpha}^2) \quad (3.14)$$

avec $\dot{x}_{\alpha}, \dot{y}_{\alpha}, \dot{z}_{\alpha}$: composantes de la vitesse $\vec{\rho}_{\alpha}$ de l'atome α .

En exprimant cette énergie dans le système de coordonnées cartésiennes pondérées par les masses ($q_x = m_{\alpha}^{1/2}x_{\alpha}, q_y = m_{\alpha}^{1/2}y_{\alpha}$ et $q_z = m_{\alpha}^{1/2}z_{\alpha}$), et sans labelliser les axes cartésiens, nous obtenons une écriture simplifiée, explicitement fonction de $3n$ coordonnées :

$$2T_V = \sum_i^{3n} \dot{q}_i^2 \quad (3.15)$$

Matriciellement, l'équation ci-dessus prend la forme :

$$2T_V = [\dot{q}]^t [\dot{q}] \quad (3.16)$$

avec \dot{q}_i : dérivée de q_i en fonction du temps $\left(\frac{dq_i}{dt}\right)$,
 $[\dot{q}]^t$: matrice transposée de $[\dot{q}]$.

Notons que dans l'espace des coordonnées cartésiennes non pondérées, nous avons :

$$2T_V = [\dot{x}]^t [M_{\alpha}] [\dot{x}] \quad (3.17)$$

3.4 Energie potentielle harmonique

La fonction potentielle s'écrit généralement comme un développement en série de TAYLOR au voisinage de la position d'équilibre. Dans l'espace des coordonnées ci-dessus, elle prend la forme d'un polynôme caractéristique d'ordre n , dont nous limitons le développement à l'ordre 2 dans l'hypothèse harmonique :

$$V = V_{eq} + \sum_i^{3n} \left(\frac{\partial V}{\partial q_i} \right)_{eq} q_i + \frac{1}{2!} \sum_{i \leq j}^{3n, 3n} \left(\frac{\partial^2 V}{\partial q_i \partial q_j} \right)_{eq} q_i q_j \quad (3.18)$$

Les coefficients de ce polynôme représentent les dérivées $n^{ièmes}$ de la fonction potentielle à la structure géométrique d'équilibre. Pour cette configuration d'équilibre, V est égale à V_{eq} ; ce terme d'ordre 0 peut être pris comme référence. De plus, si l'état électronique est un état stable,

ce qui est le cas lorsque la molécule peut vibrer, les dérivées premières $\left(\frac{\partial V}{\partial q_i}\right)_{eq}$ sont nulles. Les coefficients d'ordre 2 sont appelés constantes de forces quadratiques ou harmoniques (notées $f_{ij}^{(q)}$ dans cet espace) et constituent le champ de force harmonique. Ainsi, la fonction potentielle harmonique s'écrit :

$$2V = \sum_{i \leq j}^{3n,3n} f_{ij}^{(q)} q_i q_j \quad (3.19)$$

soit, sous sa forme matricielle :

$$2V = [q]^t [f^q] [q] \quad (3.20)$$

3.5 Équations de Lagrange

3.5.1 Espace des coordonnées cartésiennes pondérées par les masses

La détermination des $3n$ mouvements vibrationnels et de leur fréquence s'effectue en résolvant les $3n$ équations de LAGRANGE à partir de la connaissance des deux fonctions fondamentales de la mécanique, exprimées dans les deux sous-paragraphes précédents :

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) + \frac{\partial V}{\partial q_i} = 0 \quad (3.21)$$

soit :

$$\ddot{q}_i + \sum_j^{3n} f_{ij}^{(q)} q_j \quad (3.22)$$

Les solutions sont de la forme $q_i = q_i^{\ddagger} \cos \lambda^{1/2} t$, où q_i^{\ddagger} est l'amplitude maximale du mode i et $\lambda^{1/2}$ est relié à sa fréquence de vibration. Pour déterminer la valeur des $3n$ fréquences, nous injectons ses solutions particulières dans les $3n$ équations. Matriciellement, cette opération revient tout simplement à diagonaliser la matrice $[f^q]$. On obtient 6 valeurs propres nulles qui correspondent aux trois translations et trois rotations de la molécule (deux rotations si la molécule est linéaire) et $3n - 6(5)$ valeurs propres non nulles, correspondant aux vibrations de la molécule. Nous déduisons de ces valeurs propres les nombres d'ondes ϖ (exprimés en cm^{-1}) et les fréquences de vibration ω (en Hz) par la relation :

$$\lambda^{1/2} = 2\pi c\varpi = 2\pi\omega \quad (3.23)$$

avec c : vitesse de la lumière.

Lorsque certaines valeurs propres sont identiques, ce qui correspond à deux ou trois mouvements vibrationnels différents mais de même fréquence, ces modes sont dits doublement ou triplement dégénérés. Les vecteurs propres représentent le mouvement des atomes en coordonnées cartésiennes pondérées par les masses, induits par les $3n - 6(5)$ vibrations. Ces mouvements, propres à chaque vibration, se nomment modes normaux de vibration. Ce nouvel espace, représenté par un repère orthonormé où chaque dimension correspond à un mouvement vibrationnel harmonique de la molécule, constitue une base de construction de l'Hamiltonien vibrationnel dans le traitement quantique.

3.5.2 Espace des coordonnées internes : résolution par la méthode de Wilson

Le choix de cet espace permet de réduire la dimension des équations à traiter en éliminant les valeurs propres nulles correspondant aux translations et aux rotations de la molécule. Ceci est possible si nous choisissons un référentiel qui obéit aux conditions d'ECKART. De plus, les vibrations moléculaires sont étudiées en fonction des variations des longueurs de liaison et des déformations angulaires, ce qui permet d'attribuer un sens physique aux constantes de force calculées dans cet espace. Ici, l'expression de l'énergie cinétique est plus compliquée, car il est nécessaire de définir une matrice de passage de dimension $(3n - 6) \times 3n$ (notée B) entre les coordonnées cartésiennes de déplacements x et les coordonnées de déplacements r : $[r] = [B][x]$. D'après l'équation 3.17 et dans l'hypothèse d'une transformation à coefficients constants, valable pour les petits mouvements, l'énergie cinétique s'écrit :

$$2T = [\dot{r}]^t [G^{-1}] [\dot{r}] \quad (3.24)$$

où :

$$[G^{-1}] = [B^{-1}]^t [M_\alpha] [B^{-1}] \quad (3.25)$$

L'énergie potentielle s'écrit en fonction des constantes de force harmoniques exprimées dans la base des coordonnées internes :

$$2V = [r]^t \begin{bmatrix} f^{(r)} \end{bmatrix} [r] \quad (3.26)$$

La résolution des $3n - 6$ équations de LAGRANGE revient à diagonaliser le produit matriciel $[G] \begin{bmatrix} f^{(r)} \end{bmatrix}$:

$$[G] \begin{bmatrix} f^{(r)} \end{bmatrix} [L] = [\lambda] [L] \quad (3.27)$$

avec $[L]$: matrice des vecteurs propres.

3.5.3 Espace des coordonnées internes de symétrie

Une coordonnée interne de symétrie (notée s_i) est une combinaison linéaire de coordonnées internes et peut représenter un mode local de vibration. Un mode de vibration peut être constitué à son tour d'une combinaison linéaire de plusieurs modes locaux possédant la même symétrie. Cette propriété est extrêmement importante puisque, dans cet espace, il est possible de déterminer *a priori* les constantes de force harmoniques $f_{ij}^{(s)}$ nulles par simple application des règles de calcul du produit direct issues de la théorie des groupes. Les matrices $[G^{(s)}]$ et $[f^{(s)}]$ ont ici la propriétés d'être bloc-symétriques. Une étude intéressante a été menée par PULAY [67] sur la construction de cet espace en fonction des différents groupements fonctionnels de composés organiques.

3.6 Traitement quantique de la vibration

Que l'approche classique soit menée dans l'espace des coordonnées cartésiennes, internes ou internes de symétrie, la finalité est d'exprimer les coordonnées normales, seules capables de conduire au traitement quantique de l'équation vibrationnelle. Dans cet espace, les deux énergies s'expriment sous forme quadratique :

$$2T = \sum_{i=1}^{3n-6(5)} \dot{Q}_i^2 = [\dot{Q}]^t [\dot{Q}] \quad (3.28)$$

$$2V = \sum_{i=1}^{3n-6(5)} \lambda_i Q_i^2 = [Q]^t [\lambda] [Q] \quad (3.29)$$

et le Hamiltonien correspondant s'écrit :

$$\hat{H} = \hat{T} + \hat{V} = \frac{1}{2} \sum_{i=1}^{3n-6(5)} (\dot{Q}_i^2 + \lambda_i Q_i^2) \quad (3.30)$$

En associant à ces grandeurs leur opérateur correspondant,

$$\hat{Q} = Q \quad (3.31)$$

$$\hat{Q} = \hat{P} = -i\hbar \frac{\partial}{\partial Q} \quad (3.32)$$

nous obtenons l'équation de Schrödinger vibrationnelle :

$$\sum_{i=1}^{3n-6(5)} \frac{\partial^2 \Psi}{\partial Q_i^2} + \frac{2}{\hbar^2} \left(E - \frac{1}{2} \sum_{i=1}^{3n-6(5)} \lambda_i Q_i^2 \right) \Psi = 0 \quad (3.33)$$

Cas de l'oscillateur non dégénéré :

Considérons les séparations de variables suivantes :

$$\Psi(Q_1, Q_2, \dots, Q_{3n-6}) = \Psi_1(Q_1)\Psi_2(Q_2)\dots\Psi_{3n-6}(Q_{3n-6}) \quad (3.34)$$

$$E = E_1 + E_2 + \dots + E_{3n-6} \quad (3.35)$$

L'équation 3.33 revient donc à résoudre $(3n - 6)$ équations à une seule variable :

$$\frac{d^2 \Psi_i(Q_i)}{dQ_i^2} + \frac{2}{\hbar^2} \left(E - \frac{\lambda_i}{2} Q_i^2 \right) \Psi_i(Q_i) = 0 \quad (3.36)$$

Habituellement, les coordonnées normales Q_i sont remplacées par les coordonnées normales sans dimension q_i (à ne pas confondre avec les coordonnées cartésiennes pondérées par les masses définies dans la partie 3.5.1) via l'application de la relation :

$$Q_i = \left(\frac{\hbar^2}{\lambda_i} \right)^{1/4} q_i \quad (3.37)$$

Dans ce système de coordonnées, l'équation de Schrödinger vibrationnelle monodimensionnelle prend la forme bien connue :

$$\frac{d^2\Psi_i(q_i)}{dq_i^2} + \left(\frac{2E_i}{\hbar\lambda^{1/2}} - q_i^2 \right) \Psi_i(q_i) = 0 \quad (3.38)$$

Il existe une infinité de couples (E_i, Ψ_i) solution de cette équation, dont les caractéristiques dans l'espace des coordonnées normales sans dimension sont les suivantes :

$$\Psi_i(q_i) = N_{v_i} H_{v_i}(q_i) e^{-\frac{q_i^2}{2}} \quad (3.39)$$

$$E_i = \int \Psi_i H_i \Psi_i dq_i = hc\varpi_i \left(v_i + \frac{1}{2} \right) \quad (3.40)$$

avec v_i : nombre quantique vibrationnel de la coordonnée q_i , entier positif,
 N_{v_i} : facteur de normation des fonctions d'état vibrationnel (3.41),
 $H_{v_i}(q_i)$: un polynôme d'HERMITE (3.42).

$$N_{v_i} = \left(2^{v_i} v_i! \sqrt{\pi} \right)^{-1/2} \quad (3.41)$$

$$H_{v_i}(q_i) = (-1)^{v_i} e^{q_i^2} \frac{d^{v_i}}{dq_i^{v_i}} \left(e^{-q_i^2} \right) \quad (3.42)$$

Notons de plus que d'après la partie 3.5.1, ces fonctions d'état s'expriment en fonction des coordonnées normales et sont donc de fait orthogonales. Ce traitement est le plus général et peut bien entendu s'appliquer lorsque certaines valeurs propres sont dégénérées. Dans ce cas, nous ne discernons pas de façon explicite un mode à dégénérescence multiple mais de multiples modes à dégénérescence simple de même valeur propre. Nous appelerons ce type de traitement "traitement implicite de la dégénérescence" par opposition au "traitement explicite de la dégénérescence" que nous abordons dans le sous-paragraphe suivant.

3.7 La fonction potentielle anharmonique

Le concept de mode normal de vibration est basé sur l'hypothèse de déplacements infinitésimaux autour de la position d'équilibre. En réalité, les états vibrationnels excités ou les modes mous correspondent à des mouvements de large amplitude. De ce fait, l'expression du potentiel développée dans la partie 3.4 n'est plus suffisante, et les termes d'ordres supérieurs à 2 de la fonction potentielle doivent être pris en compte pour modéliser plus correctement le spectre vibrationnel de la molécule étudiée. Dans l'espace des coordonnées internes curvilignes de

symétrie s_i^2 , la forme analytique de cette fonction devient :

$$\begin{aligned}
 V_v = V_{eq} + & \sum_i^{3n-6} \left(\frac{\partial V}{\partial s_i} \right)_{eq} s_i + \frac{1}{2} \sum_{i \leq j}^{3n-6} \left(\frac{\partial^2 V}{\partial s_i \partial s_j} \right)_{eq} s_i s_j \\
 & + \frac{1}{3} \sum_{i \leq j \leq k}^{3n-6} \left(\frac{\partial^3 V}{\partial s_i \partial s_j \partial s_k} \right)_{eq} s_i s_j s_k \\
 & + \frac{1}{4} \sum_{i \leq j \leq k \leq l}^{3n-6} \left(\frac{\partial^4 V}{\partial s_i \partial s_j \partial s_k \partial s_l} \right)_{eq} s_i s_j s_k s_l \\
 & + \dots
 \end{aligned} \tag{3.43}$$

Les dérivées d'ordre 2, 3 et 4 sont appelées respectivement constantes de force quadratiques, cubiques et quartiques. La troncature de l'expression du potentiel à l'ordre 4 est, selon MASLEN [70], suffisante pour étudier correctement les modes de stretching fortement excités jusqu'à $10\,000\text{ cm}^{-1}$. Les différentes constantes de force sont déterminées, soit classiquement par calcul *ab initio* (ou DFT) de l'énergie moléculaire pour plusieurs configurations nucléaires autour de la position d'équilibre, soit par une procédure de différences finies des dérivées seconde ou premières de l'énergie électronique par rapport aux coordonnées nucléaires.

La fonction potentielle est ensuite exprimée dans l'espace des modes normaux sans dimension de manière à pouvoir construire le Hamiltonien dans cette base. Elle prend alors la forme :

$$\frac{V_v}{hc} = \frac{1}{2!} \sum_i \varpi_i q_i^2 + \frac{1}{3!} \sum_{i,j,k} \phi_{ijk} q_i q_j q_k + \frac{1}{4!} \sum_{i,j,k,l} \phi_{ijkl} q_i q_j q_k q_l \tag{3.44}$$

où ϕ_{ijk} et ϕ_{ijkl} sont les constantes de force cubiques et quartiques exprimées en cm^{-1} . Les relations entre les dérivées d'ordre 3 et 4 de l'équation 3.43 et les ϕ , qui s'obtiennent par les termes de la matrice de passage $[L]$, sont détaillées dans la référence [71]. Lorsqu'il est nécessaire d'expliquer la dégénérescence des coordonnées, nous utiliserons la notation de NIELSEN [72] dans la base des modes normaux sans dimension :

²Dans le cas d'oscillateurs fortement anharmonique, des coordonnées de type SIMONS-PARR-FINLAN [68] ou des coordonnées de type MORSE [69] peuvent être aussi utilisées.

$$V_{pot} = \sum_i^{3n-6} \frac{\omega_i}{2} q_i^2 + \sum_{\|S\|_1=3}^S K_S \prod_{i=1}^{3n-6} q_i^{S_i} \quad (3.45)$$

avec ω_i la fréquence harmonique (en cm^{-1}) associé à la coordonnée q_i . $\|S\|_1$ la somme des éléments du multi-indice $S=(S_1, S_2, \dots, S_{3n-6})$ et S le degré maximal de la PES.

Plusieurs manières permettent de déterminer les valeurs numériques des constantes de force. Il est important de noter préalablement que le nombre de coefficients de la PES est fonction du nombre de vibrateurs et de l'ordre de son développement (même si certains termes peuvent être simplement déterminés par la prise en compte de la symétrie moléculaire). On distingue trois grandes familles de méthodes pour la détermination des constantes de force :

- les méthodes analytiques : elles consistent à rechercher l'expression analytique des dérivées secondes, troisièmes et quatrièmes de l'énergie et à calculer ces dérivées à la configuration d'équilibre.

- Les méthodes numériques : il s'agit ici de calculer la valeur du potentiel pour différentes structures géométriques du système étudié, pour ensuite ajuster une fonction analytique sur cette grille de points ainsi obtenue. L'expression de cette fonction est déterminée par des procédés de régression linéaires. C'est cette approche qui est généralement la plus utilisée, donnant des résultats avec une précision satisfaisante sous condition de disposer d'une redondance d'informations convenables et d'une disposition correcte des points sur le domaine de calcul. En contrepartie, l'effort calculatoire devient très vite gigantesque, amenant l'utilisateur à devoir « dégrader » la qualité la qualité de la méthode de calcul de la fonction d'onde moléculaire pour pouvoir mener à bien l'acquisition de la PES.

- Les méthodes analytiques-numériques : cette approche proposée par Peter Pulay [73] est un compromis entre les deux processus de dérivations précédents. On construit ici par différences finies un champs de force d'ordre n à l'aide des dérivées analytiques d'ordre $n - x$ disponibles. Cette méthode requiert, certes, moins de calculs que l'approche précédente, mais est numériquement très sensible, conduisant parfois à des résultats difficilement exploitables lorsque l'on veut connaître la forme analytique d'un potentiel d'ordre 4 pour des systèmes de grande dimension. Ainsi, une molécule de 10 atomes conduit à 20475 calculs ab initio, nombre que l'on double généralement pour assurer la convergence des résultats.

Indépendamment de la famille de méthodes utilisée pour la détermination des constantes de force il est important de souligner le fait que la dimension du problème augmente de façon vertigineuse avec le nombre d'atome du système à étudier. Il suffit, pour s'en persuader, de rappeler que l'étude d'une molécule de 12 atomes conduit à 46376 calculs *ab initio*, soit plus du double de ceux nécessaire pour le système à 10 atomes précédemment cité. De plus, il est également important de rappeler que chacun des termes (constantes de force) qui constitue l'expression analytique de la PES ne contribue pas avec la même intensité à la description des couplages entre tous les modes. Enfin, compte tenu de la dimension des molécules qui nous a été nécessaire de décrire dans notre étude (de dimensions toutes supérieurs à 20 atomes), il nous est apparu très rapidement nécessaire de trouver un moyen soit, à ne pas avoir à calculer l'ensemble des termes de la PES jusqu'à l'ordre 4 soit, à chercher à ne tenir compte que des termes qui traduisent les couplages principaux entre modes en essayant d'influer le moins possible sur la précision des calculs vibrationnels résultants. Deux pistes sont actuellement en cours de développement au sein de notre équipe. L'une consiste à supprimer de la PES les monômes d'ordre 3 et 4 inférieurs à une constante de force prédéterminée par l'utilisateur. Cette stratégie ‘arbitraire’ a été employée avec succès dans le travail qui sera présenté au chapitre xx sur la famille des acènes. L'autre, développée très récemment dans le cadre d'un stage de master [74], consiste à prendre en compte la notion de ‘proximité énergétique’ afin d'éliminer de la PES les monômes d'ordre 3 et 4 qui couplent les états vibrationnels les plus éloignés énergétiquement. L'un et l'autre de ces critères est intimement lié aux formules énergétiques issues de la méthode standard de perturbation dite de RAYLEIGH-SCHRÖDINGER que nous développerons ci-après et qui permet d'évaluer la correction énergétique apportée par chaque état vibrationnel à l'énergie d'un état donné à travers l'ahnarmonicité mécanique. Le premier critère est directement corrélé à la valeur numérique des constantes de forces K_s envisagées (présentes au numérateur des formules de perturbation) alors que le second critère est, quand à lui, directement corrélé à la différence énergétique des modes couplés par cette constante (terme présent au dénominateur des formules perturbatives). cette observation devrait nous permettre, dans un proche avenir, de disposer d'un outil de contrôle efficace et moins ‘arbitraire’ pour réduire de façon efficace la dimension de la PES utile à la REVS.

3.8 Représentation matricielle du Hamiltonien, calculs d'intégrales

Le premier terme de l'équation (3.45) représente le Hamiltonien d'ordre 0. Ce terme est diagonal. De ce fait, les j ($j \in [1, N]$) fonctions propres normées $\psi_{j,v}^{(0)}$, produits d'oscillateurs harmoniques $\Theta_{j,v}^{(0)}(q_k)$ issues du traitement de l'équation vibrationnelle d'ordre 0 forment un système orthogonal complet :

$$\psi_{j,v}^{(0)} = \prod_{k=1}^{nv} \Theta_{j,v}^{(0)}(q_k) \text{ avec } \langle \Theta_{j,v}^{(0)} | \Theta_{j,v}^{(0)} \rangle = \delta_{v_kv_{k'}} \quad (3.46)$$

Ce système constitue une base de développement des i ($i \in [1, N_s]$) ($N_s \leq N$) fonctions d'ondes vibrationnelles recherchées $|v\rangle_i$ construites comme des combinaisons linéaires de fonctions propres d'ordre 0 :

$$|v\rangle_i = |v_1, v_2, \dots, v_{nv}\rangle_i = \sum_j C_j \psi_{j,v}^{(0)} \quad (3.47)$$

Dans le cas des modes doublement dégénérés, nous noterons la fonction d'onde vibrationnelle associée à l'oscillateur dégénéré i : $|v, l\rangle_i$.

Il s'agit alors de résoudre l'équation de Schrödinger vibrationnelle anharmonique :

$$\hat{H}_{v,l} |v, l\rangle_i = E_{v,l} |v, l\rangle_i \quad (3.48)$$

Ceci se fait habituellement à l'aide de méthodes de perturbation ou de variation, ou bien encore à l'aide de méthodes combinées type variation-perturbation.

Pour résoudre des systèmes d'équations tels que l'équation (3.48), l'outil informatique prend toute sa valeur. Par conséquent, il est commode de résoudre le problème sous sa forme algébrique, c'est-à-dire de le représenter matriciellement. Pour cela, nous définissons la représentation matricielle du Hamiltonien par la projection de ce dernier dans sa base de fonctions harmoniques non dégénérées et doublement dégénérées. Les éléments matriciels $H_{(v,l)(v',l')}$ sont définis par la relation :

$$H_{(v,l)(v',l')} = \int \Psi_{(v,l)}^{(0)} \hat{H} \Psi_{(v',l')}^{(0)} dq_{(i,j,\dots,nv)} \quad (3.49)$$

$$= \langle (v, l)^{(0)} | \hat{H} | (v', l')^{(0)} \rangle \quad (3.50)$$

où \hat{H} est partitionné de la manière suivante :

$$\hat{H} = \hat{H}^0 + \hat{H}^1 + \hat{H}^2 \quad (3.51)$$

\hat{H}^0 se rapporte au Hamiltonien harmonique, \hat{H}^1 et \hat{H}^2 traduisent l'anharmonicité et font intervenir les opérateurs q^3 et q^4 respectivement, ces derniers pouvant être décomposés en un produit d'opérateurs q [75] : $q^3 = \prod_{i,j,k}^3 q_i q_j q_k$, $q^4 = \prod_{i,j,k,l}^4 q_i q_j q_k q_l$. Les variables étant séparables, chaque terme matriciel se décompose en produit d'intégrales monomodes.

Cas des modes non dégénérés

Soit l'intégrale $\langle v_\alpha | q | v'_\alpha \rangle$. Les fonctions monomodes étant dans ce cas caractérisées par des polynômes d'HERMITE, noté ici \mathbf{H} , l'intégrale s'écrit :

$$\langle v | q | v' \rangle = N_v N_{v'} \int e^{-q^2} \mathbf{H}_v q \mathbf{H}_{v'} dq \quad (3.52)$$

On peut montrer qu'il existe des relations de récurrence entre les polynômes, ce qui conduit à [62] :

$$\langle v | q | v' \rangle = N_v N_{v'} \int e^{-q^2} \mathbf{H}_{v'} \left(v \mathbf{H}_{v-1} + \frac{1}{2} \mathbf{H}_{v+1} \right) dq \quad (3.53)$$

Ces polynômes étant orthogonaux, cette intégrale est non nulle si $v' = v \pm 1$ et nous obtenons :

$$\langle v | q | v+1 \rangle = \langle v+1 | q | v \rangle = \sqrt{\frac{v+1}{2}} \quad (3.54)$$

$$\langle v | q | v-1 \rangle = \langle v-1 | q | v \rangle = \sqrt{\frac{v}{2}} \quad (3.55)$$

Comme le développement limité de la fonction potentielle est tronqué à l'ordre 4 dans notre

étude, les opérateurs présents dans le Hamiltonien sont de type p^2, q, q^2, q^3, q^4 . Le résultat des intégrales correspondantes est tabulé [76].

3.9 Méthodes Perturbationnelles

Il existe deux méthodes de calcul perturbationnel : La méthode de “Transformation de contact” issue des travaux de VAN VLECK [77, 78] et la méthode standard dite de RAYLEIGH-SCHRÖDINGER [79]. Cette dernière est relativement simple à mettre en place et conduit aux expressions bien connues de l'énergie :

$$E_{(v,l)}^{Tot} = E_{(v,l)}^{(0)} + H_{(v,l),(v,l)}^{(2)} + \sum_{v \neq v'} \frac{(H_{(v,l),(v',l')}^{(1)})^2}{E_{(v,l)}^{(0)} - E_{(v',l')}^{(0)}} \quad (3.56)$$

$E_v^{(0)}$ et $E_{v'}^{(0)}$ étant respectivement les énergies harmoniques des états v et v' . Le développement de la fonction d'onde est habituellement déduit d'un traitement perturbationnel d'ordre 1 :

$$|v, l\rangle^{(1)} = |v, l\rangle^{(0)} + \sum_{v \neq v'} \frac{H_{(v,l),(v',l')}}{E_{(v,l)}^{(0)} - E_{(v',l')}^{(0)}} * |v', l'\rangle^{(0)} \quad (3.57)$$

Quelle que soit l'approche envisagée, l'expression perturbationnelle de l'énergie totale de vibration peut se formuler de la façon suivante :

$$\begin{aligned} E_{(v,l)} &= \sum_s \omega_s \left(v_s + \frac{1}{2} \right) + \sum_t \omega_t (v_t + 1) + \sum_{s \geq s'} x_{ss'} \left(v_s + \frac{1}{2} \right) \left(v_{s'} + \frac{1}{2} \right) \\ &+ \sum_{s,t} x_{st} \left(v_s + \frac{1}{2} \right) (v_t + 1) + \sum_{t \geq t'} (v_t + 1) (v_{t'} + 1) + \sum_{t \geq t'} g_{tt'} l_t l_{t'} + \dots \end{aligned} \quad (3.58)$$

soit encore³ :

³On pose pour des raisons pratiques : $q_{\pm} = q_x \pm iq_y$

$$\begin{aligned}
E_{(v,l)} = & \sum_s \omega_s \langle v_s | q_s^2 | v_s \rangle + \sum_t \omega_t \sum_l \langle v_t, l | q_t^2 | v_t, l \rangle \\
& + \sum_{s \neq s'} k_{sss's'} \langle v_s | q_s^2 | v_s \rangle \langle v_{s'} | q_{s'}^2 | v_{s'} \rangle + \sum_{s=s'} k_{ssss} \langle v_s | q_s^4 | v_s \rangle \\
& + \sum_{s,t} k_{sstt} \langle v_s | q_s^2 | v_s \rangle \left(\sum_l \langle v_t, l | q_{t+} q_{t-} | v_t, l \rangle \right) \\
& + \sum_{t \neq t'} k_{ttt't'} \left(\sum_l \langle v_t, l | q_{t+} q_{t-} | v_t, l \rangle \right) \left(\sum_{l'} \langle v_{t'}, l' | q_{t'+} q_{t'-} | v_{t'}, l' \rangle \right) \\
& + \sum_{t=t'} k_{tttt} \left(\sum_l \langle v_t, l | q_{t+}^2 q_{t-}^2 | v_t, l \rangle \right) \\
& - \frac{1}{2} \sum_{t=t'} k_{tttl} l_t^2 + \Theta(q^3)
\end{aligned} \tag{3.59}$$

où les indices s et t se rapportent successivement aux états non et doublement dégénérés, x et g sont les constantes d'anharmonicité dépendantes des termes cubiques et quartiques. Si la méthode de perturbation est facile à programmer, et malgré le fait qu'elle soit une des techniques les plus utilisées pour calculer un spectre anharmonique [80], il n'en demeure pas moins que son utilisation reste limitée puisqu'elle s'adresse uniquement, en toute rigueur, au calcul des fréquences les plus basses d'un spectre, zone spectrale où la densité d'états vibrationnels est peu élevée. Cette limitation de la théorie de perturbation est due aux phénomènes de résonance (type FERMI [81] ou bien encore DARLING-DENNISON [82]) et au fait que, tronquant la fonction à l'ordre 4 (termes (semi) diagonaux), il n'est pas possible de prendre en compte tous les termes d'interaction indispensables au traitement des bandes chaudes et de combinaisons.

Des solutions ont ensuite été développées ces dix dernières années afin de résoudre ce problème. Nous pouvons citer, par exemple, la méthode du Hamiltonien vibrationnel à transformée de contact, couplée à une procédure automatisée pour mettre en place et résoudre les résonances pertinentes propres aux systèmes de MARTIN et TAYLOR [83], la théorie de perturbation canonique de grand ordre de VAN VLECK [72] et l'approximation du champ vibrationnel auto-cohérent (VSCF, pour “Vibrational Self-Consistent Field”) qui inclue les corrections concernant la corrélation entre les modes par la théorie de la perturbation développée par JUNG, GERBER et NORRIS. Elle se nomme CC-VSCF pour Correlation Corrected VSCF [84, 85].

3.10 Méthodes Variationnelles

La résolution approchée de l'équation de Schrödinger vibrationnelle par la méthode variationnelle consiste à développer les états propres de H_v dans la base des fonctions propres connues du Hamiltonien d'ordre 0 $H_v^{(0)}$, puis de diagonaliser la représentation matricielle de ce Hamiltonien.

Le problème revient donc à calculer les coefficients C_j des fonctions propres $|v, l\rangle_i$ solutions du Hamiltonien vibrationnel H_v tels que :

- $|v, l\rangle$ soit normé :

$$\langle v, l | v, l \rangle = \sum_j^N C_j^2 \quad (3.60)$$

- l'énergie vibrationnelle :

$$E_{v,l} = \langle v, l | H_{v,l} | v, l \rangle = \sum_j^N C_j^2 \langle \psi_{j,v}^{(0)} | H_{v,l} | \psi_{j,v}^{(0)} \rangle + \sum_j^N \sum_{k \neq j}^N C_j C_k \langle \psi_{j,v}^{(0)} | H_{v,l} | \psi_{j,v}^{(0)} \rangle \quad (3.61)$$

$$E_{v,l} = \sum_j^N C_j^2 H_{jj} + 2 \sum_j^N \sum_{k \neq j}^N C_j C_k H_{jk} \quad (3.62)$$

soit minimale compte-tenu de la condition de normalisation.

- la variation sur l'énergie conduise à l'obtention d'un extremum :

$$dE_{v,l} = \sum_{j=1}^N \frac{\partial E_{v,l}}{\partial C_j} dC_j = 0 \quad (3.63)$$

où

$$\frac{\partial E_{v,l}}{\partial C_j} dC_j = 2C_j H_{jj} + 2 \sum_{k \neq j}^N C_k H_{jk} \quad (3.64)$$

Les variations (∂C_j) ne sont pas indépendantes les unes des autres puisqu'il faut qu'elles vérifient simultanément la condition de normation. La méthode des multiplicateurs de Lagrange permet de prendre en compte ces conditions. L'application du théorème de variation conduit donc à calculer la transformation unitaire (C) qui diagonalise la représentation du Hamiltonien (H) :

$$(H)(C) = E(C) \quad (3.65)$$

dans la base des fonctions propres du Hamiltonien non perturbé afin d'obtenir les valeurs propres – énergies des niveaux vibrationnels – et vecteurs propres – états vibrationnels – correspondants. Les solutions vibrationnelles seront d'autant plus proches des solutions exactes que la base (N) de développement tendra vers l'infini. Cette dernière étant toujours finie et limitée, le problème repose alors sur le choix des configurations de l'espace à diagonaliser.

Les stratégies de sélection sont nombreuses [86, 87, 88, 89, 90, 91] et très proches de celles pratiquées en chimie quantique pour les méthodes permettant d'atteindre l'énergie de corrélation électronique.

Choix de l'espace à diagonaliser

Le choix de l'espace à diagonaliser (N_s) influe grandement sur la précision et la rapidité de convergence des méthodes variationnelles. La troncature de l'espace complet (N) est un problème délicat sur lequel il est nécessaire de s'attarder tant il est à l'origine des particularités singulières des différentes approches.

Pour une meilleure interprétation d'un spectre vibrationnel résonant, il est donc préférable d'employer une méthode variationnelle, même si ce type de méthodes conduit à des écueils connus. En effet, l'ensemble des données extraites à partir de la fonction potentielle V croît radicalement avec la taille de la molécule i.e. proportionnellement au nombre de vibrateur et au terme noté g qui représente le nombre de terme non nul de la PES. On montre que cette croissance se comporte en $O(gN)$ [92]. Or, la précision de la méthode est entièrement dépendante de la qualité de V et de l'obligation de prendre en compte un maximum des N configurations vibrationnelles dans l'espace actif variationnel. La solution recherchée sera d'autant plus proche de la solution exacte que la base N tendra vers l'infini ; or, cette dernière est forcément finie et limitée et le problème réside donc dans le choix des configurations N_s à diagonaliser. Ce point est crucial car de lui dépend grandement la précision et la rapidité de convergence de la méthode variationnelle. La troncature faite sur l'espace N est alors devenue un axe de recherche majeur, qui a conduit à différentes méthodes ayant chacune leurs spécificités.

Ces méthodes de sélection sont très proches de celles employées pour le calcul de l'énergie de corrélation électronique en chimie quantique. Il y a presque autant de stratégies sélectives que de programmes basés sur l'interaction de configuration (CI, pour “Configuration Interaction”). En accord avec les deux algorithmes de WYATT et al [93], nous avons initialement commencé par implémenter une méthode dite P_MWCI (Parallel Multiple Windows CI) basée

sur la construction itérative et parallélisée (γ cycles) de $\langle H_v \rangle$ de l'espace de configuration. Les détails sont reportés dans la référence [94]. Plus récemment, une méthode plus efficace, appelée A-VCI (Adaptative Variational CI) a été mise au point au laboratoire pour sélectionner, non plus la matrice complète du problème, mais une sous-matrice suffisamment représentative du problème pour nous permettre d'être en mesure de calculer les premières valeurs propres d'intérêt à nos analyses (sans jamais avoir à diagonaliser la matrice globale du problème variationnel). Notre analyse est basée sur une méthode hiérarchique apparenté à la méthode de Rayleigh-Ritz variationnelle utilisant une nouvelle façon d'écrire l'erreur résiduelle commise entre deux itérations [95]. Cet algorithme adaptatif a donc été développé afin de corrérer trois conditions simultanément, à savoir, un espace approprié de départ qui ne comporte aucun ‘trou’ énergétique, un critère de convergence contrôlé, et une procédure d'enrichissement de l'espace actif raisonné. L'utilisation à postériori du résidu nous permet de contrôler la/les direction(s) la(es) plus pertinente(s), dans laquelle il est nécessaire d'enrichir l'espace actif de travail.

Ces méthodes ne sont cependant pas uniques et, parmi les différentes méthodes variationnelles développées, nous pouvons citer en exemple le travail innovant de BOWMAN et GERBER relatif à l'approche VSCF [96, 97, 98, 99], au développement du MULTIMODE [100] et à la très efficace approche VCC (“Vibrational Coupled Cluster”) récemment développée par CHRISTIANSEN [101]. Plus récemment encore, une extension de la VSCF par l'introduction de la théorie de la perturbation des états quasi-dégénérés (QDPT, pour “Quasi-Degenerate Perturbation Theory”) a été développée, afin d'améliorer la description de la résonance vibrationnelle dans le cas de molécules polyatomiques [102]. Il est cependant important de rappeler ici que la précision moyenne actuelle concernant les calculs variationnels sur les petits systèmes organiques est de 1 à 15 cm^{-1} , suivant la nature des mouvements étudiés. Celle-ci est à la fois dépendante de la qualité et de la forme analytique choisie pour décrire le champ de forces. Notons enfin que ces approches en cours de développement ne sont destinées, pour le moment, qu'à l'étude de systèmes comportant moins de 10 atomes au total.

3.10.0.1 Les méthodes de variation-perturbation

Face à l'impossibilité de résoudre le problème séculaire de grande dimension associé à une interaction de configuration (CI) complète, nous avons fait le choix dans ce travail d'avoir recours à une méthode de variation-perturbation similaire à celle développée dans le domaine de la spectroscopie électronique CIPSI [103] dont l'équipe ECP s'est fait une spécialité depuis plus

de 20 ans. Dans cette méthode, on cherchera à diagonaliser la représentation du Hamiltonien vibrationnel dans une base construite de façon itérative, l'espace des fonctions générées à partir d'un sous espace initial S_0 verra l'information qu'il contient traité de manière plus approximative par une méthode de perturbation Rayleigh-Schrödinger à l'ordre 2 (sur l'énergie).

Le processus itératif pour l'enrichissement de la base se fait par rapport aux coefficients de fonctions générées qui seront testées dans la correction de la fonction d'onde d'ordre 0 au premier ordre de perturbation. Seules les fonctions qui ont un poids $(C_{jj})^2$ important, défini par l'utilisateur, seront incluses dans la base pour l'itération suivante. Ce choix réside dans le fait que l'on cherche à décrire le mieux possible les fonctions d'onde associées aux états étudiés. On améliore donc la description de la fonction d'onde en lui adjoignant une correction qui tienne compte des excitations de cette dernière, à l'aide d'une méthode de perturbation de Rayleigh-Schrödinger au second ordre. Pour chacun des états la norme de la correction apportée à la fonction initiale est calculée, cette norme permet de suivre l'importance de la correction et de juger si le choix de la base (l'espace générateur) est judicieux. Le choix de la base de départ S_0 va déterminer la qualité du calcul. Il faut nécessairement y inclure toutes les fonctions vibrationnelles (fonctions produit d'oscillateurs harmoniques) que l'on juge importantes pour décrire les états étudiés. Dans la pratique, la dimension de la base peut être de l'ordre de 200 tandis que l'espace total généré peut être de plusieurs milliers ou dizaines de milliers.

Un détail plus précis de la méthode de variation-perturbation est donné ci-après.

Le but de l'algorithme est de construire une base de fonctions vibrationnelles (les fonctions propres $\psi^{(0)}$ de $H^{(0)}$) sur laquelle seront développés les états propres de H ($H = H^{(0)} + H'$)

$$H^{(0)}\psi_i^{(0)} = E_i^{(0)}\psi_i^{(0)} \quad (3.66)$$

avec

$$\psi_i^{(0)} = \prod_{k=1}^{nv} \Theta_{vk}^{(0)}(q_k) \text{ avec } v_k = 0, 1, 2, \dots \quad (3.67)$$

A chaque itération, le processus de génération des fonctions $\psi^{(0)}$ à partir de celles contenues dans S se fait suivant :

$$\langle \psi_i^{(0)} | H' | \psi_k^{(0)} \rangle \text{ avec } \psi_k^{(0)} \notin S_0 \quad (3.68)$$

Soient $(S_m^{(n)}, \psi_{0m}^{(n)}, \eta_m^{(n)}, \psi_{1m}^{(n)}, E_{0m}^{(n)}, E_m^{(n)})$ représentant respectivement le sous-espace S , la fonction d'onde d'ordre 0, le seuil de sélection, la fonction d'ordre 1, l'énergie d'ordre 0, l'énergie corrigée à la $(n)^{ième}$ itération. La diagonalisation de l'Hamiltonien dans le sous-espace $S_m^{(n)}$ nous permettra d'obtenir la fonction d'onde $\psi_{0m}^{(n)}$ associée à l'état étudié, de valeurs propres associée $E_{0m}^{(n)}$:

$$\psi_{0m}^{(n)} = \sum_{i \in S_m^{(n)}} C_i \psi_i^{(0)} \quad (3.69)$$

$$\langle \psi_{0m}^{(n)} | H | \psi_{0m}^{(n)} \rangle \quad (3.70)$$

A l'itération n , l'énergie associée à l'état étudié sera :

$$E_m^{(n)} = E_{0m}^{(n)} + \epsilon_m^{(n)} \quad (3.71)$$

$$\epsilon_m^{(n)} = \sum_{j \notin S_m^{(n)}} \frac{|\langle \psi_i^{(0)} | H' | \psi_j^{(0)} \rangle|^2}{E_{(i)}^{(0)} - E_{(j)}^{(0)}} \quad (3.72)$$

$\epsilon_m^{(n)}$ est la correction due à la perturbation de l'état m , décrit par la fonction d'onde $\psi_{0m}^{(n)}$ par toutes les fonctions générées à partir des fonctions de $S_m^{(n)}$. Ceci donne lieu à la fonction :

$$|\psi_{1m}^{(n)}\rangle = |\psi_{0m}^{(n)}\rangle + \sum_{j \notin S_m^{(n)}} \frac{|\langle \psi_i^{(0)} | H' | \psi_j^{(0)} \rangle|}{E_i^{(0)} - E_j^{(0)}} |\psi_j^{(0)}\rangle = |\psi_{0m}^{(n)}\rangle + \sum_{j \notin S_m^{(n)}} C_{ij}^{(n)} |\psi_j^{(0)}\rangle \quad (3.73)$$

Seules les fonctions générées $|\psi_j^{(0)}\rangle$ dont les coefficients du développement satisfont la condition :

$$|C_{ij}^{(n)}| > \eta_m^{(n)} \quad (3.74)$$

vont contribuer à la correction au premier ordre de la fonction d'onde vibrationnelle. Le sous espace $S_m^{(n)}$ sera enrichi par les fonctions qui ont un poids $(C_{ij}^{(n)})^2$ important (supérieur à un seuil choisi par l'utilisateur) servant à définir ainsi le sous-espace $S_m^{(n+1)}$ qui nous permettra

par diagonalisation, d'engendrer la fonction $\psi_{0m}^{(n+1)}$ et recommencer le cycle, jusqu'à obtenir un résultat stable. Toutefois, notons que cette notion de 'résultat stable' reste délicate à définir comme l'on prouvés nos récents développements concernant la notion de résidu [95]. Néanmoins, de façon générale, la fonction d'onde vibrationnelle s'améliore au cours des itérations, ceci étant réalisé en diminuant le seuil de sélection $\eta_m^{(n)}$, ainsi que le poids $(C_{ij}^{(n)})^2$ des fonctions choisies.

A la fin du processus itératif, on aura donc :

- un sous-espace S judicieusement choisi qui sera traité par une méthode variationnelle. Les fonctions propres obtenues sont de bonnes approximations des solutions vraies pour les états de plus basse énergie.
- un sous-espace généré dont on va évaluer la contribution à l'énergie vibrationnelle en faisant la somme des contributions à la correction de perturbation au $2^{ième}$ ordre de chacun des fonctions qui le composent. Les fonctions du sous-espace générateur doivent interagir faiblement avec ceux de l'espace complémentaire de manière que la fonction d'onde $\psi_{1m}^{(n)}$ soit normalisée ce qui correspond à :

$$\sum_{j \notin S_m^{(n)}} |C_{ij}^{(n)}|^2 \cong 0 \quad (3.75)$$

Chapter 4

Corrections à la Dispersion

Introduction

En mécanique quantique les molécules sont décryptées comme un assemblage de noyaux et d'électrons attaché aux effets classiques et quantiques. Même en considérant les noyaux comme fixes, le comportement des électrons constituent un problème compliqué à N corps. Pour la résolution de l'équation de Schrödinger la théorie de la fonctionnelle de la densité établie une alternative intéressante pour aborder ce problème, grâce à la description du système par sa densité électronique[104].

La DFT a eu un grand succès dans l'étude de systèmes dans lesquels la taille fait impossible un calcul ab initio Post Hartree- Fock, pour la compréhension parmi différents sujets en sciences et technologie.[105, 106, 107]

Cependant cette méthode présente déficiences au moment de traduire la dispersion dans les systèmes et les interactions du van der Waals, voire le théorème Hohenberg- Kohn prédit que si nous connaissons la fonctionnelle correcte, nous pourrions obtenir l'énergie à l'état fondamental en arrivant à son point minimum, donc la difficulté provient des fonctionnelles approximatifs. Local et semi local fonctionnelles ne peuvent pas reproduire le comportement asymptotique de $1/R^6$, comme cela affecte les molécules où il y a une considérable superposition de la densité est inconnue. Il existe une corrélation claire de l'incapacité de décrire les interactions de Van der Waals avec la fonctionnelle d'échange GGA dans la région de faible densité et gradient de densité réduite x.

L'échec de LDA et GGA pour donner l'énergie correcte pour deux densités fixes à longue distances a conduit à des corrections, surtout dans l'étude des interactions intermoléculaire qui définissent le comportement de systèmes complexes en chimie et biologie.

Ce chapitre constitue l'occasion de connaître les bases des corrections disponibles pour les calculs électroniques et qui nous aideront à avoir une approche théorique plus exacte dans les interactions des heterocycles aromatiques.

4.1 DFT-D2

En 2006 Stefan Grimme [108] propose une correction à la dispersion basée en semiempirical GGA ajustement de la correction déjà présenté pour lui même en 2004 [109], dû à l'amélioration de trois points :

- Les coefficients C_6 étaient disponibles que pour les premières atomes du tableau périodique.
- Calculs d'essaie pour molécules avec atomes du troisième période ont montré des erreurs systématiques.
- L'addition de l'énergie de dispersion à l'énergie KS-DFT conduit a des incohérences pour les valeurs normaux en thermochimie.

4.1.1 Théorie

L'approche est basée dans la fonctionnelle de Becke (GGA) [54]. La fonctionnelle B97 est fondée :

$$S_\sigma = \frac{\nabla n_\sigma}{n_\sigma^{4/3}} \quad (4.1)$$

Où n est la densité électronique et σ dénote le spin (α ou β selon la position). La partie de la densité qui dépend de l'échange-corrélation vient donnée par :

$$E_{XC} = E_X + E_{C\alpha\beta} + \sum_\sigma E_{C\sigma\sigma} \quad (4.2)$$

Où X et C sont les contributions d'échange et corrélation respectivement.

L'énergie total est :

$$E_{DFT-D} = E_{KS-DFT} + E_{disp} \quad (4.3)$$

Où E_{KS-DFT} est l'énergie KS usuel obtenue par la fonctionnelle de la densité choisie, et E_{disp} est la correction à la dispersion qui est représentée par :

$$E_{disp} = -s_6 \sum_{i=1}^{N_{at}-1} \sum_{j=i+1}^{N_{at}} \frac{C_6^{ij}}{R_{ij}^6} f_{dmp}(R_{ij}) \quad (4.4)$$

N_{at} est le nombre des atomes dans le système, C_6^{ij} le coefficient de la dispersion pour les atomes paire ij , s_6 est le scaling facteur qui dépende de la fonctionnelle choisie pour le calcul DFT et R_{ij} est la distance interatomique.

Les paramètres C_6 (Jnm^6mol^{-1}) et R_0 (\AA) radio de van der Waals utilisés dans le champ de force semiempirique de Grimme sont tabulé, selon la figure 4.1.

Element	C_6	R_0	Element	C_6	R_0
H	0.14	1.001	K	10.80	1.485
He	0.08	1.012	Ca	10.80	1.474
Li	1.61	0.825	Sc-Zn	10.80	1.562
Be	1.61	1.408	Ga	16.99	1.650
B	3.13	1.485	Ge	17.10	1.727
C	1.75	1.452	As	16.37	1.760
N	1.23	1.397	Se	12.64	1.771
O	0.70	1.342	Br	12.47	1.749
F	0.75	1.287	Kr	12.01	1.727
Ne	0.63	1.243	Rb	24.67	1.628
Na	5.71	1.144	Sr	24.67	1.606
Mg	5.71	1.364	Y-Cd	24.67	1.639
Al	10.79	1.639	In	37.32	1.672
Si	9.23	1.716	Sn	38.71	1.804
P	7.84	1.705	Sb	38.44	1.881
S	5.57	1.683	Te	31.74	1.892
Cl	5.07	1.639	I	31.50	1.892
Ar	4.61	1.595	Xe	29.99	1.881

Figure 4.1: Les paramètres C_6 et R_0

La correction D2 a été employé par Grimme [108] dans systèmes de références comme gaz diatomiques, benzène, anthracène, etc. Les clusters du deuxième groupe du tableau périodique ont donné les plus grands erreurs à cause de la difficile description de systèmes avec cette configuration électronique, laquelle donné des interactions type métalliques.

Park et al [110] étudient l'interaction benzène-benzène dans le cas graphite, ils ont trouvé bon corrélation avec les données expérimentales pour les paramètres de maille.

Plus tard Lee et al [111] rapportent les fréquences de vibration des polymorphes de cellulose I_α et I_β dans le logiciel VASP, cependant le valeur du streching OH calculé est 100 cm^{-1} plus petit que l'expérimental.

4.2 DFT-D3

Les propriétés que présente cette correction par rapport aux versions précédents sont les suivantes :

- Elle est moins empirique, les paramètres les plus importants sont calculés par KS-DFT.
- L'approche est asymptotiquement corrigée pour les fonctionnelles de la densité dans les systèmes limités (molécules) ou systèmes infinites non métalliques.
- Elle fournit une description pour les éléments plus relevant (Z=1-94).
- Les coefficients de la dispersion des atomes paires et leur cutoff radio sont calculés.

L'énergie totale DFT-D3 est donnée par la relation:

$$E_{DFT-D3} = E_{KS-DFT} - E_{disp} \quad (4.5)$$

Où E_{KS-DFT} est l'énergie KS usuel obtenue par la fonctionnelle de la densité choisie, et E_{disp} est la correction à la dispersion qui est représentée comme une somme de l'énergie de deux ou trois corps.

$$E_{disp} = E^{(2)} + E^{(3)} \quad (4.6)$$

Le plus important terme de deux corps expression est :

$$E^{(2)} = \sum_{AB} \sum_{n=6,8,10\dots} s_n \frac{C_n^{AB}}{r_{AB}^n} f_{d,n}(r_{AB}) \quad (4.7)$$

Le premier somme est sur tous les atomes paires du systèmes. C_n^{AB} indique la moyenne de nème ordre des coefficients de la dispersion pour les atomes paires AB, et r_{AB} est la distance internucléaire entre eux. Global s_n ou facteur d'écaillage est ajusté seulement pour $n > 6$ pour assurer l'exactitude asymptotique qui est remplie quand C_6^{AB} est exacte.

C_6 terme n'est pas largement écaillage, les plus hauts termes sont nécessaires afin d'adapter le potentiel pour la fonctionnelle de la densité utilisée dans sa région moyenne. Il a été constaté que les termes $n > 8$ fassent la méthode légèrement instable dans certains situations, pour cela il a été fait une troncature à $n = 8$.

Pour éviter les singularités dans petits r_{AB} et double comptage des effets de corrélation pour distances intermédiaires, la fonction damping $f_{d,n}$ est employée pour déterminer le range de la correction de la dispersion.

Il est utilisé une variante proposée par Chai et Head-Gordon[59] laquelle révèle être numériquement stable et pratique également pour les termes d'ordre supérieur de la dispersion. l'expression de cette fonction est donnée par :

$$f_{d,n}(r_{AB}) = \frac{1}{1 + 6(r_{AB}/(s_{r,n}R_0^{AB}))^{-\alpha_n}} \quad (4.8)$$

Où $s_{r,n}$ est le facteur d'écaillage dépendant du radio de cutoff R_0^{AB} . Il remplace s_6 terme en DFT-D2.

Compare au potentiel DFT-D2 le nouveau est moins contraignant pour les petites distances mais plus attrayante dans la région typique des interactions de van der waals. Il fournit une séparation plus nette entre le court terme et les effets de dispersion à longue portée.

4.2.1 Les coefficients de la dispersion

Au lieu d'utiliser une formule d'interpolation dérivée de manière empirique comme dans DFT-D2, les coefficients de la dispersion sont maintenant calculés par DFT dépendant du temps (TD-DFT), employant des relations de récursion connues pour les termes d'ordre supérieur des multipôles. Le point du départ est la formule de Casimir- Polde[112]

$$C_6^{AB} = \frac{3}{\pi} \int_0^\infty \alpha^A(i\omega) \alpha^B(i\omega) d\omega \quad (4.9)$$

Où $\alpha(i\omega)$ est la moyenne du dipôle de polarisabilité à la fréquence imaginaire ω . Les coefficients d'ordre supérieur sont calculés récursivement par :

$$C_8^{AB} = 3C_6^{AB}\sqrt{Q^A Q^B} \quad (4.10)$$

et

$$C_{10}^{AB} = \frac{49}{40} \frac{(C_8^{AB})^2}{C_6^{AB}} \quad (4.11)$$

Avec

$$Q^A = s_{42} \sqrt{Z^A} \frac{\langle r^4 \rangle r^A}{\langle r^2 \rangle r^A} \quad (4.12)$$

4.2.2 Le terme de Trois Corps

La partie à long terme de l'interaction entre trois atomes à l'état fondamental ne correspond pas exactement aux énergies d'interaction prises par paires. Au meilleur de nos connaissances, nous ne sommes pas conscients de toute considération de cet effet dans une structure modèle. Le premier terme non additif de dispersion dérivé de la théorie des perturbations de troisième ordre pour trois atomes ABC est :

$$E^{ABC} = \frac{C_9^{ABC}(3 \cos \theta_a \cos \theta_b \cos \theta_c + 1)}{(r_{AB} r_{BC} r_{CA})^3} \quad (4.13)$$

Reckien et al [113] ont mis en place dans le logiciel VASP les corrections D2 et D3. Ils ont étudié métaux, oxydes métalliques, cristaux moléculaires comme benzène, et l'adsorption sur surfaces. Tous les deux méthodes montrent qualitatif et quantitatif résultats des paramètres de maille, cependant la méthode D3 proportionne une meilleure performance.

4.3 La méthode Tkatchenko-Scheffler (DFT-TS)

Elle a été proposée en 2009 par Tkatchenko et Scheffler[114]. Nombreux études des structures cristallines ont été menées en utilisant cette correction. Par exemple Kronik et Tkatchenko[115] étudient le cristal d'Hemozoin d'importance biologique, lequel présente des interactions sur les

liaisons hydrogènes. Pour vérifier l'exactitude de la méthode ils sont calculés les paramètres de maille et les modes de vibration du Brushite et le phosphate de calcium hydraté. Bučko et col[116] ont montré que la méthode peut être utilisée dans systèmes ioniques en changeant le calcul de volume effectif et les charges des atomes par la version itératif de la partition de Hirshfeld.

Dans cette méthode la expression pour l'énergie de dispersion est officiellement identique à la méthode DFT-D2, la plus importante différence cependant c'est que les coefficients de dispersion et la fonction damping sont dépendants de la densité de charge. La méthode DFT-TS est capable de prendre en compte les variations à la contribution des interactions de van der Waals des atomes dû à leur environnement.

En effet les coefficients de dispersion, la polarisabilité et les radios atomiques sont calculés à partir des valeurs des atomes libres selon les expressions suivantes :

$$\alpha_i = \nu_i \alpha_i^{free} \quad (4.14)$$

$$C_{6ii} = \nu_i^2 C_{6ii}^{free} \quad (4.15)$$

$$R_{0i} = \left(\frac{\alpha_i}{\alpha_i^{free}} \right)^{\frac{1}{3}} R_{0i}^{free} \quad (4.16)$$

Les paramètres α_i^{free} , C_{6ii}^{free} et R_{0i}^{free} sont tabulés pour tous les éléments des premiers 6 groupes du tableau périodique des éléments, sauf pour les lanthanides. ν_i représente le volume effectif et il est calculé en utilisant la partition de Hirshfeld de la densité des électrons.

$$\nu_i = \frac{\int r^3 w_i(r) n(r) d^3r}{\int n_i^{free}(r) d^3r} \quad (4.17)$$

Où $n(r)$ est la densité totale électronique, n_i^{free} est la moyenne de la densité électronique sphérique des électrons pour les espèces atomiques neutres libres i. $w_i(r)$ est le poids de Hirshfeld et il se définit par les densités atomiques libres.

$$w_i(r) = \frac{n_i^{free}(r)}{\sum_{j=1}^N n_j^{free}(r)} \quad (4.18)$$

L'énergie de dispersion su système AB est égal :

$$E_{dis} = -\frac{1}{2} \sum_{A=1}^N \sum_{B=1}^N \sum_L \frac{C_{6AB}}{R_{AB,L}^6} f_{damp}(R_{AB,L}) \quad (4.19)$$

4.3.1 Self-consistent screening dans la méthode Tkatchenko-Scheffler

Ce variant de la méthode Tkatchenko-Scheffler a été proposé en 2012 par Tkatchenko et col[117] avec l'objectif d'inclure le dépistage des effets à longe portée des polarisabilités effectifs des atomes. Grâce à l'utilisation de l'équation self-consistent screening d'électrodynamique classique.

Cela conduit à une description clair de polarisation et dépolarisation pour le tenseur de polarisabilité des molécules et solides.

Dans cette méthode la fréquence dépendante de la polarisabilité s'obtient la équation self-consistent :

$$\alpha_{iSCS}(\omega) = \alpha_i(\omega) - \alpha_i(\omega) \sum_{i \neq j} \tau_{ij} \alpha_j^{SCS}(\omega) \quad (4.20)$$

où τ_{ij} est le tenseur d'interaction dipôle- dipôle et $\alpha_i(\omega)$ est la fréquence effectif dépendant de la polarisabilité, avec l'expression approximatif :

$$\alpha_i(\omega) = \frac{\alpha_i}{1 + (\omega/\omega_i)^2} \quad (4.21)$$

avec la fréquence d'excitation moyenne caractéristique $\omega_i = \frac{4}{3} \frac{C_{6ii}}{(\alpha_i)^2}$

Les coefficients de la dispersion sont calcules par l'intégral de Casimir- Polder :

$$C_{6ii} = \frac{3}{\pi} \int_0^\infty \alpha_i^{SCS}(\omega) \alpha_i^{SCS}(\omega) d\omega \quad (4.22)$$

Les radios atomiques de van der waals sont calcules par renormalisation du radio obtenu par la méthode de Tkatchenko (DFT-TS).

$$R_{0i}^{SCS} = \left(\frac{\alpha_i^{SCS}}{\alpha_i} \right)^{1/3} R_{0i} \quad (4.23)$$

L'énergie de dispersion est évaluée en utilisant la même équation DFT-TS, avec les paramètres corrigés C_{6ii}^{SCS} , α_i^{SCS} et R_{0i}^{SCS}

Cette méthode est disponible sur le code VASP grâce à Bučko et al.[118] qui étudiaient une large gamme de solides comme gaz nobles, cristaux ioniques et moléculaires, structures type chaînes et métaux. La méthode conduit à des valeurs raisonnablement précises des propriétés structurelles et de cohésion pour divers types de matières solides. Cependant, l'analyse des résultats a montré que les performances ne sont pas aussi bons pour tous les systèmes et il y a certains types de solides où l'approche échoue définitivement. C'est ainsi que pour les cristaux moléculaires le volume calculé est tout à fait exact, l'énergie de cohésion est exacte pour l'azote, mais surestimée dans une certaine mesure pour les autres systèmes. Ils ont montré que les effets de dépistage sont assez petites pour les systèmes constitués par interactions faibles des atomes neutres et molécules. La méthode est moins performante dans les solides ioniques.

Conclusion

Les interactions faibles ou de Van der Waals jouent un important rôle dans le comportement des systèmes en interaction, en raison de cela la compréhension de ces interactions représente un domaine d'intérêt surtout dans les calculs théoriques qui ont donné d'importantes idées aux scientifiques.

Bien qu'il existe de méthodes lesquels prennent en compte la dispersion du système, tels que MP2 ou CCSDT, la taille moléculaire représente une limitant du à la cout de ressources et temps. En conséquence différents approches pour ajouter la dispersion dans les calculs DFT fournissent d'une opportunité pour l'étude de systèmes complexes dont actuellement nous ne connaissons pas vraiment comme ils interagissent.

Sans doute chaque système présente caractéristiques spécifiques pour lesquelles une ou autre correction sera plus appropriée, cependant nombreux travaux donnent les pistes à suivre ou les bases pour la démarche dans l'analyse de molécules à l'état solide.

En définitive la DFT continue apporter un approche ajusté au molécules de plusieurs atomes dans les différents états d'agrégation.

Chapter 5

Modélisation des interactions intermoléculaires

Introduction

Les interactions non-covalentes ou interactions Van der Waals occupent une place prépondérante dans les problématiques actuelles de recherche, celles-ci étant aujourd’hui considérées comme les pierres angulaires de disciplines telles que la chimie supramoléculaire, la science des matériaux ou la biochimie. Après les liaison hydrogènes et les interactions électrostatiques fortes (*e.g.* charge- charge, charge- dipôle, dipôle- dipôle), la plus significatif interaction non covalente est probablement laquelle implique les systèmes aromatiques [119]. Attractives interactions entre cycles aromatiques ont été considérées comme responsables de la stabilité de la double hélice d’ADN et ARN, ainsi comme l’agrégation de Porphyrines [120].

L’énergie d’interaction intermoléculaire d’un ensemble de molécules se situe généralement entre 1 et 20 kcal/mol selon le nombre et le type de molécules impliquées. Cependant, l’énergie d’une liaison chimique covalente se mesure entre 100 et 300 kcal/mol, ce qui est d’un autre ordre de grandeur. De même, la portée d’une liaison chimique dépasse rarement quelques Angströms tandis que les interactions intermoléculaires s’étendent théoriquement jusqu’à l’infini (électrostatique) et de manière pratique entre 2 et 10 Angströms selon la taille et la nature du système.

Malgré toute cette importance pour grandes structures, π -stacking est un phénomène pas bien connu. L’étude expérimentale de ces interactions présente un défi car il est compliqué séparer les interactions π -stacking d’interactions secondaires ou les effets de solvant. A cause de ces difficultés expérimentales, les études en chimique computationnelle apparaissent ainsi qu’un avantage pour comprendre la nature fondamentale des interactions non covalentes, et l’influence sur les systèmes chimiques.

5.1 Bases Théoriques

La modélisation en chimique théorique de la corrélation électronique représente la principale difficulté dans les systèmes des molécules ou bien dans de complexes faisant intervenir des interactions intermoléculaires. Ce phénomène déterminé par l'impossibilité de trouver deux électrons au même endroit de l'espace, apparaît puisque il n'existe pas une solution analytique à l'équation de Schrödinger à des systèmes à plus d'un électron; plusieurs méthodes ont été développées pour traiter des systèmes électroniques à n corps.

La méthode Hartree-Fock, proposée indépendamment par Hartree et Fock en 1930 [121], met chaque électron d'un système dans un champ moyen de façon isolé, dans lequel chaque électron pris subit l'interaction des $n-1$ autres électrons du système. La méthode permet de reproduire le trou de Fermi, par contre ne prend pas compte du trou de Coulomb et il manque de la corrélation électronique, indispensable pour calculer avec la précision nécessaire des énergies et différences. Cela sera ajouté par perturbation, interaction de configurations, random phase approximation ou autres méthodes post Hartree-Fock.

D'autre part la théorie de la fonctionnelle de la densité (DFT) permet de prendre en compte à la fois le trou de Coulomb et le trou de Fermi. Elle modélise un système multi électronique par un système fictif sans interaction électron-électron ayant la même densité que le système réel.

L'énergie d'interaction intermoléculaire est une observable qu'on peut interpréter par différentes décompositions qui ont un sens physique. Buckingham [122], a proposé par exemple une décomposition de l'énergie d'interaction intermoléculaire en quatre grandes contributions : l'électrostatique E_{elec} , l'induction E_{ind} , l'échange-repulsion E_{rep} , et la dispersion E_{disp} .

$$\Delta E = E_{elec} + E_{ind} + E_{rep} + E_{disp} \quad (5.1)$$

L'interaction électrostatique est l'ensemble des interactions coulombiennes des deux densités de charges isolées, est additive et peut être répulsive ou attractive selon l'orientation

relative des molécules. Elle constitue la plus grande partie de l'interaction intermoléculaire. L'énergie d'induction est la stabilisation d'un système par la polarisation des constituants par le champ électrique qu'ils subissent de la part de densités de charge que les avoisinent. Cette énergie est non additive et toujours attractive.

La répulsion vient du principe de Pauli, qui stipule que deux électrons ne peuvent pas occuper le même spin au sein d'une même orbitale. C'est une interaction toujours répulsive et qui apparaît seulement à court distance.

Finalement la dispersion n'a pas d'équivalent classique car c'est une interaction qui est liée à la corrélation électronique de deux densités de charge en interaction. Elle est attractive et existe dans tous les complexes.

Il y a deux manières différentes de calculer une énergie d'interaction intermoléculaire. L'une est la méthode supermoléculaire, et l'autre est la construction d'une interaction par contributions directement résultant des fonctions d'onde des monomères séparés. Dans la méthode supermoléculaire l'énergie d'interaction intermoléculaire est la différence entre l'énergie totale du dimère et la somme de l'énergie totale de chaque monomère.

$$\Delta E = E_{AB} - E_A - E_B \quad (5.2)$$

A et B représentent les monomères et AB le dimère. Dans ce cas-là, une erreur subtile peut se produire autour de l'énergie d'interaction. Cette erreur est connue sous le nom de Basis Set Superposition Error (BSSE) [123]. Si nous calculons les deux monomères dans leurs bases spécifiques, et ensuite un dimère dans l'ensemble des fonctions de base des monomères, nous pourrons utiliser les orbitales virtuelles d'un monomère pour agrandir la base disponible pour la distribution de charge de l'autre monomère et vice versa. Le résultat est donc une augmentation de la qualité de la base pour le dimère vis-à-vis des monomères, et par conséquent une surestimation de l'énergie de l'interaction. Pour corriger l'erreur de BSSE, une méthode possible est de travailler dans une base complète ou saturée pour les monomères et le dimère. Une autre méthode est de calculer l'énergie des monomères dans la base du dimère, ce qui est le plus fréquemment utilisé sous le nom de counterpoise grâce à Boys et Bernardi [124].

Cette correction entraîne que pour chaque distance intermoléculaire considérée il faut calculer l'énergie totale du dimère et des monomères.

L'énergie sans correction donnée pour l'équation 5.2 peut-être modifiée pour l'estimation de la quantité pour laquelle le monomère A est stabilisé artificiellement pour la base supplémentaire du monomère B et vice versa, avec la relation suivante:

$$E_{BASE}(A) = E_A^{AB}(A) - E_A^A(A) \quad (5.3)$$

$$E_{BSSE}(B) = E_B^{AB}(B) - E_B^B(B) \quad (5.4)$$

Où l'exposant désigne la base utilisée, l'indice désigne la géométrie et le symbole entre parenthèses est le système chimique considéré. Ainsi nous soustrayons l'énergie du monomère A et ses bases de l'énergie et bases du dimer, également dans le cas du monomère B.

Pour le moment nous considérons que la géométrie des monomères ne changent pas quand ils s'approchent pour former le complexe. Normalement cette approximation simplifie les calculs et donne de bon résultat.

Le calcul direct d'une énergie d'interaction par la méthode supermoléculaire ne donne aucune information sur la nature de l'interaction. Il y a des méthodes pour décomposer l'énergie d'interaction en différentes contributions ayant en sens physique.

La première théorie mécanique quantique pour interactions intermoléculaire a été développée en 1930 par London et al [125], (la même idée étant dégagée plutôt par Wang [126]). Cette théorie est basée en le standard bas ordre de l'expansion de perturbation de Rayleigh Schrödinger avec l'hamiltonien non perturbé des monomères pas interagissant. La différence entre le totale et l'hamiltonien non perturbé est l'opérateur V d'interaction intermoléculaire, lequel est remplacé par ses multipôles expansions. La méthode de London est validée seule de façon asymptotique pour large séparation intermoléculaire.

Elle peut être améliorée par l'utilisation de l'opérateur V non étendu, l'expansion de per-

turbation résultant est dénommée comme l'approximation de la polarisation. Les composants qui sont présentes dans la polarisation et qui manquent dans la théorie de London sont visés comme l'effet de pénétration.

Autre approche, la méthode de Morokuma, qui est l'une des premières méthodes développées, modifiée plus tard par Kitaura et Morokuma, où l'énergie est divisée au niveau Hartree-Fock dans les composants électrostatique, échange, polarisation et transfert de charge. Ces constituants sont déterminés pour le change de l'énergie totale quand quelques éléments bien définis sont éliminés de la matrice d'interaction de Fock et la matrice de recouvrement. Ils dérivent d'un fonction d'onde qui n'a pas été antisymétrisé [127]. Restricted Virtual Space (RVS), qui a été proposée indépendamment par Stevens et Fink [128], corrige nombreux défauts insatisfaisants de la décomposition de Morokuma en conséquence des termes lesquels n'étaient pas bien corrigés du principe d'exclusion de Pauli. Cependant elle combine leurs termes électrostatique et d'échange et modifie l'évaluation de la polarisation et les composants de transfert de charge [129].

5.2 Interactions intermoléculaires en théorie des perturbations

A partir de la théorie des perturbations de Rayleigh-Schrödinger nous pouvons construire l'énergie d'interaction d'un système, puisque elles sont faibles. L'Hamiltonien du système est la somme de l'hamiltonien de chaque monomère $H^{A/B}$ dont on connaît la solution plus un terme perturbatif V .

$$H = H^A + H^B + V \quad (5.5)$$

$$H = H_0^A + W^A + H_0^B + W^B + V \quad (5.6)$$

Les termes $W^{A/B}$ représentent des perturbations intramoléculaire des monomères, connues par la corrélation normale de chaque fragment, et V représente la perturbation intermoléculaire.

Sans la corrélation nous avons $H_0 = H^A + H^B$ avec $E_0 = E^A + E^B$, où V devient l'interaction

electrostatique entre A et B. Nous écrivons l'équation de Schrödinger :

$$(H_0 - E_0)\psi(\xi) = (E_{int} - \xi V)\psi(\xi) \quad (5.7)$$

Avec le paramètre ξ pour moduler la perturbation. La théorie de Rayleigh-Schrödinger fournit une série de corrections pour l'énergie et la fonction d'onde [130].

$$\psi = \phi_0 + \sum_n \xi^n \phi_{pol}^{(n)} \quad (5.8)$$

$$E_{pol}^{(n)} = \langle \phi_0 | V | \phi_{pol}^{(n)} \rangle \quad (5.9)$$

Le terme "pol" signifie qu'on est dans l'approximation de polarisation, sans effet d'un échange d'électrons entre les monomères dans la fonction $\phi_{pol^{(0)}}$, violant ainsi le principe d'antisymétrie. Les orbitales ne sont ni orthogonalisées, ni relaxées, mais proviennent directement du calcul de monomères séparés.

Pour assurer la convergence de la fonction est employé une asymptote.

$$E_{int} = \sum_{n=1}^N E_{pol}^{(n)} + O(1/R^{2(N+1)}) \quad (5.10)$$

Pour retrouver des termes classiques induction et dispersion, l'énergie en deuxième ordre est encore décomposé en deux parties selon les excitations impliquées, selon la figure 5.2.

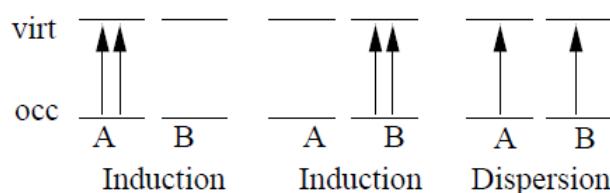


Figure 5.1: Classes d'excitations qui contribuent à l'énergie de perturbation intermoléculaire en 2^{eme} ordre

5.3 La Théorie de Perturbations à Symétrie Adaptée (SAPT)

Cette théorie est un ab initio méthode qui prends en compte les déficiences de la théorie de London et permet de calculer la surface énergie potentiel, grâce à une structure conceptuel pour comprendre le phénomène des interactions intermoléculaire.

Le point du départ est similaire à la théorie de London avec l'hamiltonien non perturbé, mais chacune des corrections à l'énergie d'interaction calculées dans cette approximation peuvent être classifiés tels en décrivant les quatre interactions fondamentales: l'électrostatique, l'induction, l'échange et la dispersion. Ainsi SAPT représente l'énergie d'interactions comme la somme des termes bien définis avec interprétation physique.

Les propriétés de cette approximation ont été étudiées depuis 1960 pour systèmes simples comme H₂⁺. Dans le cas de systèmes multiélectroniques SAPT a fallu appliquer une double perturbation, laquelle a été fait après 1970 [131]. L'hamiltonien du dimère est partitionné dans les contributions de l'opérateur de Fock de chaque monomère (F), l'interaction entre les monomères (V) et le potentiel de fluctuation (W).

$$H = F_A + F_B + V + W_A + W_B \quad (5.11)$$

L'énergie d'interaction peut être décrit comme une série perturbative :

$$E_{int} = \sum_{n=1}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} (E_{pol}^{(nkl)} + E_{exch}^{(nkl)}) \quad (5.12)$$

Le terme n dénote l'ordre en V, k et l dénotent l'ordre en W_A et W_B respectivement. E_{pol} sont de termes qui viennent de l'expansion de la polarisation et E_{exch} sont de termes résultants de l'antisymétrisation de la fonction d'onde par rapport à l'échange des électrons entre les monomères, cela qui manque dans la perturbation de Rayleigh-Schrödinger. Nous pouvons écrire l'énergie RS avec l'opérateur d'antisymétrisation comme :

$$E_{RS}^{(n)} = \frac{1}{\langle \phi_0 | \mathcal{A} \phi_0 \rangle} \left[\langle \phi_0 | V | \mathcal{A} \phi_{pol}^{(n-1)} \rangle - \sum_{\kappa=1}^{n-1} E_{RS}^{\kappa} \langle \phi_0 | \mathcal{A} \phi_{pol}^{(n-\kappa)} \rangle \right] \quad (5.13)$$

Par conséquent les termes d'interaction $E_{pol}^{(n)}$ sont complétés ordre par ordre par des termes d'échange.

$$\begin{aligned} E_{SAPT}^{(1)} &= E_{pol}^{(1)} + E_{exch}^{(1)} \\ E_{SAPT}^{(2)} &= E_{pol}^{(2)} + E_{exch}^{(2)} \end{aligned} \quad (5.14)$$

L'opérateur \mathcal{A} est représenté en pratique pour un seul échange d'électron car toute autre approche est difficile à gérer. La série 5.12 peut être effectuée à divers degrés d'exhaustivité selon la taille du système étudié et la précision souhaitée.

Historiquement plusieurs troncatures de la série ont été faites :

$$E_{SAPT0} = E_{HF} + E_{disp}^{(20)} + E_{exch-disp}^{(20)} \quad (5.15)$$

$$E_{HF} = E_{elst}^{(10)} + E_{exch}^{(10)} + E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)} + \delta E_{HF} \quad (5.16)$$

Où δE_{HF} contient tous les termes de troisième et ordre supérieur en induction et échange – induction.

$$E_{SAPT2} = E_{SAPT0} + E_{elst,resp}^{(12)} + \epsilon_{exch}^{(1)}(2) + {}^t E_{ind}^{(22)} + {}^t E_{exch-ind}^{(22)} \quad (5.17)$$

$$E_{SAPT} = E_{SAPT2} + E_{elst,resp}^{(13)} + [\epsilon_{exch}^{(1)}(CCSD) - \epsilon^{(1)}(2)] + \epsilon_{disp}^{(2)}(2) \quad (5.18)$$

Dans l'équation 5.15 E_{HF} est l'énergie d'interactions de Hartree- Fock, en 5.17 $\epsilon_{exch}^{(1)}(2) = E_{exch}^{(11)} + E_{exch}^{(22)}$, alors que dans l'équation 5.18 $\epsilon_{exch}^{(1)}(CCSD)$ fait référence à la correction de corrélation intramonomère d'échange évaluée au niveau CCSD de théorie. L'indice " resp" montre que la contribution due à la réponse de l'orbitale a été considérée, également l'équation 5.18 est équivalente à l'ordre 4 de la théorie des perturbations supramoléculaire. Ces équations ont été utilisées jusqu'à la moitié des années 90, après les développements des systèmes informatiques ont aidé à l'inclusion de termes d'ordre supérieur dans la formulation.

$$E_{SAPT}^{(30)} = E_{ind}^{(30)} + E_{ind-disp}^{(30)} + E_{disp}^{(30)} + E_{exch-ind}^{(30)} + E_{exch-ind-disp}^{(30)} + E_{exch-disp}^{(30)} \quad (5.19)$$

Pour reduire l'effort numérique, souvent les termes EHF sont remplacés par un calcul Hartree-Fock du dimère, reintroduisant les approximations d'orthogonalisation et couplage entre ordres différents de dispersion et induction par exemple. Au delà du deuxième ordre en V, les termes dispersion et induction et ind-disp perdent un peu leur signification physique et deviennent des grandeurs numériques qui augmentent ou baissent tel ou tel effet au 2^e ordre.

5.3.1 L'interaction Electrostatique

Le premier ordre de l'énergie de polarisation peut être écrit :

$$E_{pol}^{(10)} = \langle \Phi_A^0 \Phi_B^0 | V_{AB} | \Phi_A^0 \Phi_B^0 \rangle \quad (5.20)$$

Pour avoir une représentation physique nous transformons l'équation 5.20 dans :

$$E_{pol}^{(10)} = \int \int \rho_A(r_1) \frac{1}{r_{12}} \rho_B(r_2) dr_1 dr_2 \quad (5.21)$$

Avec $\Phi_{A/B}$ est la fonction d'onde sans perturbée des monomères A et B, ainsi comme $\rho_{A/B}$ est la densité de charge qui s'obtient de l'intégration sur les coordonnées de tous les électrons moins un.

L'équation 5.21 représente l'interaction entre deux distribution de charge, et on peut la nommer comme l'énergie électrostatique $E_{elst}^{(10)}$. Dans le limite de la séparation asymptotique, peut-être représentée comme la somme de l'interaction des moments des multipôles permanents. Alors que dans la région non-asymptotique il contient effet de pénétration de charge.

5.3.2 L'Induction

Si deux molécules en interaction s'approchent, le potentiel électrostatique de la molécule A déforme la distribution de charge de la molécule B et vice versa. Ce phénomène est communément appelé induction. La déformation de la charge va dans le sens du champ électrostatique, et introduit un champ électrostatique dans le sens opposé (principe de Le Châtelier). Le dipôle induit a une interaction favorable avec le champ extérieur, donc l'interaction baisse l'énergie totale. Une induction est toujours attractive.

L'energie de polarisation de deuxième ordre $E_{pol}(2)$ vient donnée par :

$$E_{pol}^{(2)} = -\langle \Phi_0 | V \hat{R} V | \Phi_0 \rangle = - \sum_{m \neq 0} \frac{|\langle \Phi_0 | V \Phi_m \rangle|^2}{E_m - E_0} \quad (5.22)$$

L'energie de l'induction s'obtient quand la somme sur tous les états de la équation 5.22 est limitée aux états excités. Ainsi l'energie totale d'induction :

$$E_{ind}^{(2)} = E_{ind}^{(2)}(A) + E_{ind}^{(2)}(B) \quad (5.23)$$

Où

$$E_{ind}^{(2)}(A) = -\langle \Phi_A | \Omega_B \hat{R}_0^A \Omega_B | \Phi_A \rangle \quad (5.24)$$

Ω_B est l'opérateur du potentiel électrostatique généré par le monomère B sans perturbé.

$$\Omega_B = \sum_{i \in A} \omega_B(r_i) \quad (5.25)$$

$$\omega_B(r_i) = \int \frac{1}{r_{ij}} \rho_B^{tot}(r_j) d^3 r_j \quad (5.26)$$

$\hat{R}_0^A = (H_A - E_A + P_A)^{-1} Q_A$ où $P_A = |\Phi_A\rangle\langle\Phi_A|$ et $Q_A = 1 - P_A$, il est le résolvante réduite de H_A à l'état fondamental.

Dans les calculs de l'énergie d'induction $E_{ind}^{(2)}(A)$ les mouvements des électrons du monomère A n'est pas corrélé avec le mouvement des électrons du monomère B. Cela signifie que la méthode négligent complètement la corrélation électronique intermonomere en ne considérant que l'interaction d'induction.

5.3.3 Dispersion

Ce terme a son origine dans le fait que la distribution de charge n'est pas statique, mais fluctue, même pour les positions des noyaux fixées. Un dipôle (temporaire) induit un dipôle sur l'autre molecule avec la loi d'induction : $E \sim (1/R_{AB})^6$.

L'energie de la dispersion de deuxième ordre $E_{disp}^{(2)}$ est définie par la différence entre l'énergie

de polarisation et l'énergie d'induction, $E_{disp}^{(2)} = E_{pol}^{(2)} - E_{ind}^{(2)}$

En utilisant les expressions 5.22- 5.24, nous écrivons directement la définition :

$$E_{disp}^{(2)} = -\langle \Phi_0 | V R_0^{\hat{A}B} V | \Phi_0 \rangle \quad (5.27)$$

Avec $R_0^{\hat{A}B} = \hat{R}_0 - \hat{R}_0^A P_B - P_A \hat{R}_0^B$ comme la partie de la dispersion du résolvante \hat{R}_0 .

La définition même de l'interaction de dispersion représente un effet de corrélation intermoléculaire pur, elle peut être vue comme l'effet de stabilisation énergétique de la corrélation instantanée des moments des multipôles du monomère Grâce au travail de Casimir et Polder [132] nous savons que l'énergie de l'interaction de dispersion peut être exprimé en termes des polarisabilités multipolaires dynamiques des monomères

5.3.4 Échange

L'opérateur d'antisymétrisation peut être écrit :

$$\mathcal{A} = \frac{N_A! N_B!}{(N_A + N_B)!} \mathcal{A}_{\mathcal{A}} + \mathcal{B}(1 + \mathcal{P}) \quad (5.28)$$

Où $N_{A/B}$ est le nombre des électrons dans les monomères et \mathcal{P} est l'opérateur d'échange des électrons entre les monomères, il peut s'exprimer comme une série :

$$\mathcal{P} = \sum_{i=1}^{\infty} \mathcal{P} \quad (5.29)$$

\mathcal{P} troque $i + 1$ électrons entre les deux monomères.

D'ailleurs le terme d'échange-induction $E_{exch-ind}^{(2)}$ représente l'échange d'électrons entre les monomères quand un est perturbé par le champ électrique statique de l'autre, et le terme d'échange-dispersion $E_{exch-disp}^{(2)}$ représente l'effet électronique d'échange pendant la polarisation mutuel des deux monomères.

5.4 DFT-SAPT

Cette méthode a été développée par Misquitta et al [?], et indépendamment par Heßelmann et Jansen [133], en poursuivant les idées de Williams et Chabalowski [134].

5.4.1 L'Energie Électrostatique et d'Échange

En utilisant les fonctionnelles d'échange-corrélation de la théorie de la fonctionnelle de la densité (DFT), nous pouvons calculer E_{elst}^1 , équation 5.21. Heßelmann et Jansen [133] et Misquita et al [135] rapportent la fonctionnelle PBE0 comme la meilleure adapté pour reproduire l'énergie électrostatique trouvée avec SAPT(HF).

Dans les calculs de l'énergie d'échange $E_{exch}^{(1)}$ où se requiert de matrices de densité d'un et deux électrons, la DFT seulement fournit de matrices d'un électron, cependant $E_{exch}^{(1)}$ est calculée avec un minimum d'erreurs.

5.4.2 L'énergie d'Induction et d'Échange- Induction

Les termes SAPT(HF) d'induction et d'échange-induction $E_{ind}^{(2)}$ et $E_{exch-ind}^{(2)}$ ne prennent pas en compte les changements qui se produisent dans le potentiel de Coulomb ou dans le potentiel d'échange- corrélation due à des changements induits dans la densité électronique. Grâce à l'utilisation de l'approche DFT, ces changements sont pris en compte par l'emploi des fonctions de réponse densité- densité [136]

5.4.3 L'Energie de Dispersion et d'Échange- Dispersion

La fonction de réponse densité- densité est employée dans le calcul du terme de dispersion [137]. l'énergie de dispersion peut être écrite par la expression :

$$E_{disp}^{(2)} \propto \sum_{p \geq q} \sum_{r \geq s} \sum_{t \geq u} \sum_{v \geq w} \int_0^\infty \alpha_{pq,rs}^A(i\omega) \alpha_{tu,vw}^B(i\omega) d\omega \quad (5.30)$$

Où $\alpha(i\omega)$ sont les fréquences de dépendance linéaire de la fonction de réponse.

$$\alpha(i\omega) \propto \sum_p \frac{2\omega_p}{\omega^2 + \omega_p^2} \quad (5.31)$$

ω_p sont les autovaleurs du produit de deux matrices Hessians calculés en utilisant l'approximation adiabatique de la densité locale (Adiabatic local density ALDA)[138] dans le noyau d'échange- corrélation avec DFT dépendant du temps (TDDFT). Dans les valeurs de l'énergie de dispersion cette approximation a montré être en concordance avec les résultats du SAPT(HF).

5.5 Méthodes Basées sur la DFT

La théorie de la fonctionnelle de la densité DFT est largement utilisée dans les calculs de l'interaction intermoléculaire des systèmes aromatiques, néanmoins il manque toujours dans les fonctionnelles employé dans la procédure Kohn Sham la contribution à la dispersion. Pendant que les calculs avec fonctionnelles locales d'échange-corrélation dirigent à surestimé l'énergie d'interaction, non local fonctionnelles d'échange- corrélation souvent sous-estimé l'énergie [139].

Une approche pour décrire correctement la dispersion en DFT est l'ajout d'un terme interatomique de type Lennard-Jones en C_6/R^6 , modulé par une fonction d'amortissement empirique. Un autre développement dans l'amélioration de la DFT a été proposé par l'utilisation de fonctionnelles double-hybrides, où une partie de l'échange et une partie de la corrélation sont calculées par MP2 à l'aide de coefficients empiriques.

5.6 CCSD(T)

Actuellement la méthode d'agrégats couplés (Coupled Cluster) de mono- et diexcitations avec des excitations triples par perturbation est devenue un standard pour calculer des énergies de corrélation dû à l'inaccessibilité du Full CI pour les systèmes de moyen taille. Elle est nommée comme l'étalon d'or de la chimie quantique.

Si nous utilisons une base assez grande d'orbitaux atomiques, la séquence suivante est généralement validée : Full CI > CCSDTQ > CCSDT(Q) > CCSDT > CCSD(T) > CCSD.

CCSD(T) est particulièrement utile pour l'étude des interactions non-covalents où l'inclusion des termes de triple excitation sont nécessaires pour avoir une précision satisfaisante, ainsi l'emploi de CCSD détériore fortement la justesse des résultats [140].

La référence pour le calcul d'énergies d'interactions intermoléculaires est ainsi CCSD(T) en base complète, *i.e.* CCSD(T)/CBS. Pour atteindre l'énergie correspondant à la base complète,

une méthode qui consiste à calculer les énergies en base double, triple, quadruple zéta puis extrapoler par une formule empirique l'énergie pour une base infinie. Riley et al [141] ont montré que cette méthode est précise dans les différents types d'interactions non-covalent, cela qui est probablement pour la compensation des excitations d'ordre supérieur qui ne sont pris en compte dans la CCSD(T) seul.

5.7 Dimère du Benzène

L'interaction intermoléculaire du dimère de benzène a été largement étudiée par méthodes expérimentales et théoriques, puisque il représente un prototype pour les interactions $\pi - \pi$ stacking.

L'énergie de liaison du dimère obtenu à partir de l'énergie de dissociation de l'ion et les potentiel de ionisation du dimère et du monomère ont été mesuré par Krause et al [142] $D_0 = 1.6 \pm 0.2 \text{ kcal mol}^{-1}$ et par Grover et al [143] comme $2.4 \pm 0.4 \text{ kcal mol}^{-1}$. La plupart des études discutent trois configurations de prototypes, qui sont affiches dans la figure 5.7.

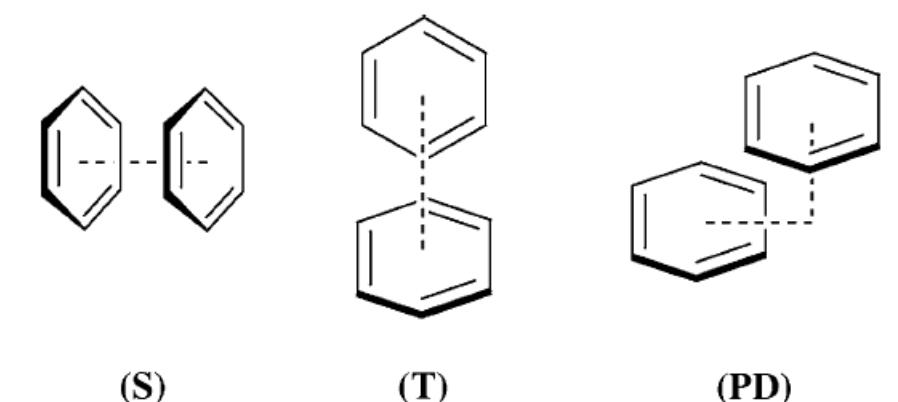


Figure 5.2: Structures des dimères de Benzène : (S) Sandwich, (T) En forme de T et (PD) Parallèle Déplacé

La stabilité relative entre la configuration en forme de T et la parallèle déplacée a été un sujet de nombreux études.

La configuration en sandwich ayant une superposition maximal, peut apparaître comme la plus favorable pour maximiser les interactions de dispersion, mais elle est rarement observée dans des systèmes contenant des noyaux phényles, alors la configuration parallèle déplacée est

souvent observée dans les structures cristallines de composés aromatiques [144, 145, 146] ou dans les interactions des chaînes latérales aromatiques dedans les protéines [144, 147].

Klemperer et al [148] rapportent la configuration en forme de T comme prédominante dans leur étude de la molécule en phase gaz, ensuite du spectre rotationnel d’Arunan et Gutowsky [149], en concordance avec Henson et al [150] qui rapportent par un étude de spectroscopie Raman que le dimère du benzène présente deux non équivalents monomères, un de bas et l’autre de symétrie élevée.

Les calculs ab initio indiquent que les configurations (T) et (PD) présentent une barrière énergétique minimal de 0.1 kcal/mol [151]. Dans la littérature il existe considérables données pour l’énergie d’interaction du dimère de benzène, le tableau 5.1 fournit un petit résumé avec les valeurs obtenue par CCSD(T), en ce qui concerne à cette thèse, nous faisons attention aux interactions dans systèmes qui se présentent à l’état solide, par conséquent nous évaluons la méthode DFT-SAPT sur le système du benzène avec la configuration (PD), puisque cela expose des interaction π -stacking.

Table 5.1: Energie d’interaction du dimère de Benzène avec CCSD(T) en kcal/mol et DFT-SAPT (PBE0)

	T	PD
Park et Lee ¹	-2,67	-3,03
Tsuzuki et al ²	-2,46	-2,48
Sinnokrot et al ³	-2,74	-2,78
Hobza et al ⁴	-2,17	-2,01
DFT-SAPT (PBE0)		-2,35

¹ref [151], ²ref [152], ³ref [153], ⁴ref [154]

La méthode DFT-SAPT donnée un valeur qui corresponde assez bien avec les résultats obtenus par CCSD(T) avec différents conditions de calculs. Pour cela nous proposons cette méthode pour l’étude des interaction dans les heterocycles aromatiques d’intérêt pour ce travail.

Chapter 6

**Molecular dynamics study of the
nano-aggregation in asphaltene
mixtures: effects of the N, O and S
heteroatoms**

We present molecular dynamics simulations (*MDS*) for interpreting the molecular aggregation of:

1. six different asphaltene molecular models based on the C5Pe molecule;
2. several other derivatives based on the PA3 and CA22 structure issued from the recent work of isolation and identification of asphaltenes by Schuler et al.¹

These simulations are based on recent X-ray (*SAXS*) and neutron (*SANS*) small-angle scattering experiments from Eyssautier et al.² which proposed a discoidal asphaltene nanoaggregate structure of 0.67 nm height and its chemical composition. For the first topic, we have investigated the effect of sulfur atom position in the asphaltene structure and we have found that when it is grafted directly to the conjugated core no effect can be deducted from the stability of dimers and trimers and all the other studied parameters. This is not the case when sulfur is located in the lateral chain, where it has a particular affinity with oxygen and hydrogen atoms of the acid chain ends of neighbor molecules. Consequently, we suggest that these configurations of the sulfur atom are more likely to be problematic in oil industry. For the second topic, the chemical characteristics of these real asphaltene molecules were screened by changing the heteroatoms on the backbone and the lateral chain-ends. The results show that the interaction energies vary for different heteroatom arrangement within a given structure and depend on the type of asphaltene. Moreover, we showed that the chain-ends have a crucial role on this phenomenon.

¹ *JACS*, **2015**, 137, 31, 9870

² *The J. of Phys. Chem. B*, **2011**, 115, 21, 6827

6.1 Introduction

Asphaltenes are the oil's insoluble phase in light *n*-alkanes and soluble in aromatic solvents like benzene or toluene[155]. Their presence is critical and can cause problems in several fields of petroleum utilization, from oil recovery and transportation to refining.[156] These complications arise from the tendency of asphaltenes to aggregate and form macro structures able to modify their flow properties (increasing oil viscosity, emulsion stability, etc.) and even, in the worst cases, stop the fluid production (plugging pores in the reservoir, well-bore tubing or flow-lines/pipelines, accumulating in separators, etc.) [157].

Asphaltenes consist of a complex mixture of polycyclic aromatic hydrocarbons substituted with alkyl side chains. Carbon and hydrogen are the major atoms, but the molecules also contain heteroatoms, mainly sulfur.[158] Although much knowledge has been gained throughout the last years, the molecular structure of this phase is poorly determined and can vary quite significantly depending on the origin, on the age of the crude oil and on the pressure and temperature conditions[159].

So, in order to fully understand their nature and to predict their structure, several analytical techniques have been used to gather information on their macromolecular structure. Small angle X-ray (SAXS) and neutron scattering (SANS) techniques have proved that the basics of their supramolecular arrangement is the discoidal π -stacking of the conjugated cores, with heights varying from 1 to 10 nm [160, 161].

From *in silico* point of view, the study of asphaltene aggregation makes use of molecular dynamics simulations. Several force-fields have proved to produce results consistent with experimental results.[162, 163, 164] Thereby, we have, in a first moment, chosen six asphaltene molecular models, as dimers and trimers in toluene, to investigate the sulfur atom position in the structure and the consequent effect on the stability of dimers and trimers in solution. These models are based on the C5Pe model module, a prototypical asphaltene molecule that has been reported and used in several studies[164, 165].

Experimental work in progress using this molecule also justifies this choice that allows us to

compare experience and theory. This work aims to investigate the influence of sulfur atoms on C5Pe's dimers and trimers stabilization during a 150 ns time-delay in toluene at 298K and 1 atm conditions. In this way, all the other possible parameters influencing the nano-aggregation of asphaltenes (size of conjugated core, lateral chains, etc...) are excluded and one can really screen the effect of sulfur atoms in this physical-chemical process.

Then, the first part of this study uses the N-(1-Hexylheptyl)-N'-(5-carboxy-lipentyl)-perylene-3,4,9,10-tetracarboxylicbisimide molecule (C5Pe, molecule 1) and three C5Pe analogues (molecules 2, 3, and 4), with the presence of sulfur atom at different positions (core and side chain) as test cases. Two other molecules, 5 and 6, respectively, were studied: the situation where all the C=O groups are replaced by C=S ones (the extreme of molecule 2) and the situation where no heteroatom is found in this position (C—H bond, quinoid structure imposed).

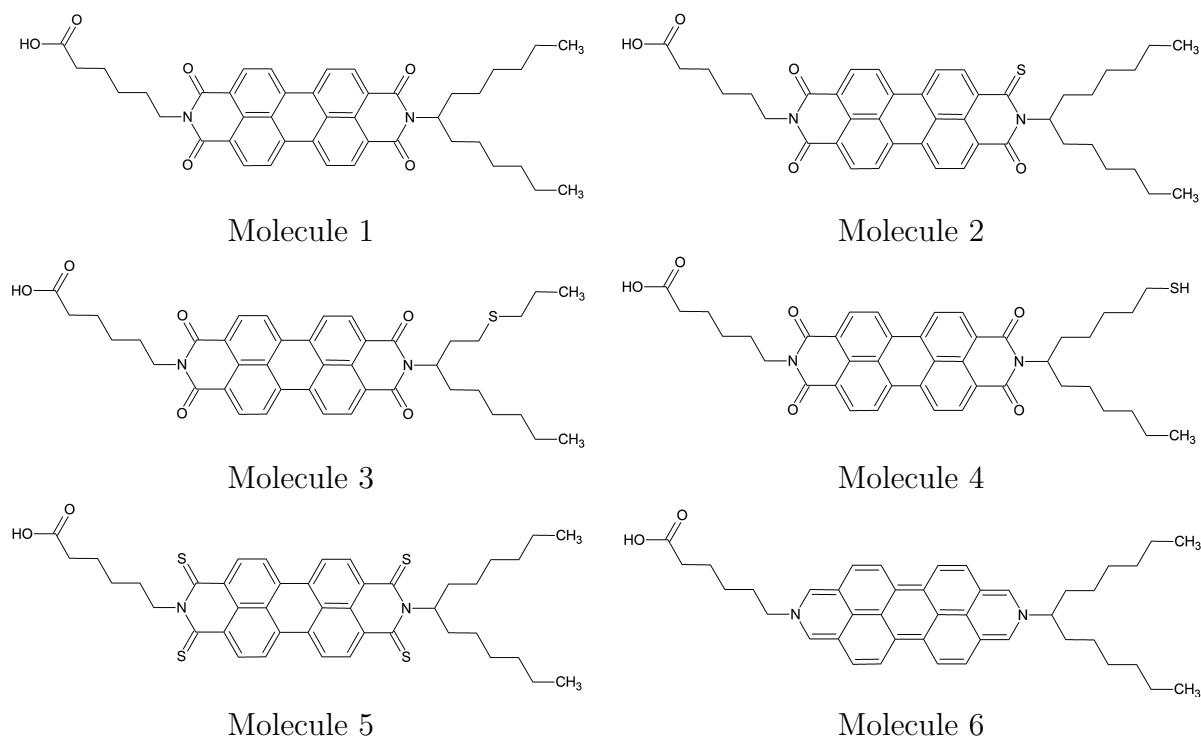


Figure 6.1: C5Pe molecular models of asphaltenes.

For each solvated asphaltene system using the C5Pe model, the initial configuration was constructed in a box with the 2 or 3 asphaltene molecules models. Two starting points were envisaged: 1) their poly-aromatic cores are anti-parallel to each other and the aggregate is already formed (this is called the “organized” configuration); and 2), the molecules were

randomly inserted in the simulation box. For the former case, the spacing between the mass centers of two neighboring asphaltene molecules is 0.35 nm in the -stacking direction. Finally, for both cases, the box was randomly filled with toluene molecules.

More recently, Schuler *et al.*'s[166] work brought more evidences on the molecular structures of asphaltenes based on Atomic Force Microscopy (AFM). They report the isolation of several distinct molecules which are in good agreement with the commonly accepted asphaltene structure: basically composed of aromatic cores, substituted by heteroatoms and presence of alkyl side chains with a variable number of carbons. We have chosen two among these molecules in order to study the effect of the heteroatom substitution on the conjugated aromatic core and the mixture in the crude oil and how the aggregation can occur in a more realistic model than just using one molecule at a time. The selected molecules, depicted in Figure 6.1, are originally labeled PA3 and CA22 after Schuler *et al.*,[166] both of them being of the continental type.[158, 155, 167] Finally, it is worth saying that the CA22 molecule is a coal asphaltene rather than a petroleum asphaltene. It is herein used since it has a conjugated core size sensibly different from the PA3 one.

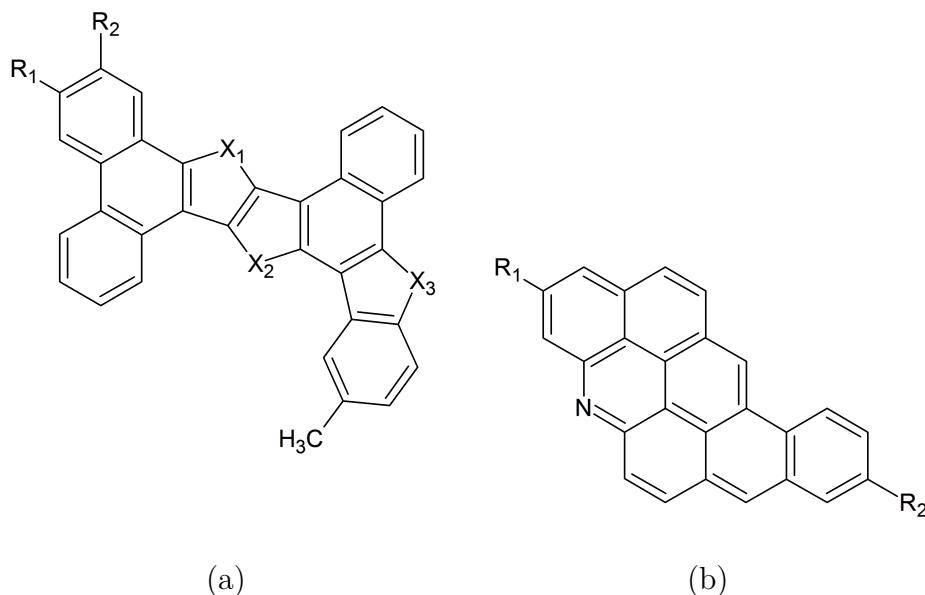


Figure 6.2: Two different asphaltene structures found in the crude petroleum: (a) PA3 and (b) CA22. R_1 and R_2 are the *n*-hexyl lateral chains which can have $-COOH$ or $-CH_3$ as chain-ends. X_n ($n=1-3$) can be O, S, or NH atom groups.

These molecules had their lateral chain-end varied as well as the heteroatom on their conjugated cores. This allowed us to have 14 different molecular structures that were studied separately and together in different proportions. In this way, the aim of this paper is to study the first steps of the asphaltene nano-aggregation by molecular dynamics simulations (MD) as a function of their chemical composition. For PA3, the choice of the heteroatom was motivated by previous Fourier transform ion cyclotron resonance mass spectroscopy (FT-ICR-MS) results indicating their presence in the aromatic core.[168, 169] In the case of CA22, nitrogen was chosen because of the relatively high stability of 6-atoms cycle compared to what can be obtained using the other heteroatoms. These molecules were declined following the tabulation in Table 6.1. They are classified following their number of carbon atoms (NbC), their Double Bond Equivalent (DBE) and the Class to which they belong accordingly to FT-ICR-MS, [170, 171, 172] which indexes might be of use for further referring.

Table 6.1: Labelling of the systems used in this study, with the chain-ends of the lateral chains. Molecules A11-A53 are of the type PA3 and AAA, AAB and AAC are of the type CA22. NbC is the number of Carbon atoms of the structure, DBE is the Double Bond Equivalent and Class stands for the nomenclature reporting the number of heteroatoms found in the structure by FT-ICR-MS. M_w is the molecular weight of each molecule.

Label	X_1	X_2	X_3	R_1	R_2	NbC	DBE	Class	M_w (g.mol ⁻¹)
A11	O	O	O	COOH	COOH	45	27	O ₇	690.8
A12	O	O	O	CH ₃	COOH	45	26	O ₅	660.8
A13	O	O	O	CH ₃	CH ₃	45	25	O ₃	630.8
A21	S	S	S	COOH	COOH	45	27	S ₃ O ₄	740.0
A23	S	S	S	CH ₃	CH ₃	45	25	S ₃	679.0
A31	NH	NH	NH	COOH	COOH	45	27	N ₃ O ₄	687.8
A33	NH	NH	NH	CH ₃	CH ₃	45	25	N ₃	627.8
A41	O	O	S	COOH	COOH	45	27	O ₆ S	706.8
A43	O	O	S	CH ₃	CH ₃	45	25	O ₂ S	646.8
A51	O	S	S	COOH	COOH	45	27	O ₅ S ₂	723.9
A53	O	S	S	CH ₃	CH ₃	45	25	OS ₂	662.9
AAA	-	-	-	COOH	COOH	37	22	N ₁ O ₄	555.6
AAB	-	-	-	COOH	CH ₃	37	21	N ₁ O ₂	525.7
AAC	-	-	-	CH ₃	CH ₃	37	20	N ₁	501.7

The DBE in this table was calculated as:

$$DBE = C - \frac{H}{2} + \frac{N}{2} + 1 \quad (6.1)$$

Where C, H and N stand respectively for the number of carbon, hydrogen and nitrogen atoms. The labeling system relies on the following logics: the first “A” stands for “asphaltene”,

the second character identifies a different family of molecules: numeric for each type of different molecule within the PA3 family and alphabetic for each type of different molecule within CA22 family. The last character identifies the number of acid chain ends: index 1 for two acid chain ends $-COOH$, index 2 for one acid chain $-COOH$ end and one $-CH_3$ and index 3 for two $-CH_3$ chain ends (no acid).

For this ensemble of molecules, three distinct scenarios were studied:

1. homogeneous boxes - boxes where each asphaltene molecule interacts with its equal;
2. heterogeneous boxes - boxes where asphaltenes of the same class (PA3 or CA22) are studied among them;
3. *super-heterogeneous* boxes - boxes where at least one asphaltene molecule is of type PA3 and at least one is of type CA22;

This variation allowed us to span a large variety of mixing scenarios and to separate the different contributions of each physical-chemical factor having a role on the aggregation of asphaltenes. Finally, it is worthy saying that this work is a follow-up of the one recently published by our group,[173] in which we have studied the influence of the heteroatom position in a model asphaltene molecule.

However, all the studies available in literature very often take into account the presence of only one type of asphaltene at a time. This approach is not completely representative of the real very complex mixture that is the crude oil and the effect of rationally changing the chemical structure on the aggregation behavior of asphaltenes has, up to our knowledge, never been reported. Some authors have indeed studied the presence of different asphaltenes in the same simulation box but their interest was, mainly, the distribution of each type of molecules across the oil-water interface.[174, 175, 176]

6.2 Simulation Method

The MD simulations were carried out using the GROMACS 5.1.1 software package.[177, 178, 179] The GROMOS96 force field[180] with the 53a6 parameter set[181] was used in all calculations. This force field has already been tested on poly-aromatic molecules and has shown to be very useful in the dynamics of this type of system.[165, 182, 183, 184]

Toluene was used as solvent in order to mimic the infinite dissolution effect and to avoid the formation of the aggregates due to polarity difference. The initial coordinates were used as input to the PRODRG 2.5 server[185] to generate the GROMACS molecular topology and structure files. The topology of toluene molecule was generated from phenylalanine amino acid fraction in GROMACS. This method proved to be precise since it yields a calculated toluene density value of 0.88 g.cm^{-3} , in good agreement with the 0.86 g/cm^{-3} value extracted from NIST WebBook (National Institute of Standards and Technology).[186]

The resulting topologies had their atomic charges adjusted in order to better reproduce the ones obtained by molecular quantum-mechanics methods, such as Density Functional Theory (DFT).

The C5Pe derivatives had all partial charges for the compounds were used in the topologies based on the GROMOS96 force field parameters. For C=S group, the charges were applied as described by Huang and co-workers.[187] For the molecules based on Schuler *et al.*'s work, the Mulliken charges of the heteroatoms and their surroundings were obtained at the B3LYP/def2-SVP level of theory [49, 188, 189, 190, 191] using Orca 3.0.3 software[192].

The first part of this study was divided into the study of dimers and trimers; separately, at a fixed concentration. For dimers, cubic boxes of $50 \times 50 \times 50 \text{ \AA}$ were either filled randomly with the C5Pe-derivative molecules (random configuration) or filled with two molecules having their poly-aromatic cores are anti-parallel to each other and the aggregate is already formed (this is called the “organized” configuration), as already mentioned. For the latter case, the spacing between the mass centers of two neighboring asphaltene molecules is 0.35 nm in the π -stacking direction. Finally, for both cases, the box was randomly filled with toluene molecules. For trimers, so that the concentration could be kept constant, the simulation boxes now had $50 \times 75 \times 50 \text{ \AA}$ dimensions. The second part of this study, using the molecules reported by Schuler *et al.*, a cubic box of $50 \times 50 \times 50 \text{ \AA}$ was randomly filled with 5 molecules of asphaltene and the solvent was then randomly added, as well.

Once the several distinct configurations of starting geometries were prepared, their energy was minimized by the steepest descent and conjugated gradients methods. Then, the systems were equilibrated in the NPT ensemble at 298 K and 1 bar pressure during 3 ns using the

V-rescale thermostat[193]³ and Berendsen barostat.[194] The production phase consisted of 60 ns of dynamics⁴ at 298 K and 1 bar of pressure also under the NPT ensemble but using the Nosé-Hoover thermostat[195, 196] and the Parrinello-Rahman[197] pressure coupling algorithm. The initial atomic velocities were set using the Maxwell-Boltzmann distribution of the specified temperature. Throughout minimization, stabilization and production phases, a cutoff of 12 Å was used for both Coulomb and Van der Waals interactions. Periodic boundary conditions were thoroughly applied in the three directions. The Verlet integration algorithm[198] has been used with a 2 fs integration step. The electrostatic interactions were computed using the Particle-Mesh Ewald[199] summation method with a fast Fourier transform grid spacing of 1.6 Å to account for long-range electrostatic interactions of the system. All the bond lengths have been constrained using the LINCS algorithm[200] and this is based on the fact that the bond vibration periods are orders of magnitude quicker than the aggregation process. The neighbor list within 12 Å radius was updated every 5 steps.

After minimization, equilibration and production phases, each system was analyzed by calculating the radial distribution function (RDF). These functions are calculated considering the distribution of the distances between the center of mass of the selected residues. In our case, the asphaltene residue is consisted of the aromatic core and the lateral chain. In order to exclude the self-interaction (intra-molecular) configuration, the RDFs were calculated within the interval comprised between 2 and 25 Å. Moreover, the geometry of the poly-aromatic cores was also quantified by the cosine of angle between the poly-aromatic planes.

6.3 Results and Discussion

The presentation of the results of both works is done separately in the next sections.

6.3.1 C5Pe-derivatives systems

Molecule 1 was already studied by MDS by Teklebrhan and co-workers[184, 165] and Gao and co-workers,[164] with results similar to the ones found in this study, *i.e.*, the dimers of these

³For the systems using the C5Pe derivatives, the Berendsen thermostat was used for convenience reasons and this has no impact on the results of the calculations.

⁴150 ns for the C5Pe derivatives systems.

molecules are stable and keep interacting throughout the simulation. Figures 6.3, 6.4, 6.5 and 6.6 show the initial (0 ns) and the final configurations (150 ns) of all systems for which the ordered or random starting point was used. Visual inspection of the MDS trajectories indicates that all systems keep their aggregated stack in place throughout the simulation, although some of them, such as 3T, has a final configuration where only two of the asphaltene molecules are in an aggregated state whereas the third is not in interaction with the stack. This is not found for the case where the starting configuration is random, as it can be seen in Figure 6.6. Further details why this happens are given in the next sections. Thus, the behavior of both dimers and trimers are explained in more details in the next sections.

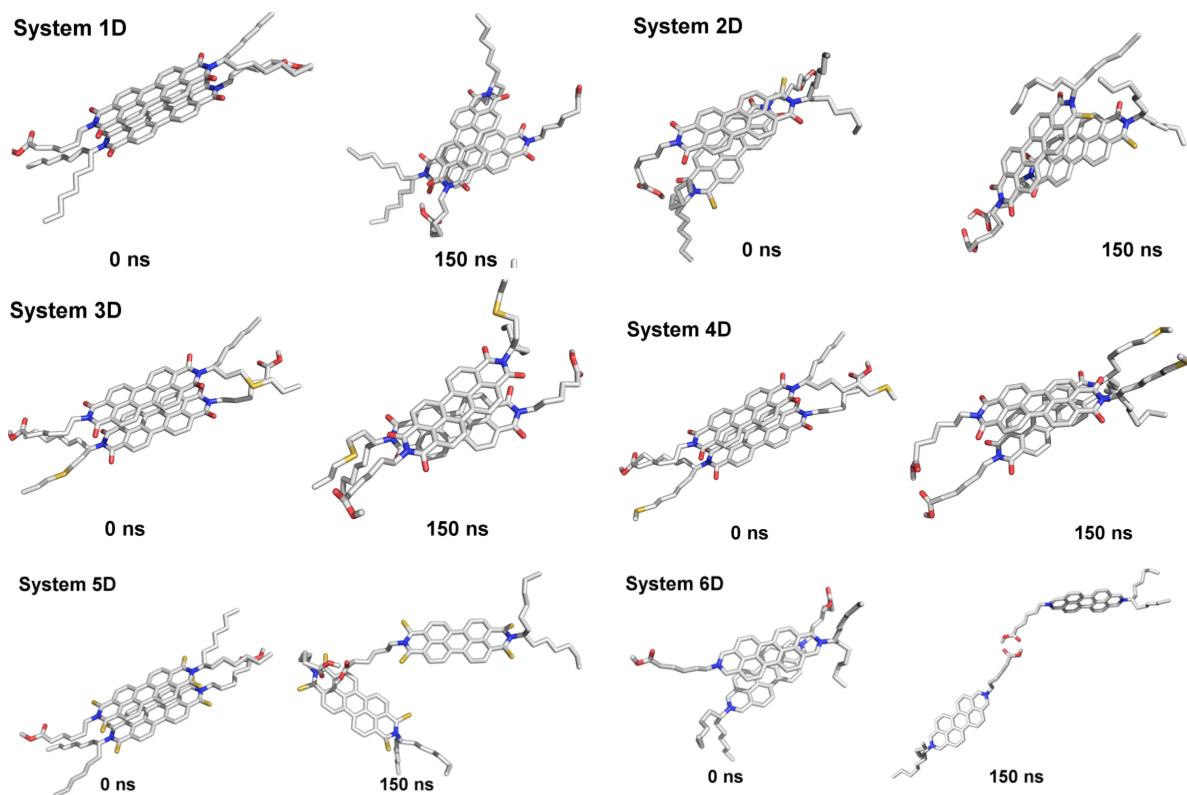


Figure 6.3: Snapshots of the initial configuration and the final frame of the dimers simulations. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.

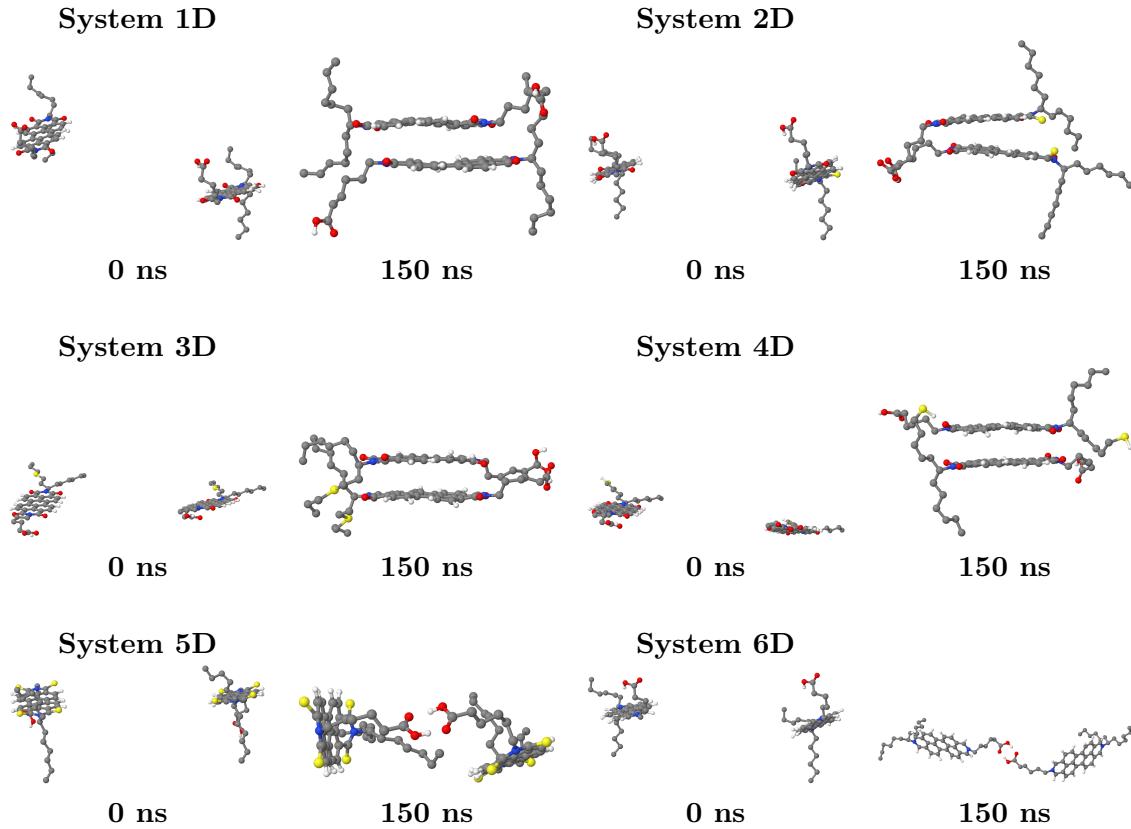


Figure 6.4: Snapshots of the initial configuration and the final frame of the dimers simulations for the random disposition case. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.

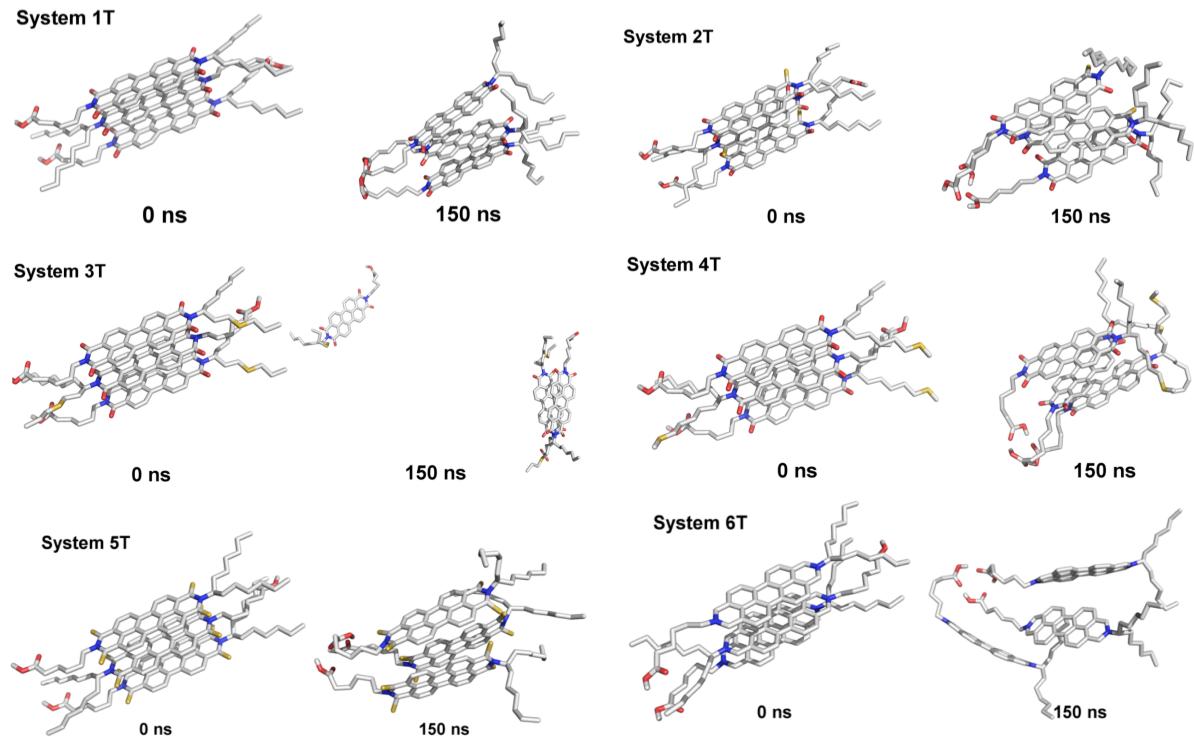


Figure 6.5: Snapshots of the initial configuration and the final frame of the trimers simulations. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.

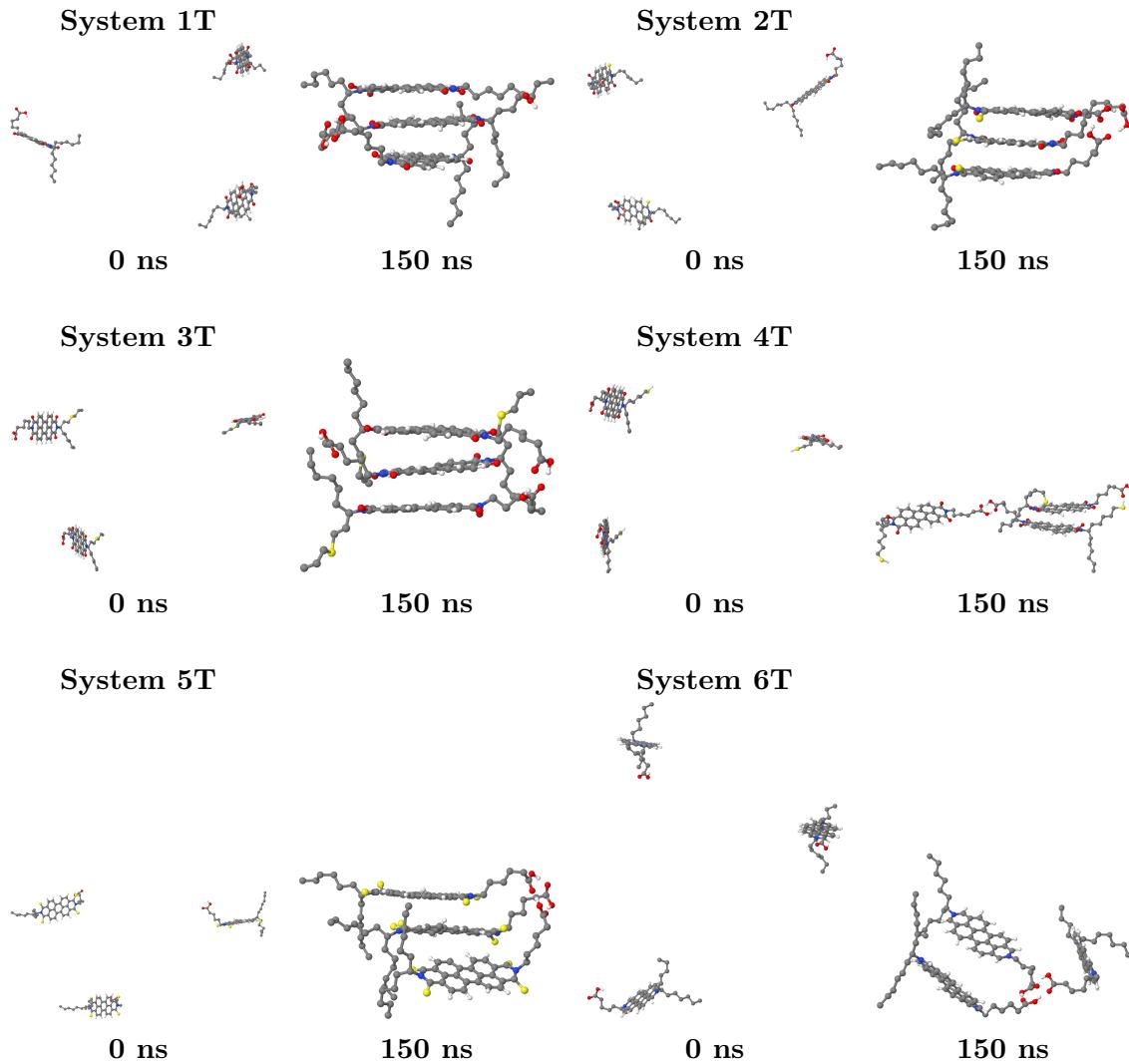


Figure 6.6: Snapshots of the initial configuration and the final frame of the trimers simulations for the random disposition case. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.

Dimers (Systems 1D - 6D): The analysis of monomer distances shows that systems 1D through 4D are stable as dimers, regardless the initial configuration within the box (Figure 6.7 and 6.8). For 1D, the average distance between the two molecules is lower than 0.4 nm during $\sim 88\%$ of the simulation, when the aggregated starting point is used and during 90% for the random starting point situation. The other systems (2D, 3D and 4D) are also stable and are found in the aggregated state during 85 (96), 94 (98) and 96 (94)% of the simulation for the “organized” (random) starting point. Visual inspection of molecular dynamics simulation (MDS) trajectories indicates that systems 5D and 6D are not stable during the molecular dynamics simulation in toluene and this regardless of the starting configuration. However, the systems presented different behavior.

For the “organized” starting configuration: the monomer from system 5D rotates at the beginning of simulation ($\sim 1\text{ns}$) and becomes separated during the MDS. At system 5T, one monomer rotates at 50 ns followed by side chains interaction. However, this configuration is not stable, since the monomers dissociated after 90 ns. Systems 6D and 6T assume many orientations and remains aggregated during all simulation. Nevertheless, they do not adopt defined geometries. This behavior is typical of an interaction done exclusively by means of the acid chain end via hydrogen bonding. System 5D shows variable distances after 30 ns (Figure 6.7). The molecules are close to each other ($< 0.4\text{ nm}$) at only 15.4% (23 ns) of the simulation. System 6D shows instability since the beginning of simulation, with distances higher than 0.4 nm during the simulation time.

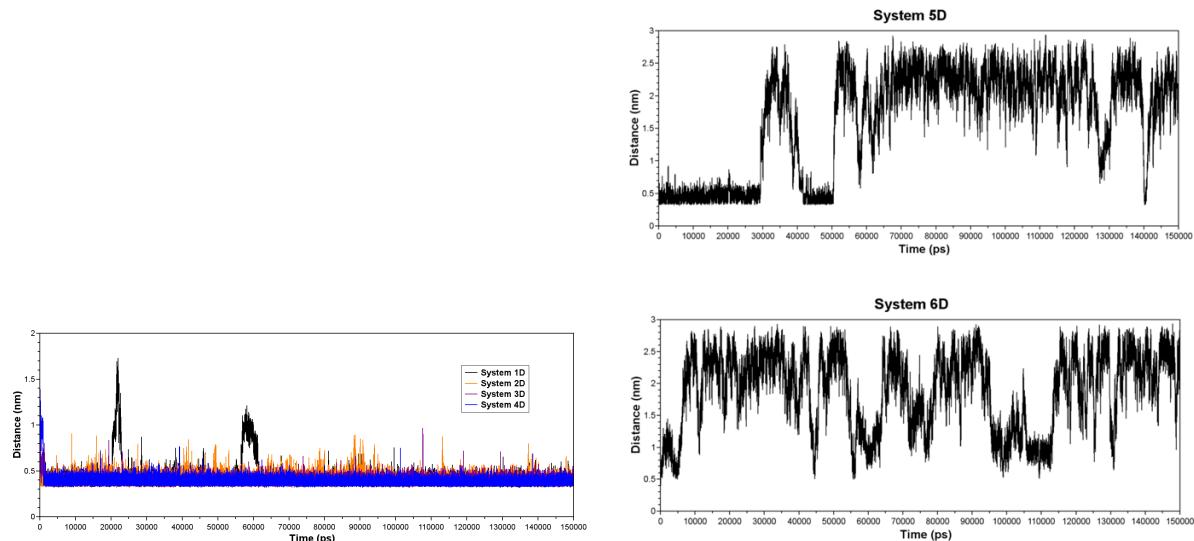


Figure 6.7: Distances between monomers for “organized” starting point.

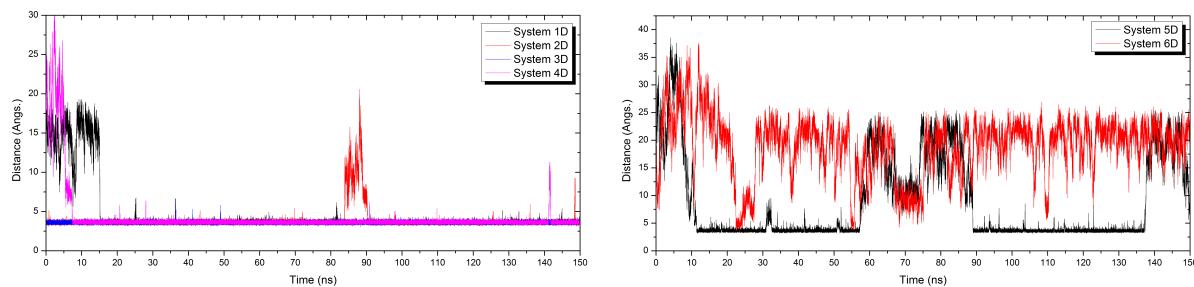


Figure 6.8: Distances between monomers for “random” starting point.

The formation of hydrogen bonds both between two acid chain ends or between an acid chain end and the carbonyl group of the conjugated core are probable to take place for all the 1D-6D systems. One should note that, as it was expected, system 2D has a lower probability of

forming such hydrogen bonds interactions because a carbonyl group was replaced by a thionyl one. The same is also true for systems 5D and 6D.

The cosine of the angle between the aromatic cores' planes is a useful complementary information towards the determination of the stability of the stack. Values of $\omega = 0^\circ$ ($\cos \omega = 1$) or $\omega = 180^\circ$ ($\cos \omega = -1$) corresponds to parallel monomers (polyaromatic cores as $\pi - \pi$ stacking), and $\omega = 90^\circ$ ($\cos \omega = 0$) corresponds to perpendicular molecules (T-shape).

For the “organized” starting point, the cosine indicates planar co-facial interacting structures, except for system 1D, which has some fluctuations during the simulation, as it can be seen in Figure 6.9. For the random starting point, these fluctuations are more common for all the systems in the beginning of the simulation and system 2D has it in the middle of the simulation time. This indicates that, although the parallel co-facial interaction is preferential, as it is also indicated by the minimum distances values, the molecules have some freedom to rotate around their normal axis during the MDS under these thermodynamic conditions. Up to this point, we cannot discern and indicate if this is an effect of the place of the sulfur atom in the structure or if it is just the fact that the hydrogen bonds are dynamic and allow the molecules to rotate. In any case, these results are in agreement with the ones found by Teklebrhan and co-workers.[184] What can be said though is that the 1D-4D systems do not perform T-shape interactions, but this is the case for the system 5D and 6D, as it is reported in Figure 6.10.

The results obtained previously for systems 5D and 6D are now reproduced if the starting configuration is random. System 5D can assume a stable face-to-face configuration during some periods of the simulation but it is not stable throughout it. System 6D is not stable in this configuration and no stable conformation can be noticed by the distances analysis.

Trimers (Systems 1T - 6T): Previous molecular dynamics simulations have shown that molecule 1 (C5Pe) has a high tendency to dimerize in toluene, but it does not form larger aggregates, since it has a long and linear alkyl side chains.[184] However, the distance between the monomers of system 1T is constant from the beginning of the simulation, indicating an average height lower than 0.8 nm. This is in agreement with experimental work,[160] which suggest a height of 0.67 nm. Systems 2T and 4T (after 33 ns) exhibit similar results.

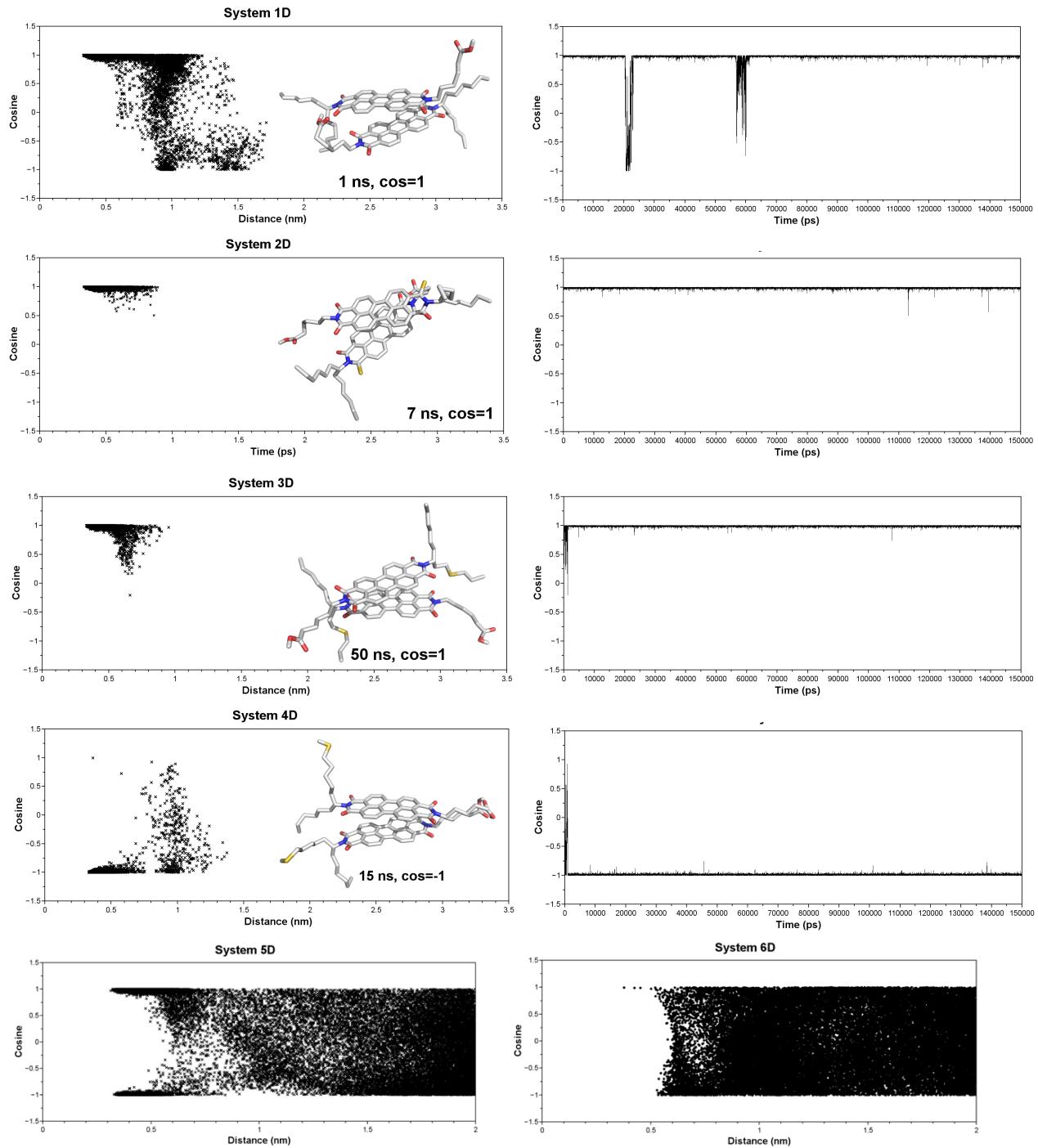


Figure 6.9: Cosine of angle between aromatic cores in function of distance between dimers monomers for the organized starting point. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.

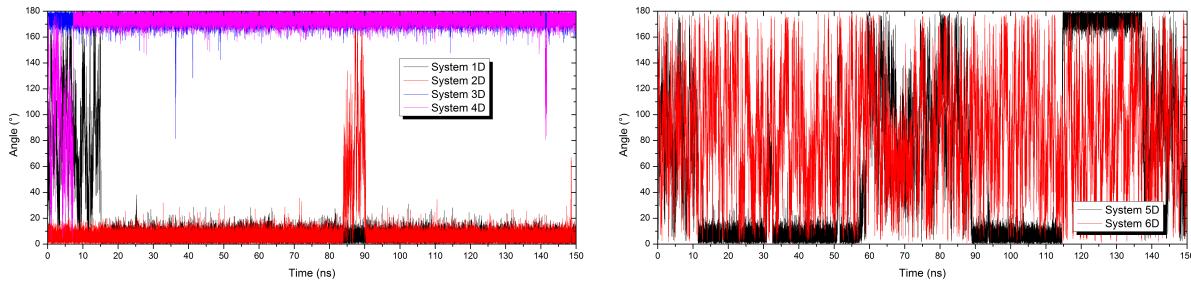


Figure 6.10: Cosine of angle between aromatic cores in function of distance between dimers monomers for the random starting point. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.

As it can be seen from Figures 6.11, 6.12, 6.13 and 6.14, the fluctuations in these distances are a function of the starting configuration. In “organized” starting configuration, system 3T shows larger distances after 45 ns and the molecules are close to each other (< 0.4 nm) at only 4% of the simulation. For system 3T, the results show variable molecular dissociation of one monomer, but the remaining dimer is stable. For the random starting configurations, systems 1T and 3T are stable as face-to-face aggregates during almost all the simulation time, whereas 2T takes 75 ns to stabilize in this configuration. The 4T now has the same effect in this situation that 3T had before: the trimer is not stable in the face-to-face configuration although the three molecules are kept together by the hydrogen bonds and the two of them assume the face-to-face configuration, as it is the case of 4D system.

When the predominant type of interactions was investigated for the case with the “organized” starting point, it was noticed they are constant during the whole simulation for system 1T, 2T, and 4T (Figure 9). Those molecules seem to be stable as a trimer. The interaction between the first two aromatic cores is not influenced when a third monomer is added, which in turn, has an identical behavior with the second molecule. It suggests that the interaction driving the aggregation is made from the poly-aromatic cores. Although, the side chain should contribute to aggregation since it interacts during all the simulation for the four systems.

The monomers geometries can be also analyzed by the cosine values of the angle between poly-aromatic cores (Figures 6.12 and 6.14). For all trimer systems, the planar configuration is more likely than others, namely the T-shape. System 3T, in the “organized” starting point (Figure 6.12) and system 4T in the random starting point (Figure 6.14) show now variations in these angles, either in the beginning or during the whole simulation. This, added to the fact that the intermolecular distance is similar for both 1-3 and 2-3 molecules, is an indication that

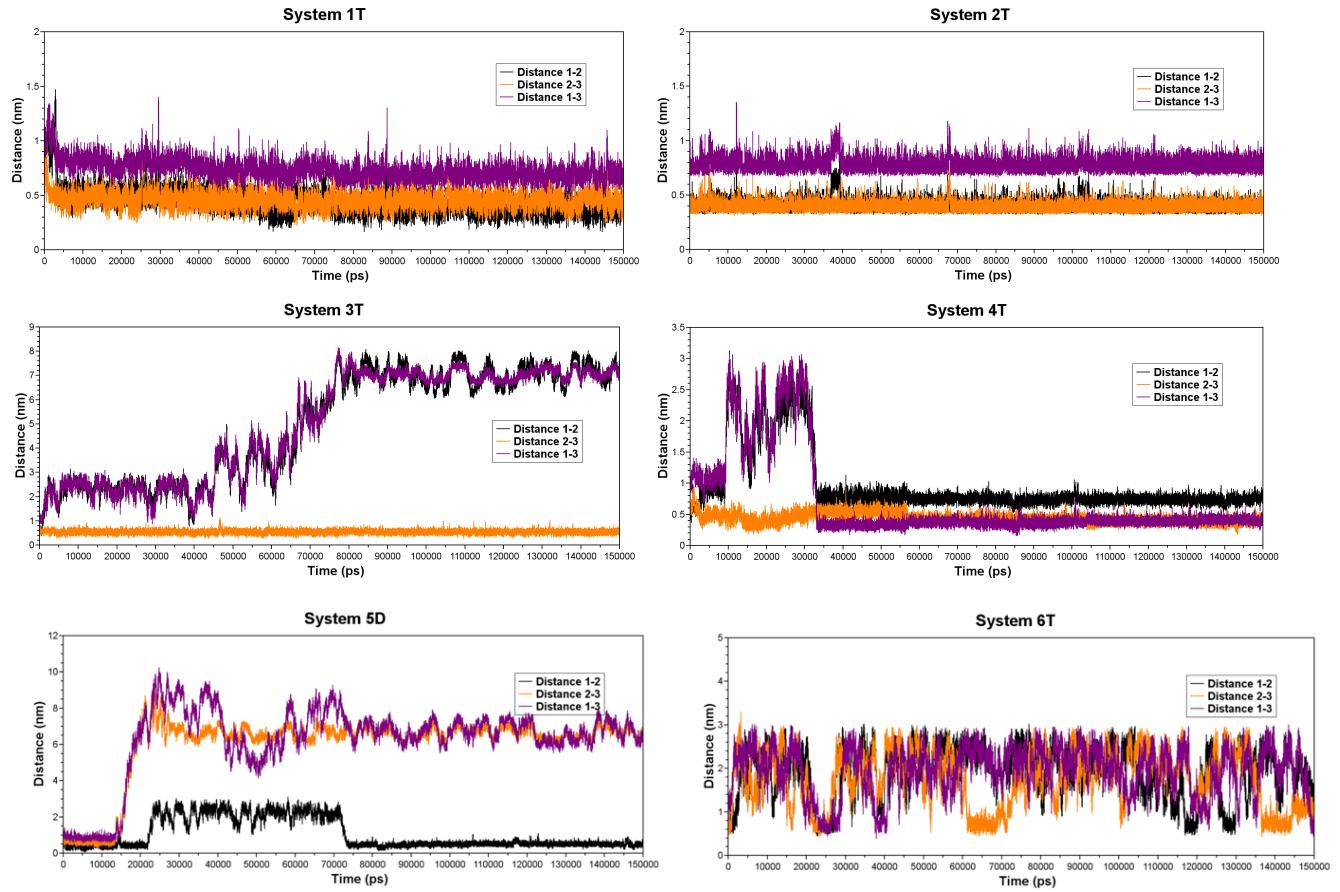


Figure 6.11: Distances between monomers for “organized” starting point.

the dimer is stable and interacts with the third molecule via the hydrogen bonds.

The distances between monomers of systems indicate greater variation for system 5T and 6T. System 5T shows larger distances after 30 ns and the molecules are close to each other (< 0.4 nm) at only 13% (19.5 ns) of the simulation. System 5T shows larger distances during the entire simulation, due to monomers dissociation. The trimers geometry was also analyzed by the cosine values of the angle between poly-aromatic cores. For both systems, the cosine does experience large variation and does not have association to short distances.

The stable trimers also interact through hydrogen bonds during the simulation. On systems 2T, 3T and 4T, the hydrogen bonds interactions occur basically between carboxylic groups of side chains (Figure 11). For systems 4T, the frequency of these hydrogen bonds is higher than 96% of the total simulation. For systems 3T, the numbers decrease to 63%, since only side chains from monomers 1 and 2 interact. Furthermore, only for system 1T, the numbers of hydrogen bonds between the side chains and the carbonyl group of poly-aromatic core are

significant (80%). Moreover, the presence of carboxylic acid groups is important to aggregation, but they were not enough present to maintain the trimer in a parallel orientation, as observed in systems 1T, 3T and 4T (from random starting point).

The side chains from trimers interact through hydrogen bonds during the simulation. At systems 5T, the frequency of hydrogen bonds between carboxylic groups of side chains is higher than 50% (75 ns) of total simulation, but for system 6T, this number decreases to 26% (39 ns). Moreover, the presence of carboxylic acid group is important to aggregation, but they were not enough to maintain the aggregated in a parallel orientation.

General effect of the Sulfur atom: Up to now, the properties of the aggregates of molecules 1-4 has been discussed for dimers and trimers. Although some differences in their dynamics in solution could be noticed during MDS, no final conclusion on the role of the sulfur hetero-element can be extracted so far. In order to circumvent this shortcut, we have introduced, for the random starting configurations, the study of the radial distribution functions (RDF) for the S–X pairwise interactions, where X can be: 1 - the oxygen atoms of the aromatic core (aromatic carbonyl groups), 2 - the oxygen atoms of the carboxylic acid group (COOH), and 3 - the hydrogen atom of the same group. This was motivated since it is known that sulfur atoms can form non-covalent (Coulomb and/or van der Waals) interactions with oxygen and hydrogen,[201] sometimes strong enough to lock conformations in place. These graphics can be found in Figure 6.16 for the dimer and 6.17 for the trimer systems.

In dimer and trimer systems, all the peaks found in the S-aromatic carbonyl oxygen RDFs are due to intra-molecular configurations. This is known by measuring the S–O distances of an isolated molecule of each type. Minor peaks, as it is the case of the ones found around 3, 8 and 14 Å for system 2D, are due to the interaction with oxygen atoms of the other molecule interacting with the one bearing the reference sulfur atom. The other systems show similar behavior with wider peaks now due to the fact that sulfur atoms are in flexible parts of the molecule. This allows us to state that there is no particular affinity between sulfur atoms and aromatic carbonyl atoms in the studied structures and this cannot be a likely interaction mechanism.

The S–O (COOH) and S–H have now new features that could not be observed in the pre-

vious case. Considering both dimer and trimer systems, these two RDFs have intra-molecular peaks centered at around 18 Å for molecule 2 (systems 2D and 2T), 22 Å for molecule 3 (systems 3D and 3T), and 26 Å for molecule 4 (systems 4D and 4T - out of the measurement range, but one can see the peak in the end of the graph). The other peaks are due to inter-molecular interactions between adjacent molecules. Molecule 2, regardless if it is a dimer or a trimer, does not have these additional peaks, indicating that the presence of the sulfur atom on the conjugated backbone does not have any particular affinity with neither oxygen or polar hydrogen atoms of the other molecules. On the other hand, molecule 3 (system 3D) has a “bump” for both S–O (COOH) and S–H (COOH) RDFs centered at 8 Å that does not come from an intra-molecular interaction and is a signature of inter-molecular affinity between the sulfur atom and these oxygen and hydrogen atoms. This behavior is more pronounced in trimer systems, as it expected, since this interaction can take place more easily.

This behavior is still more pronounced for molecule 4 (system 4D and 4T): the S–O and S–H inter-molecular interactions are now very likely to take place and, considering the peak centered in 2.5 Å of the S–H RDF that is not present in the S–O curve (see Figure 6.9), one could say that there is also an affinity of this terminal –SH group with the acid chain ends of the neighbor molecule. This same characteristic is found for trimer systems (Figure 6.10).

Based on these observations, one can state that sulfur atoms grafted directly to the conjugated backbone has a little influence on the dynamic behavior of dimer and trimer asphaltene systems and one might not be able to differentiate this with what is observed for the case where there is no sulfur atom (molecule 1). Oppositely, the presence of sulfur atoms on the lateral chains, either in the middle of it or as a chain end, has a more pronounced effect on the properties of these molecules (molecules 3 and 4). This is due to the fact that this atom can now interact with acid chain ends of neighbor molecules by means of non-covalent interactions like Coulomb and/or van der Waals.

Although the differences herein highlighted can, in our scale, not be of uppermost importance, it is the case in oil industry where huge quantities of molecules are being treated at the same time. These results can help towards the comprehension why crude oils with an elevated sulfur content are more difficult to crack, as a matter of fact.

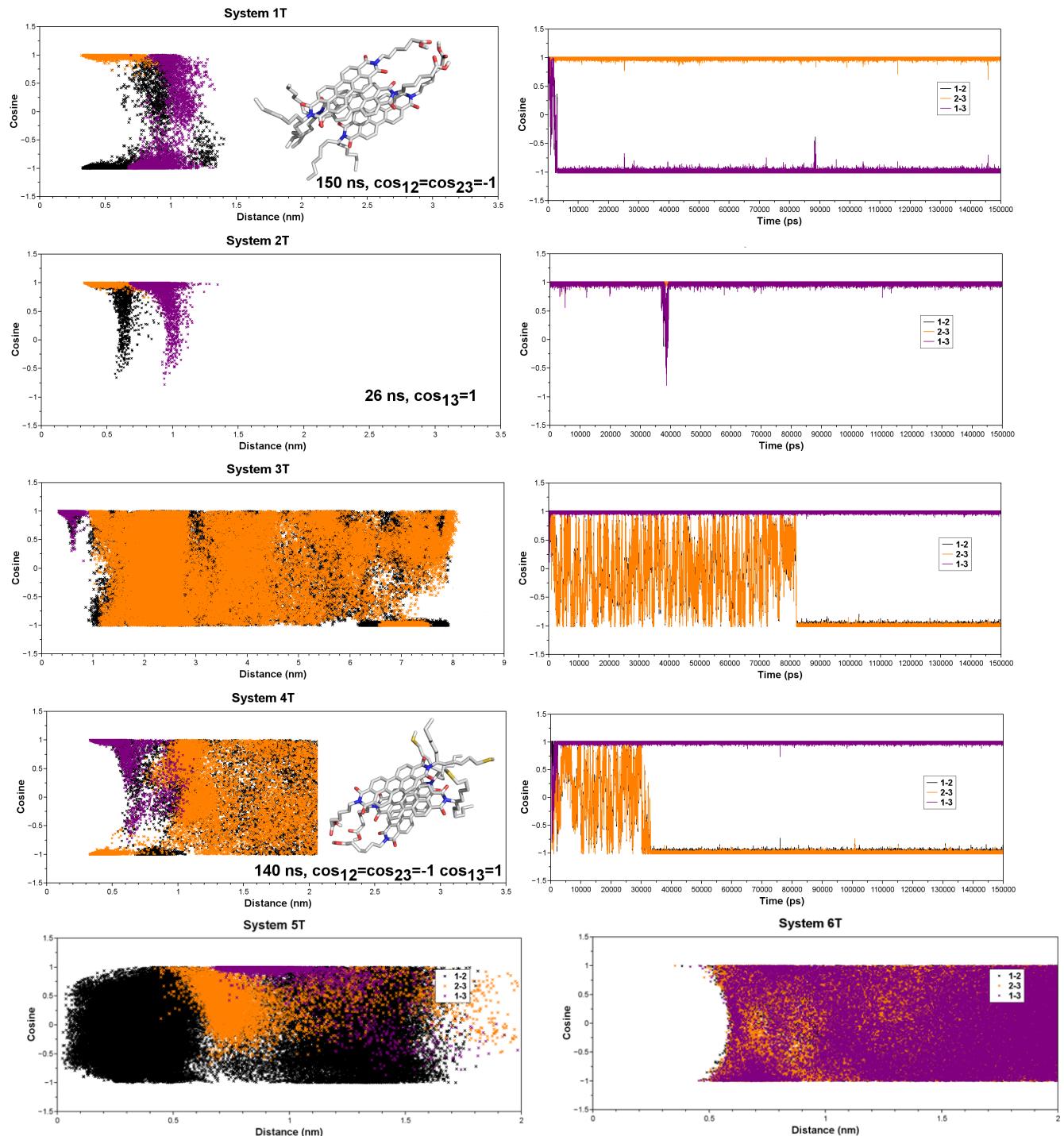


Figure 6.12: Angle between monomers for “organized” starting point.

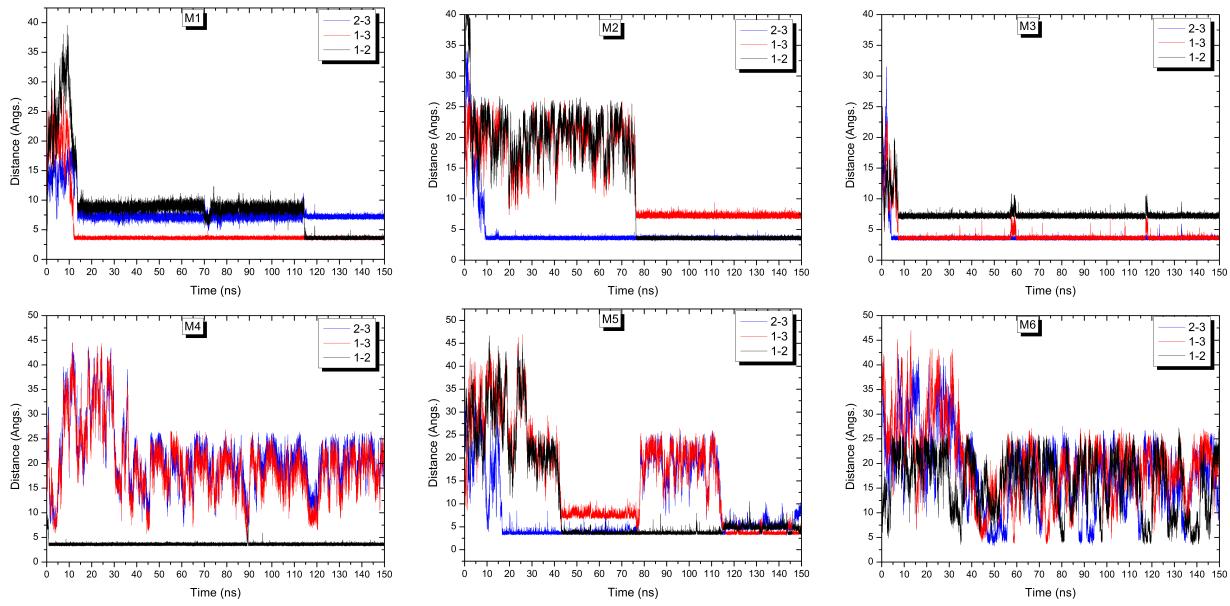


Figure 6.13: Distances between monomers for “random” starting point.

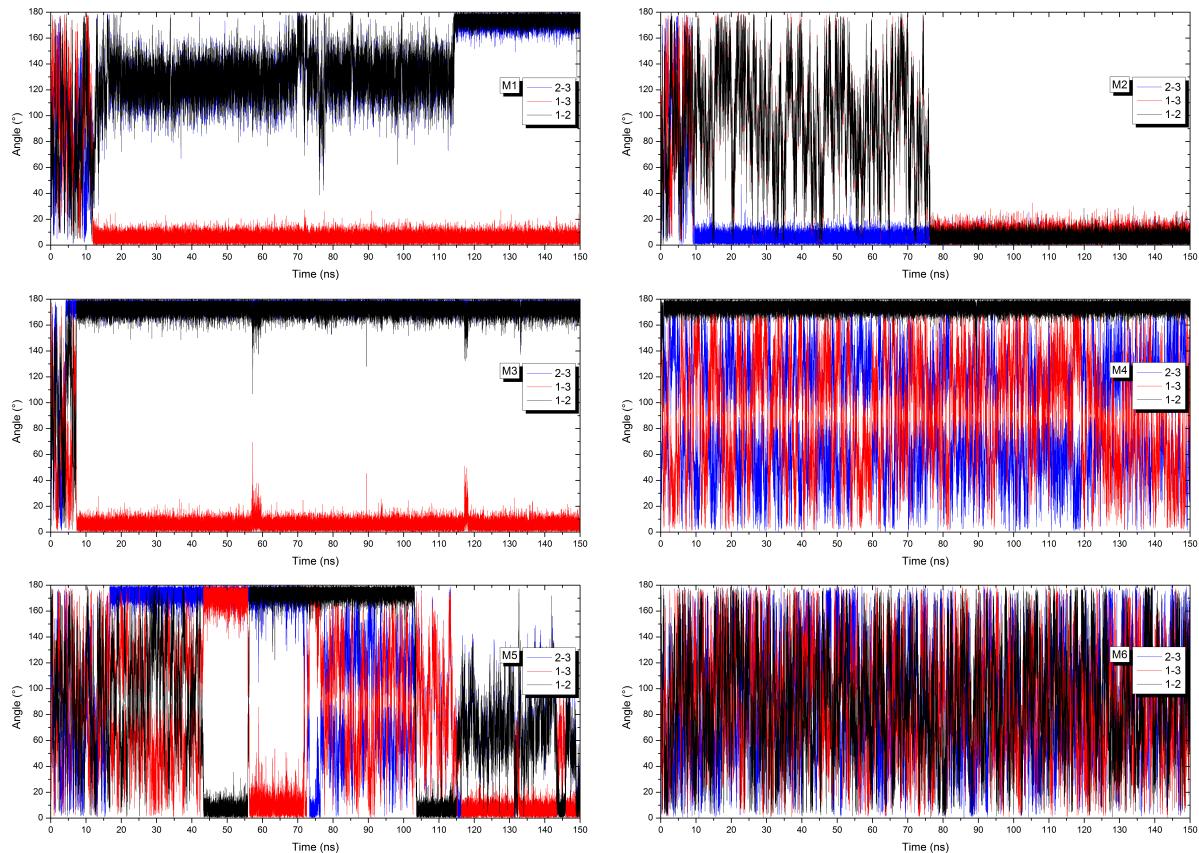


Figure 6.14: Angle between monomers for “random” starting point.

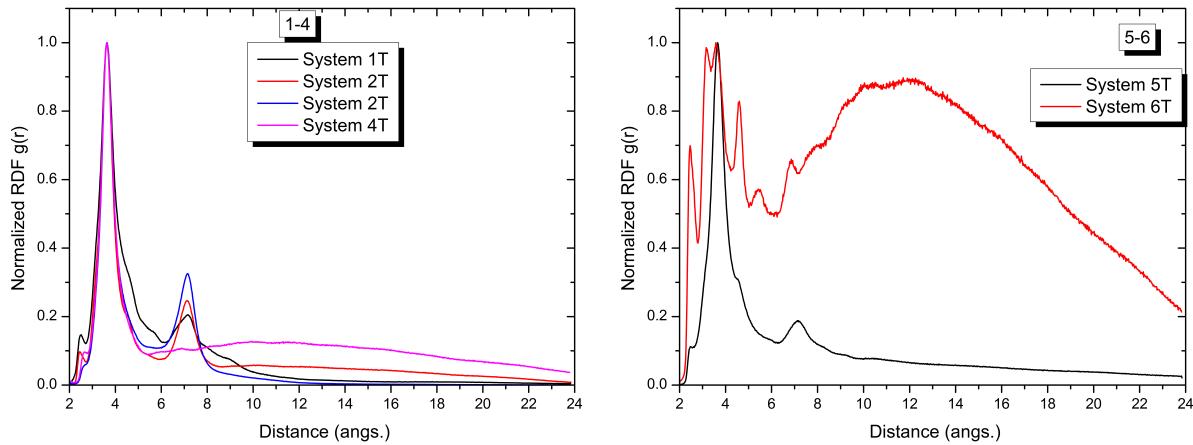


Figure 6.15: Normalized Radial Distribution Functions for the trimer systems.

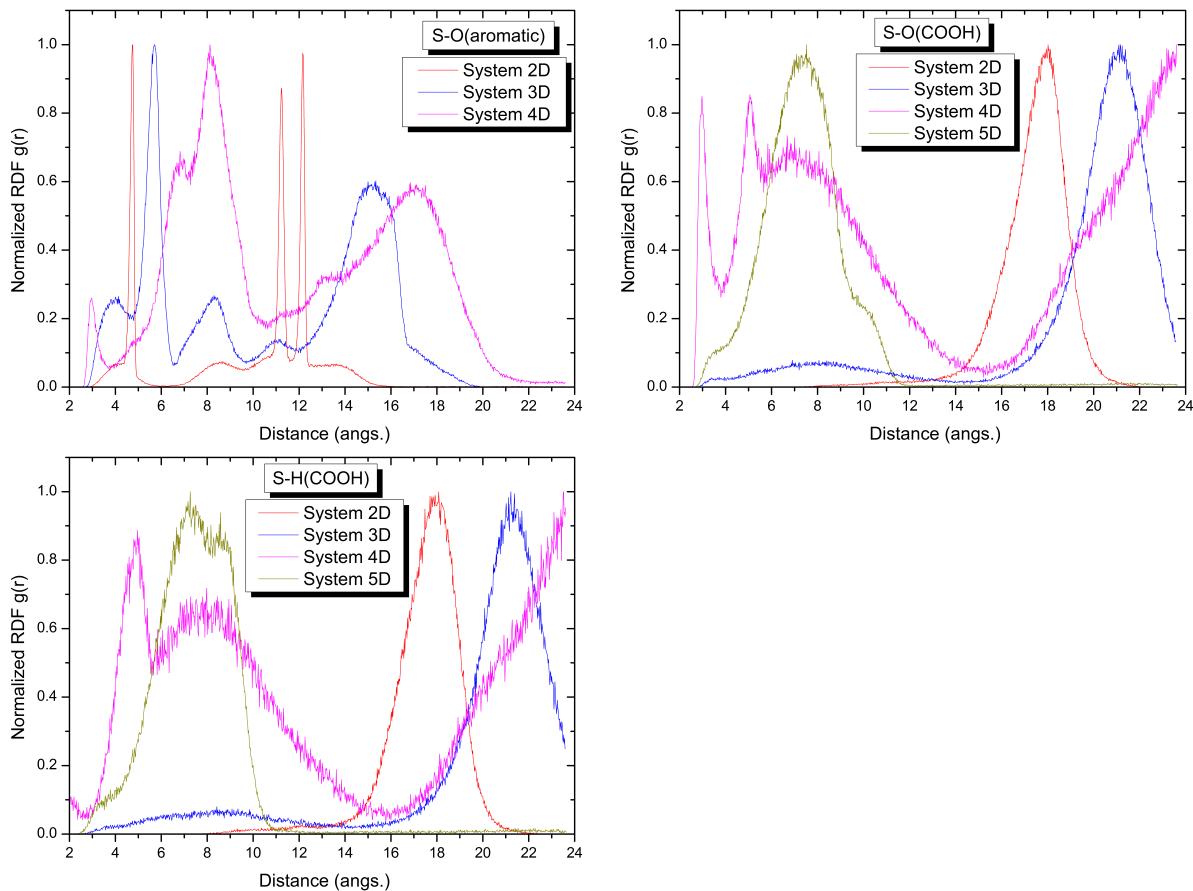


Figure 6.16: Normalized Radial Distribution Functions for the S–X pairwise interactions, where X can be the oxygen atom of the aromatic core, the oxygen atom of the carboxylic group (COOH) or its hydrogen for the dimer system.

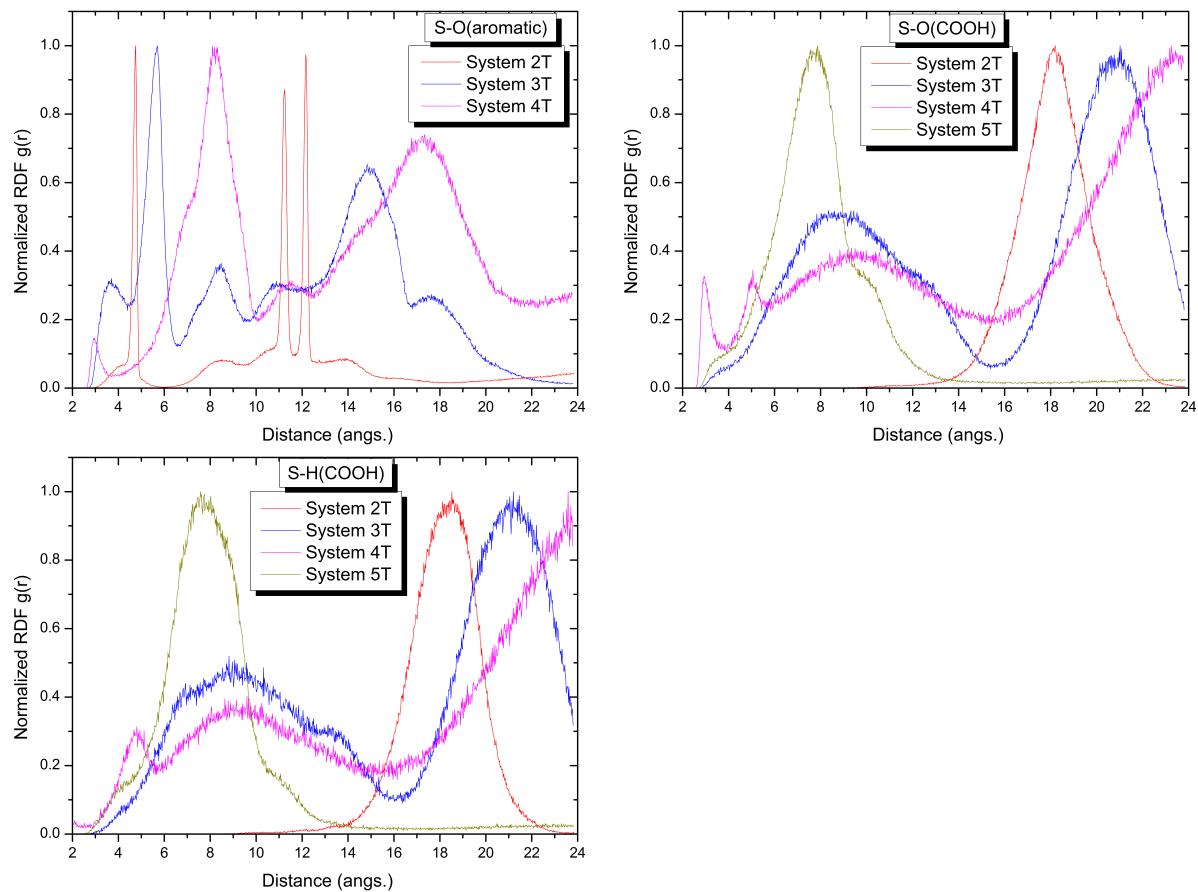


Figure 6.17: Normalized Radial Distribution Functions for the S–X pairwise interactions, where X can be the oxygen atom of the aromatic core, the oxygen atom of the carboxylic group (COOH) or its hydrogen for the trimer system.

6.3.2 Schuler's-derivatives systems

The systems were arranged so that to determine the separated effects of the different heteroatoms and the chain-ends.

6.3.2.1 Effect of the heteroatom

In order to study the effect of the heteroatom on the aggregation process of the PA3-type molecule, we screened the three free positions letting X assume the values of O, S and NH. Moreover, the chain-end for this purpose was chosen to be $-\text{CH}_3$ in order to avoid any effect due to hydrogen-bonding on the aggregation mechanism, as it will be studied later on. Thus, the following systems were considered: A13, A23, A33, A43, and A53. Although only one heteroatom is found in family CA22, *i.e.* nitrogen, we also present here, for the sake of comparison, the results obtained for the AAC molecule. In Figure 6.18, one can find the normalized RDF functions of these systems under 298 K and 1 bar conditions.

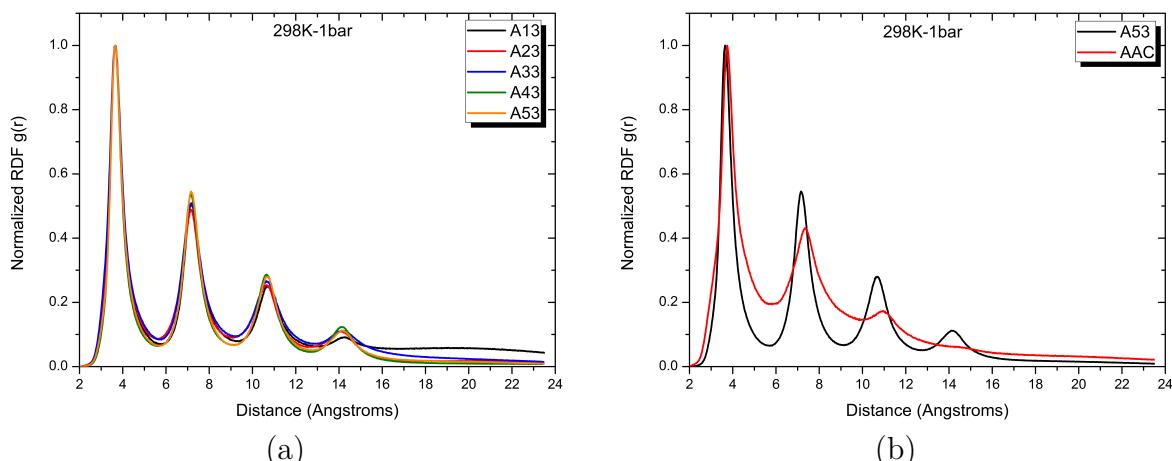


Figure 6.18: Aggregation mechanism with $-\text{CH}_3$ ended chains: (a) effect of the heteroatom (A13, A23, A33, A43, A53) and (b) effect of the asphaltene type (A53/AAC). Each peak of the RDFs corresponds to a neighbor molecule in the π -stacking structure presented in Figure 6.3.

These curves describe the organization structure in liquid state of the asphaltene solution. Such regular separation between peaks with an interval of around 3.9 Å indicates that the asphaltenes are organized in a stack structure. The first peak centered at ~ 3.9 Å stands for the interactions with the first neighbors generated by the π -stacking interaction. Depending on the position of the molecule within the stack, it can have one or two first neighbors. The other peaks represent the second, third and fourth neighbors in the stack. In this way, as the RDFs describe a probability density, it is almost twice more probable for an asphaltene to

have a first neighbor at 3.9 Å than a second neighbor at 7.8Å.

Different heteroatom substitution has a slight effect on the shape of the RDFs, although there are intensity changes (not seen here because the normalized version is presented). This is probably due to the fact that the total density of the liquid is slightly changed in function of the heteroatom (as they have different masses and different Van der Waals radii). The most important change concerns the CA22 asphaltene type AAC. This molecule has an aggregation pattern that also follows a stack architecture but high-order neighbors are not as probable as they are for the PA3 type. This is indicative of a short-range order whereas a long-range one is not kept during the simulation. This is evidenced by the snapshots of their molecular architecture after 60 ns of molecular dynamics simulation at 298K and 1 bar in Figure 6.19.

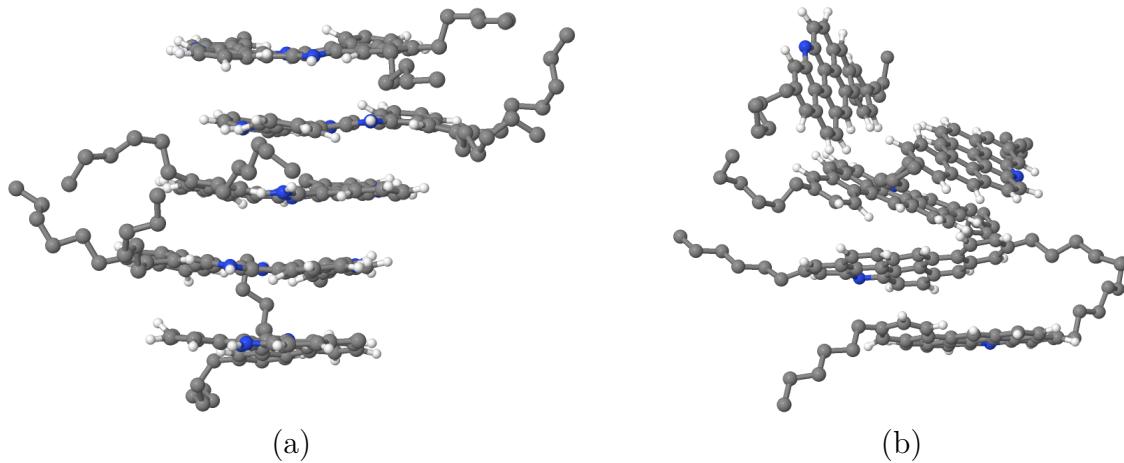


Figure 6.19: 60 ns-snapshots of 5 molecules of each asphaltene in toluene at 298K and 1 bar conditions: (a) A33 and (b) AAC. Carbon atoms are presented in gray and nitrogen atoms in blue. Aromatic or polar hydrogen atoms are presented in white. Toluene molecules are omitted for clarity.

Whereas PA3-type molecules, regardless of the heteroatom, aggregate forming co-facial planes, CA22-type molecules stack in a not completely planar way, leaving room to the formation of tilted architectures between adjacent molecules. This is the case during the whole simulation. This explains the lower intensity peaks present in the RDF. In order to deeply understand these features, the analysis of the energies of interaction is then a powerful tool to elucidate the effect of changing the heteroatom on the aggregation strength of these molecules. We studied the dimers of each molecule in vacuum dynamics for 10 ns using the same conditions as during the stabilization step. The time-averaged interaction energy $\langle E_i \rangle$ was determined by:

$$\langle E_i \rangle = -(\langle E_d^p \rangle - 2 \cdot \langle E_m^p \rangle) \quad (6.2)$$

$\langle E_d^p \rangle$ stands for the average value of the potential energy of the dimers after it has stabilized and $\langle E_m^p \rangle$ is the average value of the potential energy of a single isolated molecule after it has stabilized. The same method of calculation has been used by Liu *et al.*[163] The interaction energies of PA3 and CA22 dimers in vacuum, reported in Table 6.2 (diagonal values), corroborate the interpretation extracted from the analysis of the RDF functions: for any given molecular structure within the same class, changing one heteroatom by another has a slight effect on these molecules' dimers aggregation. The energies of interaction fluctuate around the average value of 55.7 kcal.mol⁻¹ (3.8x10⁻¹⁹J) for the PA3 class of molecules. For the CA22 molecule type, the interaction energy is lowered by around 10 kcal.mol⁻¹. The out-of-diagonal values reported in this Table will be discussed in the next sections.

	A13	A23	A33	A43	A53	AAC
A13	56.2	40.9	37.4	-	-	42.2
A23	40.9	52.2	46.7	-	-	47.4
A33	37.4	46.7	57.0	-	-	35.2
A43	-	-	-	58.9	-	-
A53	-	-	-	-	54.0	-
AAC	42.2	47.4	35.2	-	-	45.8

Table 6.2: Interaction energies ($\langle E_i \rangle$) in kcal.mol⁻¹ calculated for PA3- and CA22-type dimer molecules.

One should keep in mind that these molecules present different contributions of the Van der Waals and the Coulomb terms to the potential energy. Whereas one should expect the Van der Waals energy term to be higher for molecules containing Sulfur atoms than molecules containing Nitrogen and Oxygen atoms (motivated by the increase of the C₆ and C₁₂ Van der Waals terms, respectively). On the contrary, the Coulomb terms involving Oxygen should be higher than Sulfur's and Nitrogen's, motivated by the increased local dipole moments. In this way, the interaction energies result of both effects taking place at the same time, besides other contributions due to the angle, dihedral and improper torsion terms that are not equivalent elsewhere. The Van der Waals interactions are understood as the combination between Lennard-Jones repulsion terms (with a r^{-12} dependence) and the London dispersion/attractive terms, which is originated from induced dipole-induced dipole interactions (with a $-r^{-6}$ dependence). Finally, it is worth noting that the Coulomb interactions, besides having

a repulsive/attractive r^{-1} character, occur between fixed dipoles/charges-fixed dipoles/charges.

Decomposing the total potential energy in these terms, we found that the highest contribution to the interaction energy comes from the non-bonding Van der Waals potential terms. These contributions represent 86, 81, 84, 85, 82 and 73 % of the total interaction energy for A13/A13 ... AAC/AAC interaction pairs, respectively. The Coulomb energy contributions are not higher than 4%. This result indicates that, for the situations where a molecule interacts with a similar one, the Van der Waals forces govern the aggregation between the adjacent molecules.

6.3.2.2 Effect of chain-end

Having such results in mind, we studied the presence of $-COOH$ functions on the chain-ends for the same molecules under the same conditions, for A11, A21, A31, A41, A51 and AAA. Molecules A12 and AAB are presented in a separate plot that can be found in the ESI. The normalized RDF functions for these systems can be found in Figure 6.20.

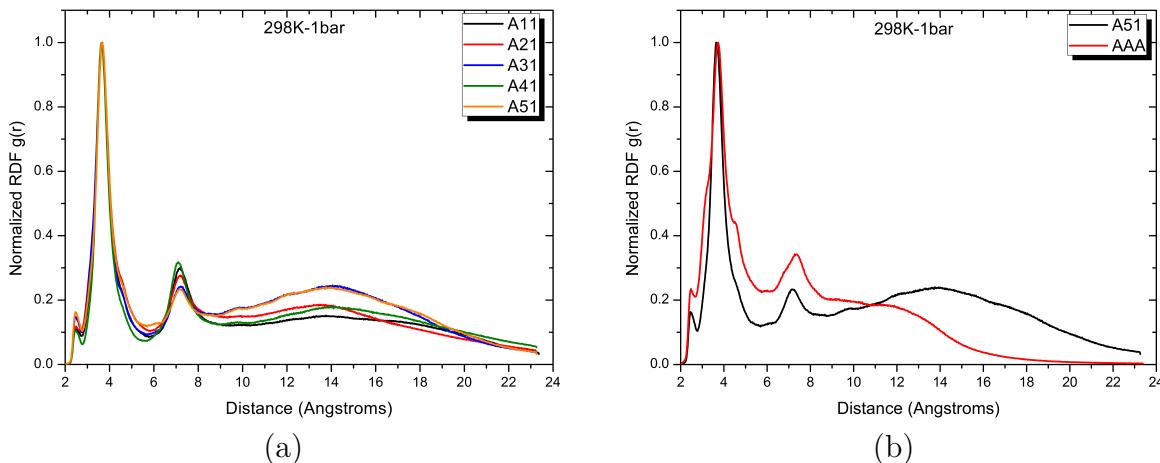


Figure 6.20: Aggregation mechanism with $-COOH$ ended chains: (a) effect of the heteroatom (A11, A21, A31, A41, A51) and (b) effect of the asphaltene type (A51/AAA). Each peak of the RDFs corresponds to a neighbor molecule in the π -stacking structure presented in Figure 6.19.

The RDFs differ significantly from the previous case with no acid chain: they are characteristic of liquids where the aggregates concern the stack of only two molecules, longer stacks presenting a lower probability. The interactions are now governed by another mechanism and the structure of the aggregate is less ordered in comparison to the stack structure, probably due to the presence of H-bonds, which signature can be seen as the 2 Å small peaks. The

acid chain-ends have an effect of overwhelming the π -stacking interactions beyond the second neighbor, hiding the third- and fourth-neighbor peaks.

The loss of symmetry due to the formation of H-bonds is identified in the RDFs by the wide peaks appearing for lengths superior to 10 Å. This wide peak is attributed to the interaction between asphaltene aggregates that are not rigid and have gained degrees of freedom from the presence of the H-bonds. Moreover, these wide poorly-defined bands can hide underneath it the peaks due to π -stacking aggregation. This can be visualized in the snapshots of A31, AAA, A33 and AAC presented in Figures 6.19 and 6.21.

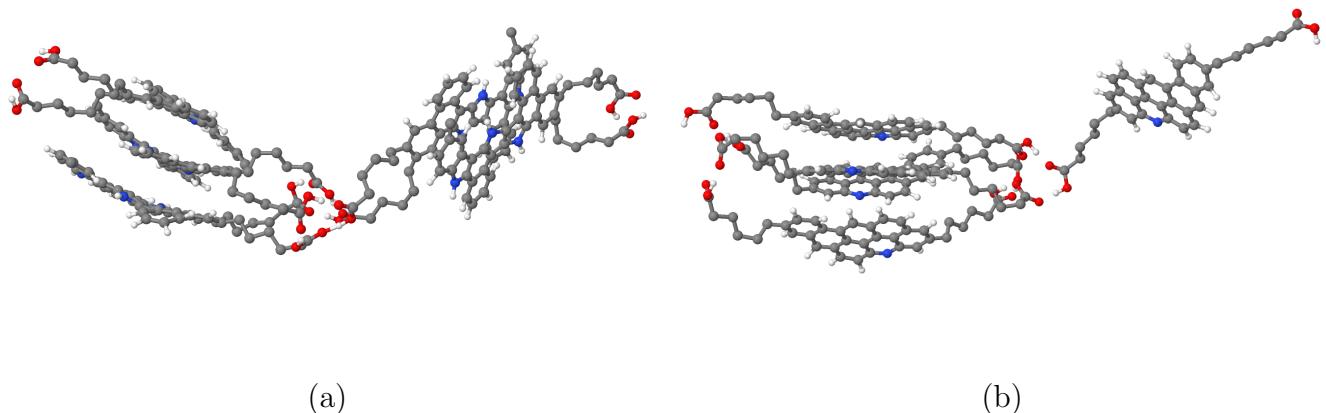


Figure 6.21: 60 ns-snapshots of 5 molecules of each asphaltene in toluene at 298K and 1 bar conditions: (a) A31 and (b) AAA.

Interestingly, the formation of H-bonds are stable enough under these conditions to avoid the self-stacking as it is the case of the $-\text{CH}_3$ -ended molecules. This characterizes the formation of an elongated cluster in the presence of more than five molecules and could also be responsible for forming larger clusters more rapidly. The interaction energy of two *n*-hexanoic acid chains alone was calculated under the same conditions as before and we have found it to be as high as 22 kcal. \cdot mol $^{-1}$ for each pair of $-\text{COOH}$ groups. When decomposed, this energy is consisted of a balance between the Coulomb and Van der Waals interactions. The intensity of the binding character of Coulomb interaction is around 10 times bigger than the repulsion character of Van der Waals interaction.

This result shows that the presence of acid chain-ends and the consequent formation of

H-bonds within the crude oil is also a great aggregation factor. This can explain why these molecules are so difficult to treat in oil industry's production chain. Total interaction energies can be estimated to be as high as 75 kcal.mol^{-1} , including both the $\pi - \pi$ and a pair of H-bond interactions, as aforementioned. Such a strong interaction between two adjacent molecules is almost equivalent to a covalent bond. For larger aggregates, this energy should be multiplied by the number of aromatic cores in interaction and the total number of acid chain-ends.

As it has been reported experimentally,[21] the macro-aggregates should be linked to each other by weak dispersive interactions between their lateral chains or simply because they are more polar than their environment, causing that they remain close together. This should explain why, experimentally, stirring the sample is enough to separate these aggregates in smaller pieces. In an underneath scale, the interactions should be governed by H-bonds which are much more robust and should keep the nano- and/or micro-aggregates in place. Finally, the same H-bonds and the π -stacking between the conjugated planes as herein studied are supposed to be the interactions ruling the aggregation phenomena at the nano-scale. This multi-scale model of the organization of asphaltene aggregates was first proposed by Mullins *et al.*[158, 155, 167]

6.3.2.3 Aggregation within heterogeneous boxes

Next, we will present the results of aggregation for boxes containing different asphaltenes, be it the same family or not. For the simulation of the mixtures between asphaltenes of the same family, we continued on using the same conditions but varied the relative concentration between the constituents.

As it was aforementioned, the effect of the heteroatom on the asphaltene aggregation with a given family can be considered as important as the difference of energies between two different classes of them. This means that, the effect of the heteroatom cannot be, at this point, ignored since the difference in energy found when changing it with a the same family of molecules (say A13 and 133) is of the same order of magnitude of the energy difference found when the asphaltene family is changed (say A13 and AAC). But, as a single model cannot be formulated to capture the effect of all the possible combination of heteroatoms, we excluded from now on the structures containing two different heteroatoms on the conjugated core of the same molecule (A41, A43, A51 and A53). This allows us to still have a large variety of

possible combinations.

Firstly, we study the binary mixtures within the PA3 molecular family. The combinations were done in order to screen the effect of relative concentration keeping constant the total number of five molecules within the simulation box. The concentrations were screened replacing a molecule of one type of asphaltene by another one at a time. This scheme is better understood after seeing Table S2 in the ESI (Mix 1a to 1d).

The simulation of these systems produces two different RDFs: the self-interaction and the cross-interaction curves. The first one describes the probability of any given molecule to interact with a system of the same nature during the simulation and the second one describes the probability of interaction between two different molecules. The curves for the A13:A23 mixture are presented in Figure 6.22. A13:A33 and A23:A33 mixtures can be found in Figure S1.

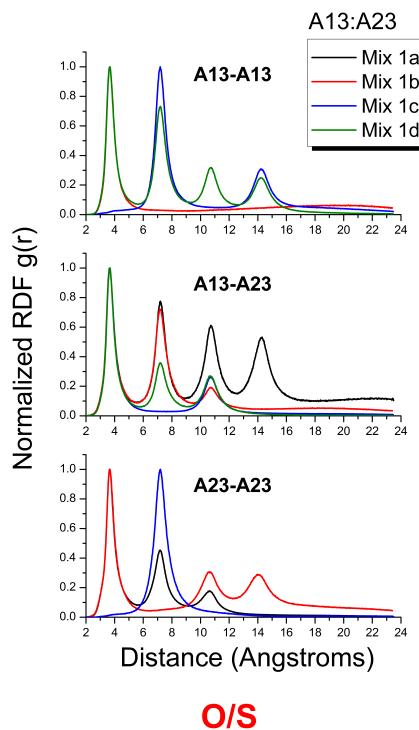


Figure 6.22: Effect of the mixture of A13 and A23 asphaltene molecules with $-\text{CH}_3$ lateral chain-end. This mixture concerns different ratios between Oxygen and Sulfur in the nano-aggregate.

These curves indicate that, regardless of the heteroatom, the asphaltenes have similar aggregation mechanisms, *i.e.*, the self-aggregate forming π -stacks with equidistant separation. This is represented by the peaks in the RDFs. As a matter of fact, A13 has the same inter-

action pattern with itself regardless of the concentration of A23 and the same is also valid for A23. This is an indicative that the presence of an asphaltene containing a different heteroatom on the backbone does not interfere on the aggregation of the parent molecule. Furthermore, the fact that not all the peaks are present for a given pair of molecules, e.g. A13-A13, indicates that the reference molecule has no similar to it as a first neighbor but it does have one as a second neighbor, what means that the asphaltenes are randomly intercalated among each others. Similarly, when the reference is A13, for instance, the probability of having A23 neighbors increases with the increase of A23 concentration, as it is expected. These very same conclusions can also be found for the other mixtures, concerning oxygen and nitrogen and sulfur and nitrogen ones. This allows one to say that, on the basis of these results, the aggregates formation is similar regardless of the combination among the heteroatoms. The major differences in these curves are originated from the random intercalation of the different molecules in the formation of the stack.

As previously, in order to discriminate if there is indeed any difference in the nano-aggregation process, the energies of interaction between each different asphaltene were studied and they are presented in Table 6.2 (off-diagonal values). It can be seen that the pairwise interaction energy changes for different combination of asphaltenes, *i.e.*, every molecule interacts more strongly with an identical molecule than with a different molecule. This effect might be directly associated to the Coulomb and Van der Waals interactions between these molecules. Upon mixing, the Van der Waals energies are expected to reduce. This is a consequence of the combination rule used to determine the Van der Waals coefficients C_6 and C_{12} of two distinct i, j atom pair ($C_{ij}^{6,12} = \sqrt{C_i^{6,12} \times C_j^{6,12}}$) induces that the cross-interactions coefficients to be smaller than the original ones.

6.4 Conclusions

Several asphaltene structural models have been proposed in the literature, such as the C5Pe molecule. Based on experimental work of Eyssautier and co-workers,[160] we have chosen six different model molecules for investigating the role of the sulfur atom position within the asphaltene structure. The molecular dynamics simulations have allowed understanding the stability of asphaltene dimers and trimers in toluene as well as showing that the position of the sulfur atom has an effect on the intermolecular non-covalent interactions that can take

place in these systems.

Although the polar side chain is important for determining the aggregation, the main interactions between the molecules are produced by the poly-aromatic cores, which are responsible to the parallel orientation of molecules. This geometry is confirmed by the cosine of angles between the poly-aromatic cores. We suggest that the presence of one sulfur atom in the poly-aromatic core (molecule 2) do not prevent the aggregation, since systems 2D and 2T are stable and the radial distribution functions of the pairwise interaction of this atom and the oxygen atoms of the neighborhood do not show any particular affinity. However, the presence of four sulfur atoms in poly-aromatic core avoids the molecules aggregation in a major part of the simulation and this may be due to both steric hindrance, Coulomb repulsion between sulfur and nitrogen atom of the adjacent molecule. Moreover, this molecular structure can not form hydrogen bonds between the acid chain end and the conjugated core of the molecule. On the other hand, sulfur atoms in the middle of the aliphatic chain, as well as a chain end, have a measurable effect on the stability of both dimers and trimers and this is due to the fact that they have a low, still measurable, interaction with the acid chain end oxygen and hydrogen atoms. Finally, these results can help towards the comprehension why heavy crude oils containing high values of sulfur are experimentally more difficult to treat and to crack.

Secondly, we have also studied the effect of mixing different asphaltene molecules on their nano-aggregation and which screening was based on experimentally isolated and identified molecules by Schuler and co-workers.[166] Molecular dynamics simulations were used in order to identify the nano-aggregation in homogeneous, heterogeneous and *super-heterogeneous* boxes. The first one consisted of only a type of asphaltene at a time; the second one had, within the same class, different asphaltenes mixed for different concentrations varying the heteroatom and the presence of acid chain-ends; and the third type included the mixing of asphaltenes of different classes with variation of the relative concentration. It seems that the nano-aggregation of these molecules is dependent on the heteroatom on the conjugated core and is more dependent on the chain-end and conjugated core size. The energies of interaction for each mix were calculated and are of the same order of magnitude. However, the energy of interaction between the same type of asphaltene tends to be always higher than the energies between different types and/or classes of molecules. The contribution of the acid lateral chains to the interaction energy

was calculated and found to be as high as 22 kcal.mol^{-1} . Moreover, changing the heteroatom can induce changes up to 10% of the total interaction energy if compared to other heteroatom. Finally, one should conclude that the aggregation mechanism might be much more strongly correlated to the size of the conjugated aromatic core and the size and nature of lateral chains than the presence of a particular heteroatom. These results are of great importance to oil industry and can help understanding how the molecular cartography of asphaltene molecules can induce stronger interactions or not.

Chapter 7

Acenes

Introduction

The strategy in the present work is to perform high-precision quantum mechanical calculations based on harmonic (solid state) anharmonic (gaz phase) electrical and mechanical approximations to compute spectra for individual molecules and dimers. With the anharmonic approach, unscaled computed outcomes including harmonics, overtones, combinations, intermolecular vibrations and other details of spectra can be compared directly with experimental data. The detailed identification of bands in experimental vibration spectra for crude oils and oil fractions may eventually permit the identification of the presence of members of a family of molecules, and the determination of the significance of specific types of molecular interaction arising in these fluids through spectral decomposition. Potential impacts include: improved speciation, and hence fluid characterization; and improved understanding of association phenomena, such as $\pi - \pi$ interactions, and hence transport property prediction.

Acenes, a family of linear polynuclear aromatic compounds possessing only $\pi - \pi$ interactions, for which high precision experimental data are available up to pentacene [202], comprise a benchmark example that illustrates both the potential and the computational challenges found in this line of inquiry. In this respect, various calculation hypotheses used for these large molecules are presented, together with the impact of these calculation hypotheses on the quality vis-à-vis experimental data and the computational cost of vibrational spectra.

Spectral assignments, determined on the basis of both wavenumber and relative intensity values, are made primarily using computed spectra for individual molecules, and dimers. In addition, based on experimental observation assessing the presence of deformed molecules in tetracene and pentacene crystals, impact on vibrational spectra of these deformations is evaluated and the phonon spectra is evaluated for the tetracene system. For some cases, assignment ambiguity is only resolved by tracing trends for vibration modes for the acene family as a whole. In the exposition that follows, each of these topics is tackled sequentially. Calculations with tetracene and pentacene monomers (individual molecules) are treated first. Then calculations for dimers and the acene family collectively are treated, and the results synthesized.

7.1 Calculation of vibrational spectra of Tetracene and Pentacene Monomers: Comparison with experimental data

7.1.1 Experimental data

Experimental solid-state infrared spectra of tetracene and pentacene [202] were acquired at room pressure and temperature. The measurements were performed on as received samples. Thus the crystallinity as well as the possible presence of defects in these solids are unknown. For tetracene, two crystal structures have been detected by x-ray diffraction experiments [203]. At room temperature and pressure, tetracene is triclinic [204] with space group $P\bar{1}$. The unit cell contains two independent molecules at the $(0, 0, 0)$ and $(1/2, 1/2, 0)$ inversion sites. Pentacene possesses two polymorphs [205, 206]. One comprises a triclinic layered structure with a herringbone arrangement in the layers, with two equivalent molecules per unit cell [207]. The other, commonly found at room temperature, adopts a 14.1 \AA $d(001)$ -spacing morphology with an inversion center on both molecules in the unit cell [208, 209, 210, 211, 212].

7.1.2 Computational details: construction of the potential energy surface

The Becke Three-Parameters Hybrid Functional (B3LYP), a combination of the Becke three-parameter exchange functional (B3) [49] that includes a mixture of Hartree-Fock exchange with the density functional theory (DFT) exchange-correlation and the LYP correlation functional [188], has proven to be a reasonable choice for predicting geometries of aromatic molecules [213, 214]. Stephens et al. [189] showed that the B3LYP force field yields infrared spectra in very good agreement with experimental data. It has become a standard method for studying vibrational spectra of organic molecules (in the far- and mid- Infrared) in the gas phase [215, 216]. Although the B3LYP method coupled with a sufficiently large basis set is a good candidate to study vibrational spectra of polynuclear aromatic molecules [202], this standard DFT calculation approach fails to describe long-distance interactions. Consequently, it does not reproduce spectra of interacting molecules [49, 189, 217, 218] where the interpretation of experimental vibration spectra requires consideration of both isolated and interacting molecules.

In parallel to the development to ab initio methods, some progress has been made on implementing the London dispersion force effect into DFT methods, permitting calculations with

larger individual molecules and binary pairs. Although the CCSD(T) method in combination with a sufficiently flexible basis set is the most accurate method for reproducing rotation-vibration spectra and to describe intermolecular complexes between aromatic hydrocarbons [219], such calculations are impractical for the description of large molecules [220]. The MP2 method [221], also able to handle dispersive interactions such as $\pi - \pi$ stacking [222, 223], is a computationally intensive and inadequate method for large molecules too [224]. Chai et al.⁷⁸ developed a functional based on optimized long-range corrected hybrid density functionals [225], which employs 100% Hartree-Fock (HF) exchange for long-range electron-electron interactions, ω B97X, to which they added an empirical dispersive interaction correction [?]. This method was tested by Salzner et al [226] for π -conjugated oligomers and it overcame the difficulties encountered with standard DFT functionals when dealing with interacting π -systems.

In this work, for isolated molecules and dimers, extended basis-sets consisting of atomic orbitals expressed as fixed linear combinations of Gaussian functions are employed for the calculations. A split valence 6-311G and 6-311G(p,d) [227, 228] was needed to provide correct calculation outcomes for naphthalene monomer [229]. Validation tests for the ω B97X-D/6-311G calculation method were also performed. The accuracy of DFT calculations in predicting molecular geometries and vibrational frequencies depend on the density functional employed. Thus geometrical optimizations were performed for isolated tetracene, using the ω B97X-D functional as well as the more conventional B3LYP method. The computational results are benchmarked using experimental data obtained by Campbell et al.[204]

The solid state calculations in the present work were carried out with the use of two approaches: First, the calculations were carried out with the projector augmented wave (PAW) implementation of Vienna Ab initio Simulation Package (VASP)[230]. The generalized gradient approximation (GGA) in the PBE parameterization was considered as the exchange correlational functional [46]. Structural optimizations (triclinic tetracene crystal) were achieved by setting the convergence criteria for total energies are below $5 * 10^{-6}$ eV, residual forces to be less than $1 * 10^{-3}$ eV/Å, and stresses limited to 0.02 GPa. To correct the missing dispersion interactions, we have used several recently proposed methods such as PBE functional with pair potential methods D2, D3 (BJ), TS, and TS+SCS. The details of the implementations and usage of these methods can be found elsewhere [108, 231, 114, 117, 232]. To obtain accurate band gaps, we have used the GW approximation⁸⁶. For getting accurate quasiparticle eigenvalues, we used 200 bands for the summation over the bands in the polarizability and the self-energy

formulas, and the polarizability matrices were calculated up to a cutoff of 200 eV.

Second, periodic calculations for tetracene were also performed in order to calculate IR intensities with CRYSTAL14 program[233], using PBE functional and STO-6G (Pople's STO-6G) basis set. Default criteria were used as consider accuracy and SCF convergence. Vibrational frequencies at the Γ Point were calculated for the optimized geometry. The empirical dispersion correction by Grimme [108] was performed to take into account weak interactions.

7.1.3 Geometrical parameters of isolated tetracene molecule optimized using B3LYP and ω B97X-D methods associated with 6-311G and 6-311G** basis set

Geometry optimizations were performed using two different DFT methods (B3LYP and ω B97X-D) implemented in Gaussian 09 software [234], and two basis sets (6-311G and 6-311G**). While Saeki et al. [229] reported a change of symmetry for the naphthalene molecule, from D2h to C2h (based on MP2/6-311G calculations), optimized geometries for the tetracene molecule, which also possesses D2h symmetry, remained planar with D2h symmetry in all cases. Calculated bond lengths based on different levels of theory are listed in Table 1 along with their mean deviations from experimental measurements (see also Figure A1 in the supplementary data). For these calculations, irrespective of the level of theory or the basis set employed, the agreement between computed bond lengths and experimental values is very good. The mean deviation from experimental values is less than 0.03 Å. These results show clearly that the choice of calculation method and basis set has a minor influence on the geometry of tetracene. Results obtained with pentacene monomer are comparable. This basis-set insensitivity for DFT approaches has been illustrated previously[235]. It permits the use of smaller basis sets and provides an additional argument for the use of DFT over ab initio methods, since larger molecules can be treated. The ω B97X-D method, which contains a London-force dispersive term, and is reliable for acene molecules was selected for this study and it is used in combination with the 6-311G basis-set.

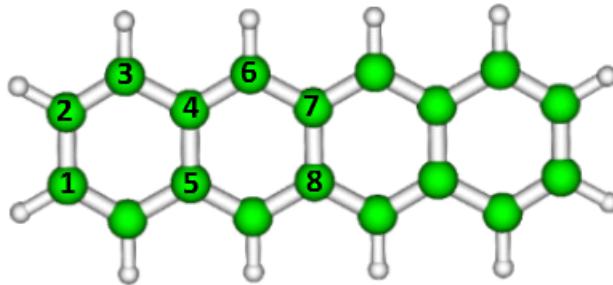


Figure 7.1: Representation of a tetracene molecule.

	$C_1 - C_2$	$C_2 - C_3$	$C_3 - C_4$	$C_4 - C_5$	$C_5 - C_6$	$C_6 - C_7$	$C_7 - C_8$
Experimental data ^a	1.46	1.38	1.42	1.42	1.39	1.40	1.46
B3LYP/6-311G	1.43	1.36	1.43	1.43	1.40	1.42	1.43
B3LYP/6-311G**	1.43	1.36	1.43	1.43	1.40	1.42	1.43
ω B97X-D/6-311G	1.43	1.37	1.44	1.45	1.39	1.41	1.45
ω B97X-D/6-311G**	1.43	1.37	1.44	1.45	1.39	1.41	1.45

Table 7.1: Method and Basis set dependence of C-C bond length (\AA) a for tetracene monomer, and mean deviation from experimental data^b

7.1.4 Truncation of the potential energy surface

Vibrational spectroscopy modelling outcomes, whether at the harmonic or anharmonic level, are directly dependent on the quality of the potential energy surface (PES). The PES is generally expressed as a polynomial determined by adjustment of a set of geometric structures of a molecule. Consequently this type of calculation depends on the quality of the electronic calculations of the wave functions as well. While it is now common to determine the PES accurately for small molecules (3-5 atoms) [94], this becomes almost infeasible when solving anharmonic approximations for systems of larger size where powerful computers are required to implement rigorous methods for solving the Schrödinger equation. Truncation of the analytical shape of the PES (degree of the potential), is unavoidable. For example, a tetracene molecule contains 30 atoms, and therefore has $3N - 6 = 84$ normal modes apart from the translational and rotational modes. 84 modes with an order 4 necessitates the calculation for 22171 force constants. Thus, truncation of the PES constitutes the first part of this study. Four levels of approximation were tested. The fundamental bands obtained in these cases, using second-order

perturbation theory (PT2), are compared to the ones calculated with the complete PES. The conclusion of these tests are summarized in Table 2 and elaborated here:

1. It is possible to truncate the PES in terms of only significant coupling. For this, it suffices to identify the constant of cubic and quartic force of very low intensity. We have realized calculations by successively suppressing force constants smaller than 2, 5, 8 and 10 cm^{-1} . This study showed that the discrepancy regarding the calculated frequencies of fundamental modes is not higher than 2.8 cm^{-1} , and 3.3 cm^{-1} for the combination modes. These disparities are small enough to justify use of this approximation. Moreover this allows setting 7, 33, 48 and 55 % of the forces to zero when suppressing constants smaller than 2, 5, 8 and 10 cm^{-1} , respectively, thus saving computational cost and enabling variational calculations for larger systems. Note, however, that this restriction is potentially acceptable only without any Fermi or Darling-Dennison resonance, often characterized totally or partially using these finest couplings. Regarding the time saved when suppressing the forces lower than 10 cm^{-1} while the wavenumber values of fundamental modes remain the same, it would be tempting to truncate the PES this way, but this would represent a serious risk of losing information regarding combination bands. Indeed our tests show that the upper limit of the truncation to maintain all the information of combination bands consists in considering all the forces higher than 5 cm^{-1} . Removal of these forces would disregard combinations whose weight is less important, resulting in the loss of details that provide a complete interpretation of infrared spectra. Thus, for the following tests of truncature and for calculations presented in this study, we will consider a PES calculated with force constants higher than 5 cm^{-1} .
2. The second truncation tested in this work consists of separating modes of high and low wavenumber. We assumed that the description of the low-wavenumber modes (mainly torsion modes), were weakly coupled with bending and scissoring modes occurring at higher wavenumber. This hypothesis has been tested and validated on the base of calculations on the tetracene molecule (Table 2); indeed each of the first 15 modes calculated using the PES truncated to torsional and libration modes are in perfect adequacy with values calculated a posteriori using complete PES.
3. We have also tested the association of a complete grid of points of the surface at the fourth order, with a fit at the sixth order on the diagonal, in order to refine coupling between

closely coupled modes. Because this refining is demanding in terms of computing time, we applied it on the 10 first modes only, supposing that the modes of higher wavenumber were correctly described with a polynome of 4th order. The results, listed in Table 2, show that this refining doesn't bring any improvement for describing the vibrational spectrum of tetracene.

4. The last approximation used in this work consists of applying a second adjustment of the PES; this adjustment would aim to eliminate the weight of vibrators we excluded in point (i), to refine the fit of remaining force constants so that they reflect, in fine, only the couplings between remaining modes. If this had proved to be an acceptable approximation, we could have extended this methodology to the study of specific molecules vibrators of undifferentiated dimension, by calculating a limited number of specific points to sought vibrators. In accordance with what Bowman et al [236] and Begue et al [237] showed on the study of water clusters, our calculations show that only the separation between libration modes with all other active modes is possible.

To conclude, we have shown that partial truncation of PES could be effective for our study. This truncation consists of extracting from a global fit, the force constants of higher order that allow the description of the problem beyond the harmonic approximation. These constants can then be directly used to solve Schrödinger's vibrational equation.

This methodology was used in the remainder of our work for the characterization and assignment of the spectra of tetracene and pentacene molecules.

7.1.5 Resolution of the vibrational Schrödinger equation (non-core bands with intensity)

The presence of overtones and combination bands in infrared spectra is a manifestation of the breakdown of the double-harmonic approximation, usually implemented in electronic structure codes. The calculation of intensities is essential in order to determine which non-fundamental transitions are active and to assign all experimental bands. Since both mechanical (anharmonicity of the potential) and electrical (non-linear dependence of the dipole moment on the normal coordinates) anharmonicities are expected to give non-zero intensities to non-fundamental transitions, these two effects must be considered in the treatment of transition energies and related vibrational wavefunctions. The wavenumber calculations in the mechanical anharmonic

approximation were carried out using a variational method developed by Bégué et al [94] Using this method, implemented in the P_Anhar.v2.0 program [90], it was possible to compute all the vibrational frequencies (fundamental, combination bands and overtones) that contribute to the mid- and near-infrared spectrum of tetracene and pentacene. In addition, the activity of each mode was also calculated in the electrical anharmonic approximation using a method developed by Baraille et al [237, 238] , which is also available in the same software.

The experimental spectra of tetracene and pentacene, extracted from the work of Michaelian et al [202] , are shown in Figure 2. The corresponding calculated infrared active transitions both fundamental and combination bands overlay the data and are listed in Table 3 and Table 4 where they are further compared with available experimental and theoretical values. Threshold intensity for IR-activity of 0.001 km/mol was chosen for the calculations. This threshold is justified following a discussion of band assignments for each molecule, and presentation of vibrational bands common for the two molecules.

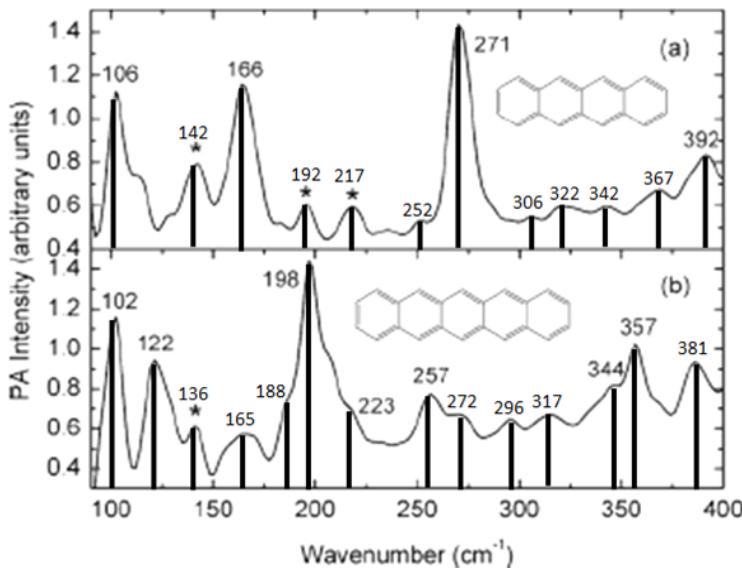


Figure 7.2: Far-infrared PA spectra of (a) tetracene; (b) pentacene. Spectra were acquired using phase modulation at a frequency of 5 Hz and an amplitude of 15λ ($\lambda = 0.6328 \mu\text{m}$), extracted from Michaelian et al [202]

For tetracene molecules, three bands were identified previously in the range $100 - 400 \text{ cm}^{-1}$ [202, 239]. Two were noted as infrared active at 166 cm^{-1} and 271 cm^{-1} , and one (388 cm^{-1}) was assigned as Raman-active (Table 3). By calculating intensities for all bands, including combination bands and overtones, 8 additional infrared- active vibrations (7 combinations and one second-harmonic) are identified, and the band at 388 cm^{-1} previously assigned as Raman-

active is identified as an infrared-active band arising from the B3u mode plus an overtone. Three similar cases where a vibrational transition could correspond either to a combination band, an overtone, or a Raman-active band, are also shown in Table 3.

Raman-transition active in infrared: molecular deformation

Raman transitions can become active in infrared spectre, when a molecule is deformed in the solid state. For example, Homes et al.⁵⁸ showed experimentally that acene molecules in the solid state are not planar and attributed this observation to crystal-packing forces that induce deformation; additionally, deformation effect on lattice-phonon frequencies has been measured, then enabling discrimination of different solid phases of tetracene [203] and pentacene [205, 240]. In order to underpin this contention, we performed calculations on a deformed tetracene molecule, shown in Figure 3. In this conformation the molecule does not possess D2h symmetry. The molecule does still possess symmetry, however. The deformation concerns dihedral angles [C1-C2-C3-C4], [C2-C3-C4-C5], [C4-C5-C6-C7], [C5-C6-C7-C8], [C9-C10-C11-C12] and [C10-C11-C12-C13], equal to 175° , 178° , -178° , -178° , -178° , -178° , respectively, instead of $\pm 180^\circ$ in the case of the molecule in the D2h conformation. The bond lengths are unchanged. Wavenumbers of fundamental modes and associated intensities calculated with the harmonic approximation are listed in Table A1 in the supplementary data. The calculated vibrational spectrum exhibits only positive wavenumbers, which attests that this deformation is located at a minimum of the PES. The intensities corresponding to the R-denominated-bands of the deformed molecule, calculated at the harmonic level, are given in Table 3. For the deformed molecule, modes that would normally be inactive in infrared (Raman active or silent modes), present significant intensities, ranging from 0.002 to 0.05 km/mol.

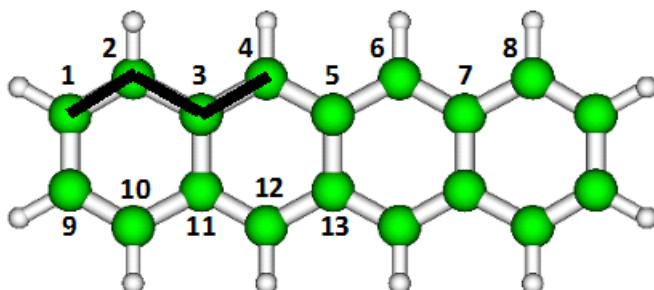


Figure 7.3: Tetracene molecule. The deformed dihedral angle C1-C2-C3-C4 is shown in black.

Assignment based on intensity value

Transitions for which there is a double hypothesis (142 , 192 and 392 cm^{-1}) can be resolved on the basis of relative intensity values. Thus the band at 142 cm^{-1} could be attributed to a second harmonic whose intensity would be three times lower than the Raman-active one that becomes optically active for the deformed molecule. The band at 192 cm^{-1} could be either a combination of an IR-active mode and a Raman-active mode, whose intensity would be ten times lower than a Raman-active band calculated at 196 cm^{-1} . The band at 392 cm^{-1} could either be a combination of an overtone and an infrared active mode, or a Raman-active band ten times more intense. From the relative intensities, in each case, the assumption of the Raman-band becoming active is more likely. For other transitions, there is a single hypothesis. The transition, denoted v_8 , located at 322 cm^{-1} is a Raman-active band while the bands located at 252 , 306 and 342 cm^{-1} are combination bands.

For pentacene, Table 4, the situation is similar. 11 bands arise from combination vibrations. One, found experimentally at 257 cm^{-1} , was previously assigned to a predicted Raman-active mode. Calculations on a deformed pentacene molecule lead to an IR active band at a comparable wavenumber and with a comparable intensity. Thus, this one assignment remains uncertain on the basis of these calculations.

Use of trends for assignment

The trends for normal modes of infrared active vibrations found for both tetracene and pentacene molecules (see types of vibration in Table 3 and Table 4) are shown in Figure 4 where they are also compared with experimental data. By viewing the wavenumbers in this way, wavenumber assignments for the two molecules can be evaluated in parallel. From the similarity of the trends for tetracene and pentacene, as well as the proximity of the respective computed wavenumbers to experimental ones, the assignment choices - combination bands over Raman-active ones – are supported and this outcome is independent of intensity values. Only one band remains unassigned in the experimental spectrum of tetracene, 106 cm^{-1} . Michaelian et al [202] suggested that it originated from lattice vibration. The pentacene band attributed to a Raman-active vibration, 102 cm^{-1} , may be of the same type as well. These bands and intermolecular bands more broadly are addressed by performing calculations with tetracene and pentacene dimers.

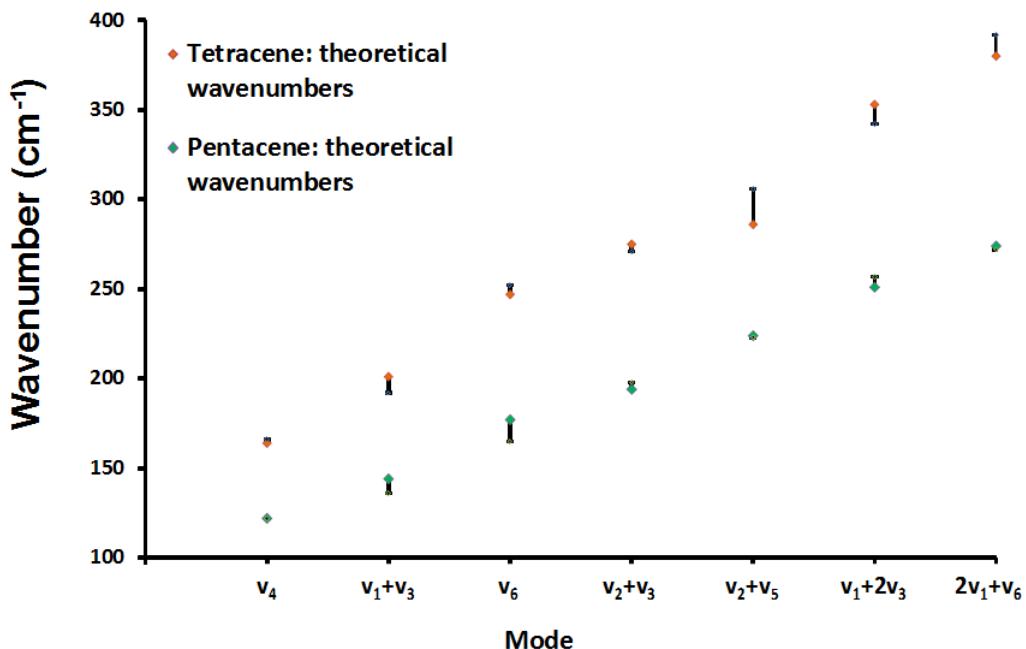


Figure 7.4: Comparison of theoretical and experimental wavenumbers of tetracene and pentacene for the same normal modes of vibration. The deviations between experimental and theoretical wavenumbers are represented with dark bars.

7.2 Calculation of vibrational spectra of acene dimers: Comparison with experimental data.

π -stacking is responsible for the cohesive structure of acene crystals [204, 210, 241, 242, 243]. Dimers constitute a benchmark for describing π -stacking phenomena. In order to generalize the effect of π -stacking on intermolecular vibrations for the acene family as a whole, the calculations described here include those for naphthalene and anthracene dimers in addition to the tetracene and pentacene dimers. Inclusion of these smaller dimers also permits calibration of the computational methodology employed.

7.2.1 Calculated geometry and wavenumbers of dimers and comparison with experimental data

Far-infrared spectra provide information about intermolecular and inter-ring interactions and about π -stacking in particular. Hineno et al [244] measured spectra of naphthalene and anthracene in the region $190\text{--}50\text{ cm}^{-1}$ at liquid helium temperature and observed that some bands were due to lattice vibrations. In the same way, bands in the experimental vibrational spectra of tetracene and pentacene that cannot be interpreted based on computed wavenumbers for a

single molecule are expected to arise from lattice modes. A first step toward understanding $\pi - \pi$ interaction of acene compounds consists of considering the interaction of acene dimers. The choice of the calculation method for describing these electronic interactions is justified, and then the differentiation of inter- and intramolecular vibrations in the acene family, is addressed.

Naphthalene dimer: to assess the performance of the ω B97X-D method for $\pi - \pi$ interaction calculations involving aromatic molecules, the robustness of the ω B97X-D/6-311G method for naphthalene dimers, for which a broad literature exists [229, 240] 50,97,98,99,100, is discussed in light of results based on reliable data and accurate computational results. For example, Gonzalez et al.102 and Tsuzuki et al [245] calculated ground-state structures and binding energies of van der Waals dimers of naphthalene using an aromatic intermolecular interaction model, at the MP2 and CCSD(T) levels of theory, respectively, using Pople's basis sets of various size. Their calculations yield two low energy dimers of similar energy: D_{2d} (crossed) and C_{2h} (parallel-displaced), and two C_{2v} T-shaped structures. Walsh et al [246, 247], using calculations at the MP2/6-31*(0.25) level found eight stationary points for naphthalene dimers: the four reported previously and four others of lower symmetry. Following these studies, Saeki et al [229] performed vibrational analysis of naphthalene dimers at the MP2/6-311G level of theory and their results are compared with the values obtained in this work. In the present work, geometry optimizations of naphthalene dimers in various configurations - parallel displaced (PD), parallel displaced according to C_{2h} symmetry, crossed-parallel according to C_2 and D_{2h} symmetry - using the ω B97X-D method, with the 6-311G and 6-311G(p,d) basis sets were performed. All stationary points were characterized by frequency calculations to ensure that the optimized structures were located at a local minimum. The basis set superposition error (BSSE) correction developed by Simon et al.106 was included in the calculations for all complexes. Calculations without this correction were carried out to characterize the importance of this error. Negative binding energies indicate exothermic complex formation [248]. Analysis of dimer configurations showed that only the low-level symmetry configurations, C_2 and parallel-displaced (PD), represented in Figures 5a and b, respectively, correspond to minima on the PES. Vibrational calculations for the symmetric configurations exhibited negative frequencies. This outcome is in agreement with Walsh et al [246] who showed that structures of low symmetry are the most stable. The center of mass separation (r), rotation angle (Θ) for the C_2 structure, and the distance of relative translation (x_1 and x_2) for the PD structure

are given in Table 5. For these two configurations the interaction energies at the different levels of theory considered, with and without BSSE correction, as well as the first harmonics, are reported in Table 6. Interaction energies and associated wavenumbers were obtained using ω B97X-D and 6-311G, 6-311G** and cc-pVDZ basis sets. Computed values are compared with results obtained by Walsh et al [246], Saeki et al [229] and Rubes et al [247]. BSSE-correction lowers the absolute value of the energy by 1.5 kcal.mol⁻¹ on average, and represents only 0.04% of the binding energy. Thus BSSE-correction does not affect the order of stability of the tested configurations. This outcome is in full agreement with the cited prior work. Further there are no trends for the impact of basis-set and BSSE-correction on the geometry of the dimer configurations. However, by calculating the mean difference of computed values, this work, with those of Saeki et al [229] shows the 6-311G basis-set combined with BSSE-correction to be a preferred option.

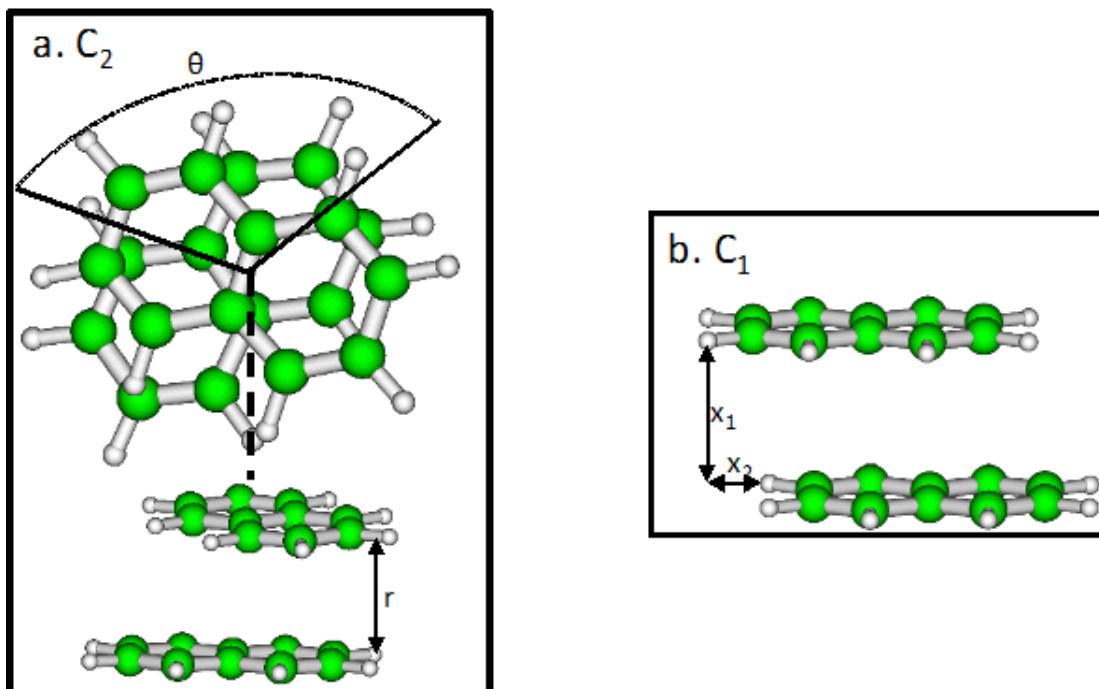


Figure 7.5: Optimized geometries at the ω B97X-D/6-311G level of theory for naphthalene dimers in the most stable configurations a) C₂, and b) Parallel-displaced (PD)

Vibration modes of naphthalene dimers: Wavenumber calculations for stable dimers were performed using the ω B97X-D/6-311G method with BSSE correction. From vibrational analysis of the monomer (Table A2 in the supplementary data) and the dimer (Table 7) the vibrational modes arising from intra- and intermolecular modes can be discriminated. The assignments are reported in Table 7 for sets of calculations performed as part of this work and

from Saeki et al[229]. Calculated values are compared with experimental data when available. The calculations were performed with counterpoise correction for both the C₂ and PD configurations. The calculated intra- and intermolecular wavenumbers, irrespective of calculation details, are readily distinguished for naphthalene. The intermolecular wavenumbers are at \sim 100 cm⁻¹ and lower, while the intramolecular wavenumbers are greater than \sim 100 cm⁻¹.

7.2.2 Pi-stacking in the acene family

Optimization at the ω B97XD/6-311G level with BSSE correction gives rise to the expected parallel tilted configuration [243] for anthracene, tetracene and pentacene dimers, as shown in Figure 6. This configuration closely resembles that found in triclinic tetracene[204] and pentacene[208] crystals. The associated interaction energies are reported Table 8. Other configurations were found to be unstable. They possess negative calculated harmonic values and were clearly not located at a minimum of the PES. So, only naphthalene has a more energetically stable C₂ configuration. Computations related to both the C₂ and PD configurations for naphthalene are shown in Table 7. The PD configuration results are used for comparisons with other acene family members where it is the most stable configuration. For a given geometry, in this case parallel-tilted, the binding interaction energy increases with the number of interacting cycles in a non-linear manner. This result is consistent with previous studies by Hohenstein et al [249] and Grimme et al[119] who compared interaction energies of naphthenic and aromatic dimers of increasing size using the B2PLYP/TZV(2d,p) level of theory. They found that arene dimers were strongly preferred over equivalent naphthenic dimers, especially for anthracene and tetracene, which were stabilized by 3 to 4 kcal.mol⁻¹. The calculations reported here are in accordance with these prior results and confirm the prevalence of π -stacking interactions.

The wavenumbers assigned to intermolecular (Table 9) and intramolecular (Table 10) modes of vibration for the acene family were calculated at the π B97XD/6-311G level of theory using BSSE correction. The characteristic wavenumbers for intermolecular vibrations fall in the same range for all acene dimers. The lower bound wavenumbers are 15 - 33 cm⁻¹ and the upper bound wavenumbers are 94 - 112 cm⁻¹. The wavenumbers for the intramolecular vibrations are found at higher wavenumbers but begin to overlap with the intermolecular vibration wavenumbers starting with anthracene where the intermolecular mode at a wavenumber of 111 cm⁻¹ with an intensity of 0.8 is not readily discriminated from the intramolecular mode at 112 cm⁻¹ with an

intensity of 1.2, on the basis of wavenumber or intensity value. Thus, inter and intramolecular vibrations are easily discriminated for naphthalene, and the inter- and intramolecular vibration modes are uncorrelated. For anthracene, tetracene and pentacene, the lowest intramolecular vibration modes fall within the range of intermolecular vibrations.

Four vibration modes, given in Table 9, corresponding to rotation, translation, rotation of the dimer combined with translation and rotation movements of one molecule relative to the other, are common for all four dimers. The last two intermolecular modes are compound-specific. For tetracene, the two additional modes comprise twisting movements of molecular pairs in the same (87 cm^{-1}) and in the opposite direction (112 cm^{-1}). For pentacene, one is a wagging movement (74 cm^{-1}) and the other comprises an opposite direction wagging movement combined with vertical translation normal to the two molecules (95 cm^{-1}). For anthracene, one mode is a rotation of one molecule relative to the other, and the second is split into two modes. These latter two modes, at 111 and 112 cm^{-1} , consist of a butterfly movement by one molecule with a half-twisting movement by the other simultaneously.

The other vibration modes, shown in Table 10 correspond to intramolecular vibrations. For every type of vibration, simultaneous vibration of each molecule is observed, either in the same or opposite direction, resulting in two wavenumbers. These alternating motions are characterized by wavenumbers and intensities that depend on the value of the dipole moment. The butterfly vibration mode presents very different wavenumbers depending on whether the molecular movements are in the same or opposite direction. Differences range in value from 8 to 42 cm^{-1} . The impact of relative movement increases with the number of aromatic rings, due to the correlation with twisting modes at higher frequencies.

Trends in values for specific bands in the acene family: In order to visualize the effect of the number of aromatic rings on the wavenumbers of specific vibration modes, intramolecular vibration modes and intermolecular vibrations are presented in Figure 7 and Figure 8, respectively. Intramolecular vibration modes, shift to lower wavenumbers with increasing molecule size irrespective of the vibrational mode considered, as was observed for monomers. Butterfly (B), Twisting (τ), Wagging (ω), Stretching (σ), and Rocking (ρ) movements behave similarly. By contrast, the intermolecular vibration modes, found at similar wavenumbers for all the molecules in the acene family, appear specific to this family, and may be used as a reliable fingerprint. In addition, twisting vibration bands appear for dimers of tetracene at (208 cm^{-1}) and pentacene at (168 cm^{-1}) that replace corresponding twisting bands for monomers at 154

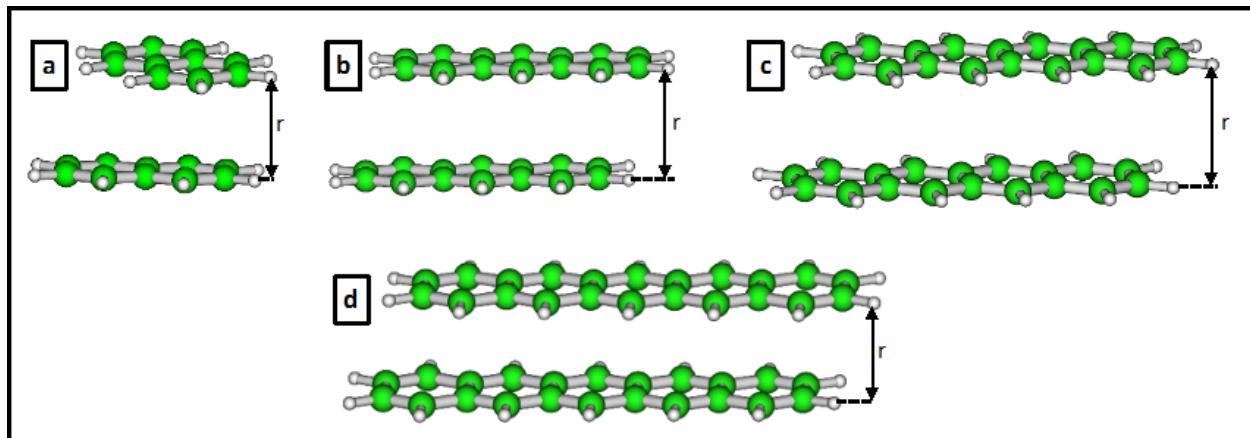


Figure 7.6: Acene dimers in the parallel displaced (PD) configuration: a) naphthalene, b) anthracene, c) tetracene, d) pentacene

and 104 cm^{-1} . The absence of a mode at 154 cm^{-1} for tetracene dimers supports the hypothesis that the feature observed experimentally at 142 cm^{-1} for monomers (see Table 3) is an overtone.

7.2.3 The impact of molecule distortion on the vibration modes of acene dimers

The impact of molecular distortion on vibration modes for dimers was evaluated in the same manner as for monomers. The molecular conformations were perturbed as illustrated in Figure 3 and the distance r (defined in Figure 6) between the molecules was set equal to 3.4, 3.4 and 2.9 Å for naphthalene, anthracene and tetracene dimers, respectively. As distortion reduces the aromaticity of the molecules, the stability of π -stacked dimers is expected to be reduced or eliminated. Table 11, where interaction energies for these systems are listed, highlights this loss of stability. For example, no conformations for distorted pentacene dimers correspond to local minima in the PES. Distortion did not affect intramolecular modes of vibration overall for dimers as shown in Table A3 in the supplementary data. Only butterfly vibration modes for anthracene and tetracene, whose wavenumbers get closer to the ones observed for monomers, shown in Table 10, appear to be affected. Some other modes, for which two wavenumbers are given in Table A3, do not exhibit the combined movement of the two molecules either in the same direction or in the opposite direction (as observed for the other intramolecular modes), but movement of the molecules simultaneously. These exceptions are a consequence of the decoupling between the molecules forming the distorted dimer. However, the intermolecular

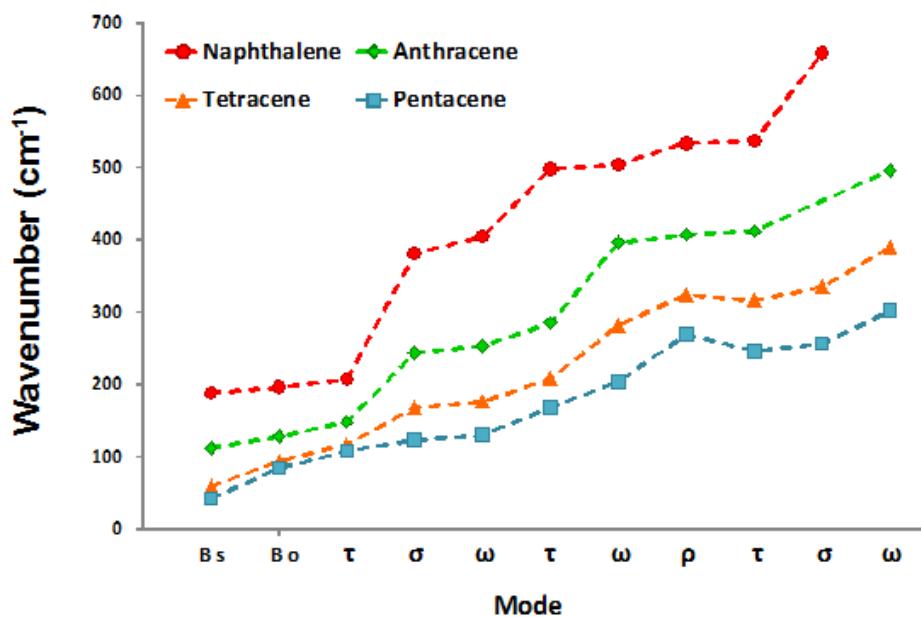


Figure 7.7: Comparison of theoretical intramolecular vibration modes for naphtalene, anthracene, tetracene and pentacene dimers: B s and B o designate butterfly modes, same (s) and opposite (o) direction respectively; τ , σ , ω , and ρ take place for twisting, stretching, wagging and rocking modes respectively

vibration modes, shown in Table 12, are shifted to lower wavenumbers compared to the symmetric analogues shown in Table 9. For the symmetric dimers, the intermolecular vibration modes are effectively independent of molecule size, while for the distorted molecules, the vibration modes tend to lower values as molecular size increases. The behavior of the symmetric dimers is attributed to a balancing of opposing effects. One would expect intermolecular vibration modes to shift to lower wavenumbers with increasing ring number (as is observed for intramolecular modes). However, the intensity of the intermolecular interactions increases with the number of rings (Table 8) for symmetric dimers, and this tends to increase wavenumbers. This compensative effect is diminished or absent for the distorted dimers, hence the difference in behavior, and wavenumber invariance for a given intermolecular vibration mode may be a characteristic of π -stacking interactions.

7.2.4 Identification of unassigned peaks in the experimental spectra of tetracene and pentacene

By subtracting the intermolecular wavenumbers from the symmetric dimer vibrational spectra of tetracene and pentacene, and comparing the resultant spectra with the spectra of the corres-

153. Calculation of vibrational spectra of acene dimers: Comparison with experimental data.

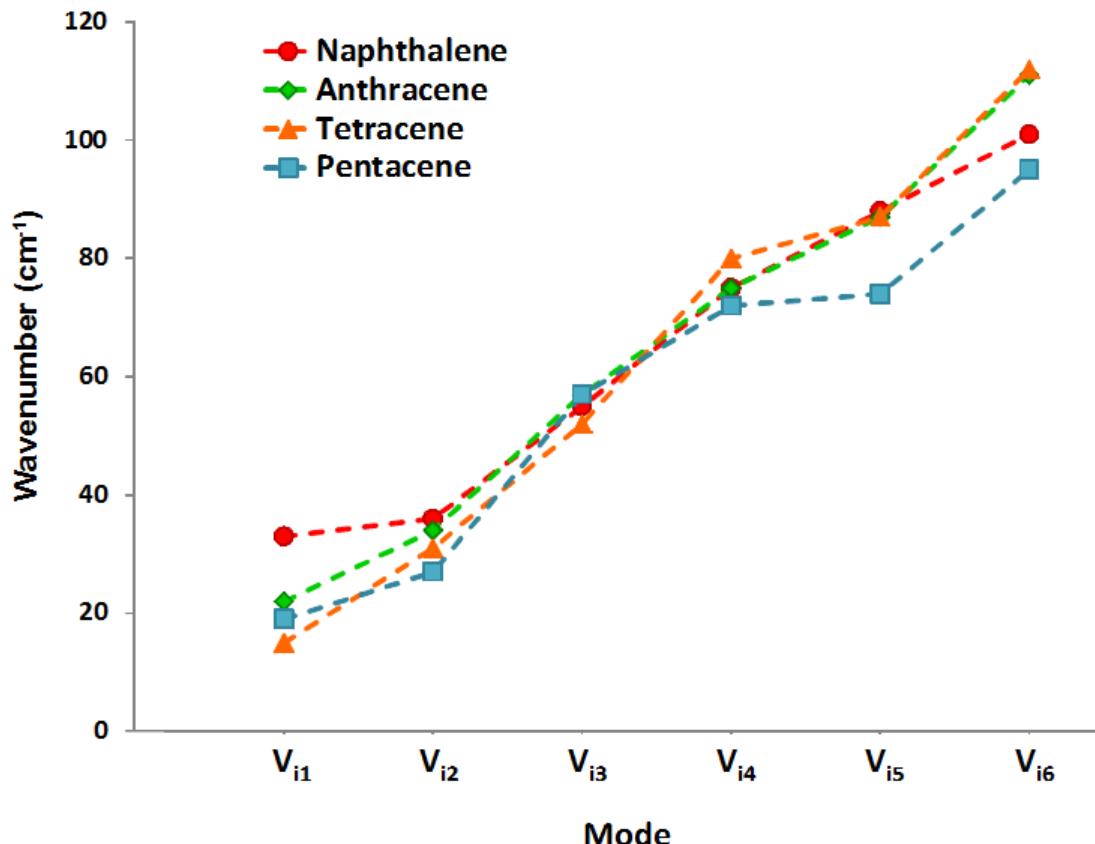


Figure 7.8: Comparison of theoretical intermolecular vibration modes for naphthalene, anthracene, tetracene and pentacene dimers

ponding monomers, there are two peaks, one for tetracene at 117 cm^{-1} and one for pentacene at 108 cm^{-1} , with intensities in the same order of magnitude as the ones calculated for fundamental bands for tetracene (0.34 km/mol) and pentacene (0.15 km/mol). These peaks arise because for every vibrational mode occurring for monomers, there are two quasi-degenerate modes arising for dimers depending on whether the molecules vibrate in the same or in the opposite direction. Thus modes that are Raman-active for a monomer become infrared-active for dimers vibrating in the opposite direction.

With this final set of assignments, the inter- and intramolecular vibration analysis for the acene family of compounds is complete. Figure 9 gives a direct comparison of experimental and calculated spectra with our conclusion as for assignment. There are no residual unassigned peaks in the available experimental spectra and ultra low wavenumber intermolecular peaks are hypothesized. While these have yet to be measured experimentally, identification and discrimination of inter- and intramolecular bands at low wavenumbers will play a more important role in the future as the range of IR measurements broadens. For example, by using coherent

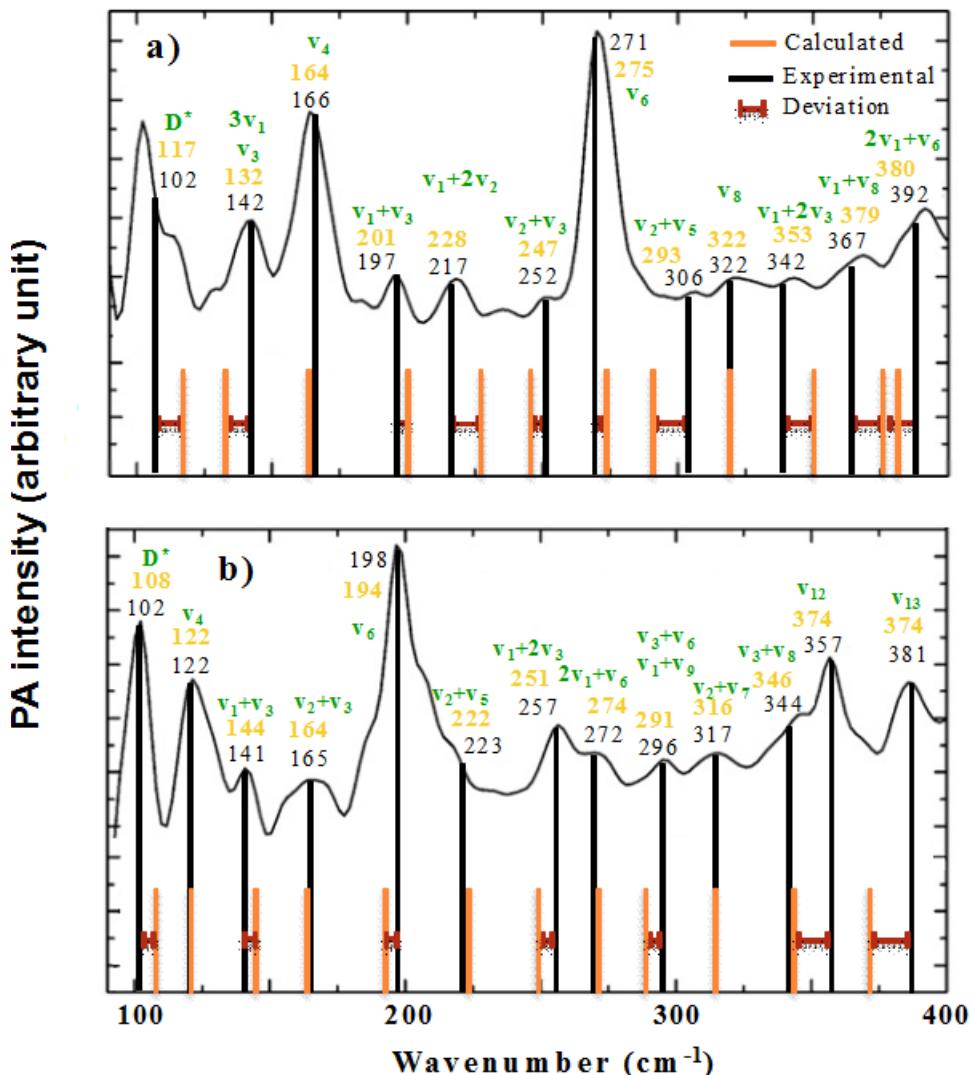


Figure 7.9: Final assignment of the experimental infrared transitions for a) tetracene and b) pentacene. Calculated wavenumbers are indicated on the spectra with orange lines, and corresponding values above. Experimental values are indicated for comparison, and represented with dark lines

synchrotron radiation it is becoming possible to make experimental photoacoustic IR measurements in the range 7 to 30 cm^{-1} as demonstrated recently by Billinghamurst and Michaelian[250]. DFT calculations have much to contribute to this development, and the interpretation of results obtained.

7.3 Normal Vibrational Modes in Crystals

7.3.1 Basic theory

The behavior of a system of interacting electrons and nuclei is determined by the solutions of the time-dependent Schrödinger equation:

where V is the potential describing the coulombian interactions. The Born-Oppenheimer (or adiabatic) approximation (valid for $\hbar \ll m_e c$) allows to split the problem into an electronic problem depending upon nuclear positions :

where and

I.E.

and a nuclear problem under an effective interatomic potential determined by the electrons:

determines the Potential Energy Surface and the equilibrium geometry. At equilibrium, forces on nuclei vanish:

Harmonic approximation: the interatomic potential energy is expended to 2nd order.

The resulting Hamiltonian transform into a sum of independent oscillators.

Normal mode frequencies, and displacement patterns, for Cartesian component of atom, at atomic position are determined by secular equation:

where K is the matrix of inter-atomic force constants, i.e. second derivatives of the energy with respect to atomic positions:

In crystal, normal modes are classified by a wave-vector \mathbf{k} . Phonon frequencies, ω , and displacement patterns, \mathbf{u} , are determined by the secular equations:

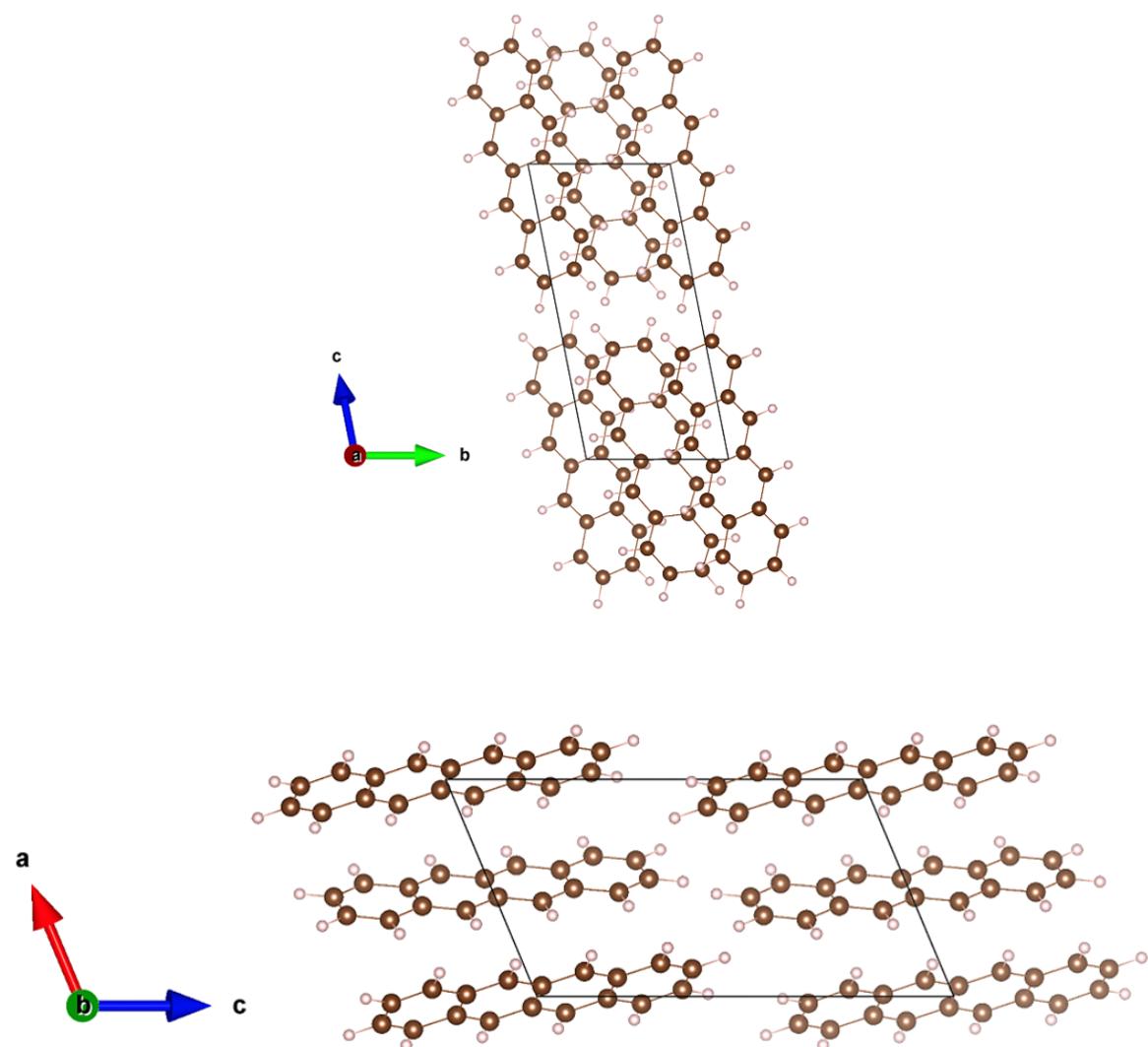
Introduce monochromatic perturbation to atomic positions as

where \mathbf{r}_i represents the lattice vector and the equilibrium position of the i -th atom in the unit cell. Fourier transform of force constants at \mathbf{k} are second derivatives of the energy with respect to such monochromatic perturbations:

7.3.2 Results and discussion

We performed first-principle calculations for tetracene in order to clarify its structural and electronic properties of the solid-state system. These calculations strongly suggest the necessity to go beyond the standard DFT approach. First, we have shown that dispersion corrected methods are necessary to provide an improvement in the calculated lattice parameters and volume

of the unit cell to compare with experimental values. The complete list of resulting lattice parameters, angles, and volumes along with experimental results for all the compounds are presented in Table 13. Furthermore, the mean absolute deviations observed for the calculation of these parameters for all the methods under consideration depend on the type of dispersion correction, and the relative errors on structural properties calculated using each method, in particular concerning the volume, may vary significantly. As mentioned by Lebegue et al [251], it has been demonstrated that one has to test all the dispersion corrected DFT methods for the C-, H-, N- and O- based energetic solids before concluding the ground state properties. In the present work, we show that the situation is the same for the S-based energetic solids. Since the D3 correction yields the most accurate structural descriptions of this model system, it has been used to calculate the vibrational, structural and electronic properties.



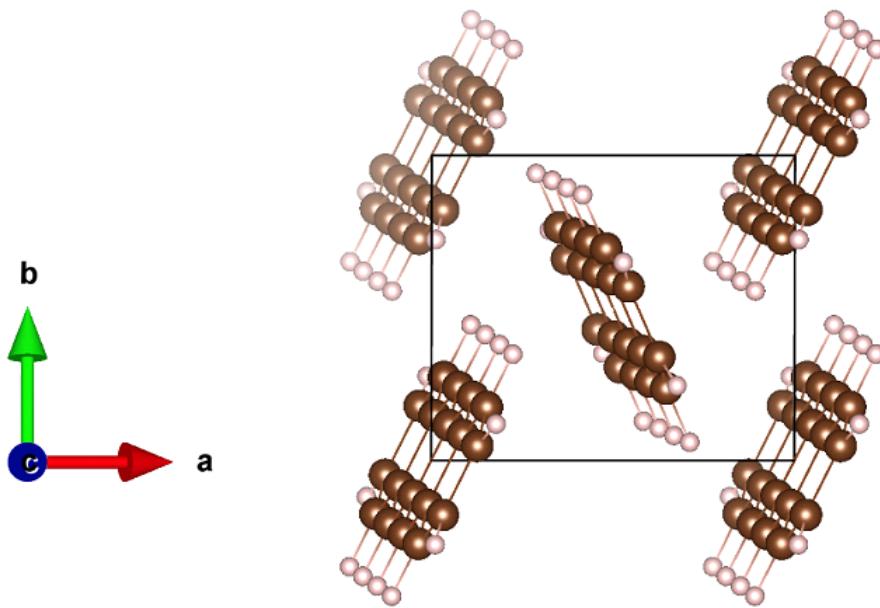


Table 14 reports all the results obtained for tetracene, from the isolated up to the solid-state system. Beyond the numerical differences due to the various Hamiltonians used for these calculations, the results show significant differences on both vibration frequencies and intensities that are to be attributed to the effect of the chemical environment. Indeed, the signatures of some particular modes (such as ν_2 , ν_8 and ν_{10}) are modified with respect to the isolated system due to intermolecular interactions. Besides, our previous study on tetracene dimers revealed that intermolecular interactions induced the appearance of some characteristic spectral signatures at very low wavenumbers. For instance, the mode experimentally found at 106 cm^{-1} , which has been assigned, with aid of calculations, to a libration mode, is part of the characteristic signatures of intermolecular vibrations of acenes; the same mode was also assigned at 102 cm^{-1} for the pentacene system. In fact, this mode cannot be only associated to a dimer: it results from the vibrations of the whole tetracene molecules within the network that constitutes the aggregate or the crystalline system. Our VASP calculations clearly show that this vibration mode is linked to distortions of the crystalline network, which can be seen as the « breathing » of the network. Such a deformation was already partially taken into account in the study of a dimer. Some other modes of the crystalline network constitute, to a lesser extent, the remaining characteristic signatures of the intermolecular vibrations. More difficult to detect experimentally, these modes are theoretically found at 69.7 cm^{-1} (ν_1), 128.5 cm^{-1} (ν_3) and $320.8\text{--}323.8\text{ cm}^{-1}$ (ν_8). Unfortunately, there is no experimental data available below 100 cm^{-1} for the

tetracene system. But it seems that the ν_3 and ν_8 modes can now be identified as the modes of medium intensity reported from experimental data at 142 and 322 cm^{-1} , respectively. No fundamental mode nor combination mode of the isolated tetracene can be associated with the presence of those new experimental signatures.

From all the modes calculated below 400 cm^{-1} , whatever their intensities are, they do not seem to be coupled to the vibration of the crystalline network. It suggests that calculations of the isolated molecule (or its dimer) are a sufficient approximation to identify these vibration modes. In the present case, it explains the fact that no assignation problem was found for regions beyond 150 cm^{-1} during the molecular study of tetracene and its dimer. At the molecular level, calculations on small aggregates can be sufficient to identify the modes of the crystalline network that are significantly impacted.

To conclude, the use of a model able to describe the chemical environment was essential to identify the spectral signatures of the inter-molecular modes that we try to characterize in our study of asphaltenes (or, at least, of acenes systems). The calculated phonon spectrum thus seems to answers, even if our calculations up to now do not account for both mechanical and electrical anharmonicities. Such conditions are all the more necessary for an exhaustive study of the modes of each isolated molecule. Particularly, it allows an identification of the combination modes as well as the harmonics for which the signatures have been clearly evidenced by photo-acoustic measurements. Only a double approach (both molecular and solid-state) can allow one to completely identify the photo-acoustic results of acenes molecules in the crystalline state.

Conclusions

The DFT calculations presented in this work provide detailed and accurate descriptions of experimental photoacoustic far IR spectra for tetracene and pentacene and for far IR spectra of the acene family as a whole. The impacts of molecular distortion on monomer and dimer spectra and dimer stability, and the interpretation of peaks in photoacoustic IR spectra of crystalline solids are discussed in detail. Distortion breaks molecular symmetry and this adds to the complexity of the interpretation of the resulting spectra. Inter- and intramolecular vibrations in the acene family of compounds are discriminated, and combinations of inter- and intramolecular vibrations are identified specially with the use of solid-state conditions. Peak assignment ambiguity, and the potential impacts of molecular distortion on peak intensity are addressed

for tetracene and pentacene. By examining computed and experimental spectra for the family of acene molecules concurrently, trends in vibration modes with molecular size were used to reinforce assignments, and to reattribute assignments made previously in the experimental tetracene and pentacene spectra. In particular, both dimer and solid state calculations permitted re-assignment of an infrared active peak arising in tetracene at 142 cm^{-1} and assignment of large infrared active peaks at 117 and 108 cm^{-1} for tetracene and pentacene respectively that may permit their identification in an acene mixture. The intermolecular vibrations for naphthalene, anthracene, tetracene and pentacene, are found in the same wavelength range of the far infrared spectrum and are not readily discriminated one from the other. The specificity of the acene π -stacking interactions, reinforced by calculations with distorted dimers (and solid-state), allows not only the identification of this family of molecules, but also highlights the presence of π -stacking interaction in a mixture as a whole.

Supplementary Data

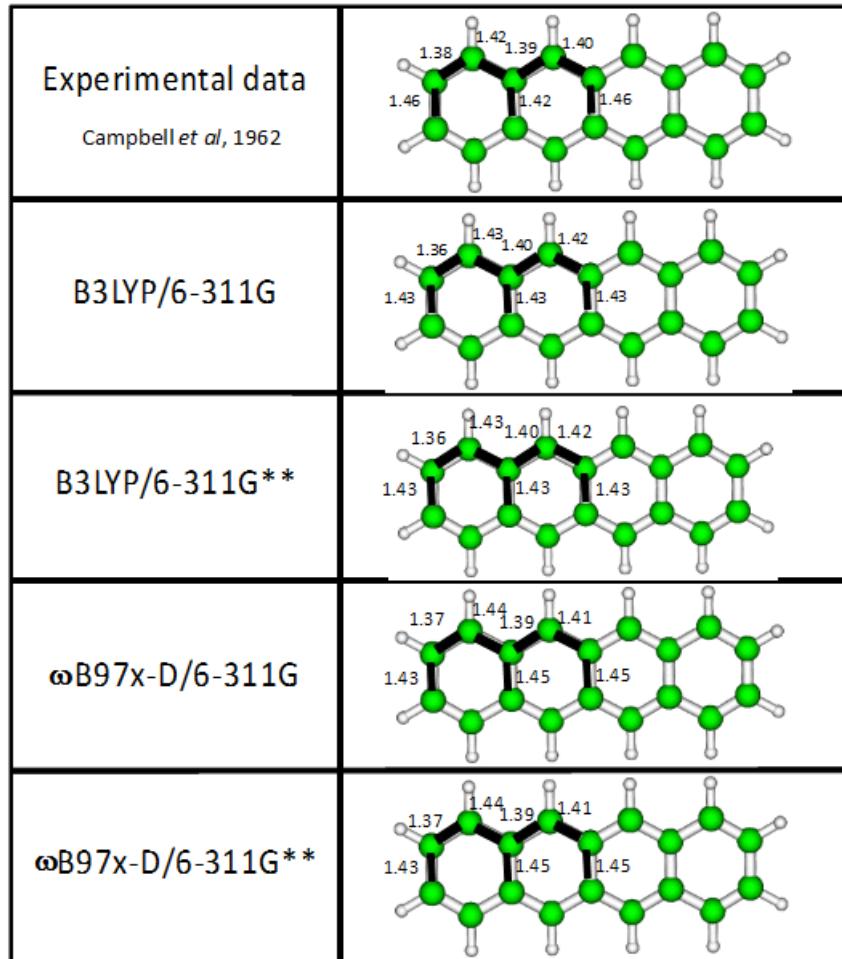


Figure 7.10: Geometrical data for a tetracene molecule by Campbell et al [204] and values optimized in the framework of various DFT calculations, used in this work

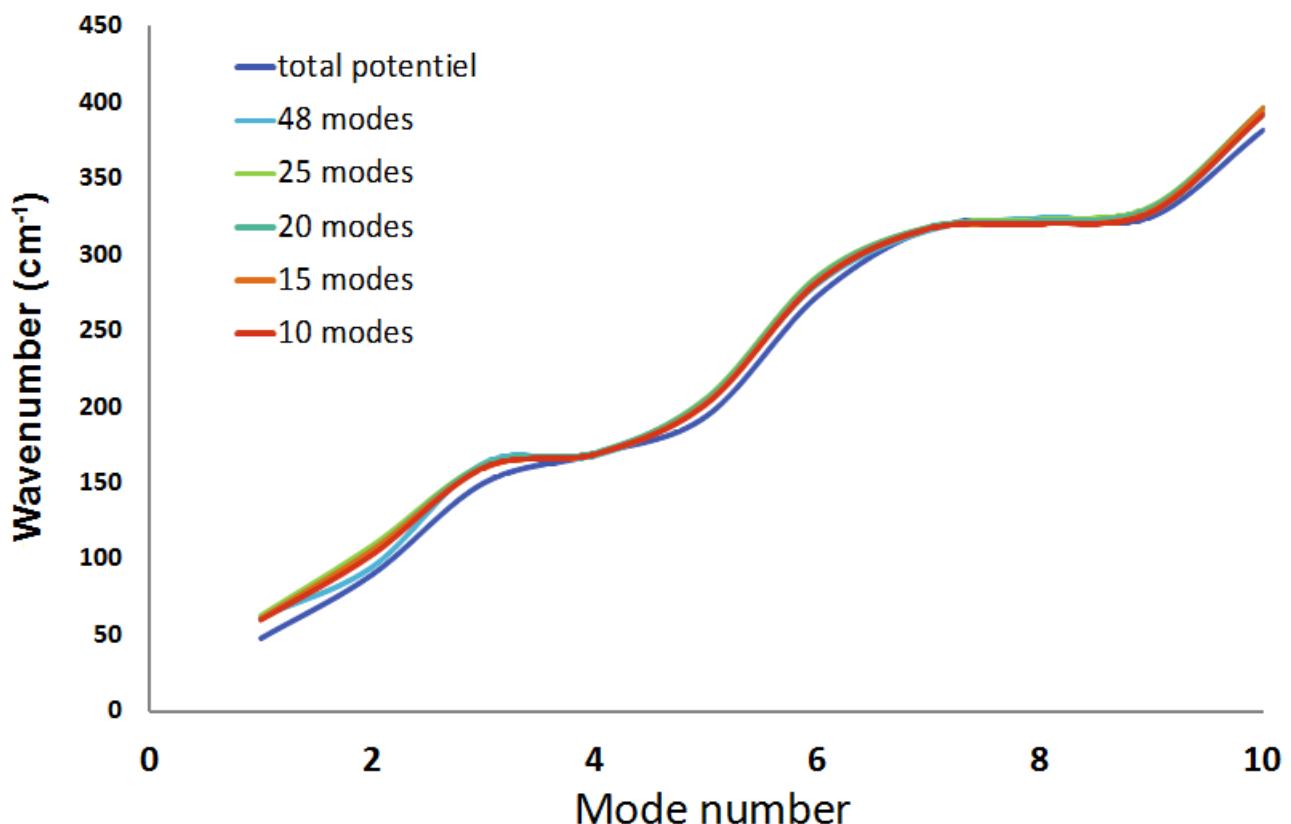


Figure 7.11: Effect of mode truncature on the wavenumber obtained using ω B97X-D/6-311G for the first ten vibration modes of a tetracene molecule, with force truncature ($f > 5\text{cm}^{-1}$)

Table 7.2: Calculated vibrational frequencies (cm^{-1}) of the monomer, dimer and solid-state (PBE tetracene system.

	Monomer	Dimer	Experimental (P.A.) this work	VASP/CRYSTAL
Assignment	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$	$\omega(\text{cm}^{-1})$ Int(relative)
ν_1	57 (0.8)	15 (<0.01)		
		31 (0)		
		52 (0)		
		59 (0.6)		69.7 (8.2)
		80 (1.0)		
		87 (0.1)		
ν_2	93 (0)	94 (0)		
		112 (0)		
		124 (0)	106 (m) + sh	102.4 (2.1) 105.7 (9.5)
$3\nu_1$	132 (<0.01)			
ν_3	153 (0.01)		142 (m)	128.5 (7.1) 148 (0.1)
ν_4	164 (1.3)	162 (0)		
		168 (1.7)	166 (s) + sh	164.6 (15.5)
		169 (0)		
ν_5	196 (0.05)	176 (0.2)	197 (w)	170.1 (0.1)
$\nu_1 + \nu_3$	201 (<0.01)	208 (0)		
		212 (0.09)	217 (w)	
$\nu_1 + 2\nu_2$	228 (<0.01)			
$\nu_2 + \nu_3$	247 (0.01)		252 (vw)	
ν_6	275 (1.03)	282 (1.6)	271 (m)	267.4 (3.0)
		286 (0)		274.6 (2.6)
$\nu_2 + \nu_5$	293 (<0.01)		306 (vw)	
		316 (0)		
		316 (0.03)		
		324 (0)	322 (w)	320.8 (0.9)
ν_8	322 (<0.01)	325 (0.06)		323.8 (0.5)
		335 (0.03)	342 (w)	

Table 7.3: Computed wavenumbers and intensities for distorted tetracene based on the harmonic approximation at the ω B97X-D/6-311G level of calculation. Comparison with symmetric tetracene (D_{2h})

Mode	Tetracene with D_{2h} symmetry		Distorted Tetracene	
	Harmonic wavenumber (cm ⁻¹)	Harmonic Intensity (km.mol ⁻¹)	Harmonic wavenumber (cm ⁻¹)	Harmonic Intensity (km.mol ⁻¹)
ν_1	57	0.86	58	0.87
ν_2	93	0	93	0.03
ν_3	154	0	154	0.01
ν_4	168	1.24	168	1.22
ν_5	196	0	197	0.05
ν_6	276	1.03	277	1.06
ν_7	316	0	316	0.002
ν_8	324	0	324	0.002
ν_9	328	0	328	0.02
ν_{10}	389	0	389	0.04

Table 7.4: Calculated and Experimental wavenumbers (cm-1)[124] of the vibrational modes of a naphthalene molecule. Values in parenthesis represent the deviation between the calculated and experimental wavenumbers.

Mode	Symmetry	B3LYP/6-311G		ω B97X-D/6-311G		Saeki et al[229]	Exp[252]
		Harmonic	PT2	Harmonic	PT2		
ν_{32}	b_{1g}	177(+1)	176(0)	177(+1)	176(0)	168(-8)	176
ν_{28}	a_{1u}	189(-6)	186(-9)	190(-5)	187(-8)	182(-13)	195
ν_{40}	b_{3u}	373(+14)	372(+13)	379(+20)	377(+18)	356(-3)	359
ν_{20}	b_{1u}	401(+15)	394(+8)	404(+18)	398(+12)	377(-9)	386

Table 7.5: Intramolecular vibration modes (cm^{-1}) and the corresponding intensities in brackets (km/mol) for distorted naphthalene, anthracene and tetracene dimers. Values of corresponding dimer modes of molecules in D_{2h} configuration are written below in italics.

Type	Naphthalene	Anthracene	Tetracene
Butterfly sw	190(8.8)	101(0.003)	77(1.6)
	<i>188(3.8)</i>	<i>111(0.8)/112(1.2)</i>	<i>59(0.6)</i>
Butterfly ow	183(0.03)	103(4.2)	54(0.005)
	<i>196(0)</i>	<i>128(0)</i>	<i>94(0)</i>
Twisting sw	206(0.2)	137(0.05)	106(0.4)
	<i>207(1.2)</i>	<i>148(0.6)</i>	<i>117(0.3)</i>
Twisting ow	208(0.06)	141(0.005)	102(0.04)
	<i>208(0)</i>	<i>149(0.02)</i>	<i>124(0)</i>
Scissoring sw	382(2.1)	244(1.9)	168(1.8)
	<i>381(2.2)</i>	<i>244(2.05)</i>	<i>168(1.7)</i>
Scissoring ow	382(0.2)	243(0.2)	168(0.1)
	<i>380(0)</i>	<i>243(0)</i>	<i>169(0)</i>
Wagging sw	411(1.9)	251(0.3)	164(0.03)
	<i>405(0)</i>	<i>250(0)</i>	<i>162(0)</i>
Wagging ow	410(0.03)	248(0.2)	154(0.1)
	<i>405(0.02)</i>	<i>253(0.17)</i>	<i>176(0.2)</i>
Twisting sw	502(0.1)	282(0.05)	203(0.03)
	<i>498(0)</i>	<i>285(0.02)</i>	<i>208(0)</i>
Twisting ow	503(0.3)	281(0.003)	201(0.2)
	<i>499(15.2)</i>	<i>286(0.08)</i>	<i>212(0.09)</i>
Wagging sw	492(59)	400(0.01)	281(2.2)
	<i>504(45.0)</i>	<i>396(0.2)</i>	<i>282(1.55)</i>
Wagging ow	495(1.2)	398(0.01)	279(0.2)
	<i>507(0)</i>	<i>396(0.04)</i>	<i>286(0)</i>
Stretching sw	533(0.02)	407(0.007)	324(0.007)
	<i>534(0)</i>	<i>407(0)</i>	<i>324(0)</i>
Stretching ow	532(0.01)	407(0.02)	324(0.07)
	<i>534(0.09)</i>	<i>407(0.04)</i>	<i>325(0.06)</i>
Rocking sw	536(0.1)/536(0.2) ¹	413(0.03)	316(0.01)/316(0.004) ¹
	<i>537(0)</i>	<i>412(0)</i>	<i>316(0)</i>
Rocking ow	536(0.1)/536(0.2) ¹	413(0.009)	316(0.01)/316(0.004) ¹
	<i>537(0.06)</i>	<i>412(0.01)</i>	<i>316(0.03)</i>
Twisting sw	660(0.06)		331(0.03)/331(0.02) ¹
	<i>658(0.4)</i>		<i>335(0.03)</i>
Twisting ow	660(0.007)		331(0.03)/331(0.02) ¹
	<i>659(0)</i>		<i>336(0)</i>
Wagging sw		487(66.8)	389(0.02)
		<i>496(70.5)</i>	<i>390(0)</i>
Wagging ow		490(1.07)	389(0.02)
		<i>497(0.5)</i>	<i>390(0.02)</i>

¹The vibrational modes for which two values are indicated correspond to modes for which one molecule vibrates at a time

Appendices

Références bibliographiques

- [1] Kamran Akbarzadeh, Ahmed Hammami, Abdel Kharrat, Dan Zhang, Stephan Allenson, Jefferson Creek, Shah Kabir, A Jamaluddin, Alan G Marshall, Ryan P Rodgers, et al. Asphaltenes—problematic but rich in potential. *Oilfield Review*, 19(2):22–43, 2007.
- [2] Matthew Neurock, Abhash Nigam, Daniel Trauth, and Michael T Klein. Molecular representation of complex hydrocarbon feedstocks through efficient characterization and stochastic algorithms. *Chemical engineering science*, 49(24):4153–4177, 1994.
- [3] Long-li ZHANG, Chun-lan WANG, Yuan-sheng ZHAO, Guo-hua YANG, SU Mei, and Chao-he YANG. Speciation and quantification of sulfur compounds in petroleum asphaltenes by derivative xanes spectra. *Journal of Fuel Chemistry and Technology*, 41(11):1328–1335, 2013.
- [4] Roberto Rodrigues Coelho, Ian Hovell, and Krishnaswamy Rajagopal. Elucidation of the functional sulphur chemical structure in asphaltenes using first principles and deconvolution of mid-infrared vibrational spectra. *Fuel processing technology*, 97:85–92, 2012.
- [5] Lamia Goual. *Petroleum asphaltenes*. INTECH Open Access Publisher, 2012.
- [6] Otto P Strausz, Thomas W Mojelsky, and Elizabeth M Lown. The molecular structure of asphaltene: an unfolding story. *Fuel*, 71(12):1355–1363, 1992.
- [7] José A Calles, Javier Dufour, Javier Marugán, José Luis Peña, Raúl Giménez-Aguirre, and Daniel Merino-García. Properties of asphaltenes precipitated with different n-alkanes. a study to assess the most representative species for modeling†. *Energy & Fuels*, 22(2):763–769, 2007.

- [8] M Barbelet, JM Dolle, and JP Poitevin. Analyses élémentaires (carbone, hydrogène, oxygène, azote, soufre) des fractions lourdes du pétrole. *Revue de l'Institut Français du Pétrole*, 34(2):293–306, 1979.
- [9] James G Speight and Speros E Moschopedis. Some observations on the molecular nature of petroleum asphaltenes. *Am. Chem. Soc., Div. Pet. Chem., Prepr.;(United States)*, 24(CONF-790917-(Vol. 24)(No. 4)), 1979.
- [10] I Merdrignac and D Espinat. Physicochemical characterization of petroleum fractions: the state of the art. *Oil & Gas Science and Technology-Revue de l'IFP*, 62(1):7–32, 2007.
- [11] JG Speight. Petroleum asphaltenes-part 1: Asphaltenes, resins and the structure of petroleum. *Oil & gas science and technology*, 59(5):467–477, 2004.
- [12] Oliver C Mullins, Bruno Martínez-Haya, and Alan G Marshall. Contrasting perspective on asphaltene molecular weight. this comment vs the overview of aa herod, kd bartle, and r. kandiyoti. *Energy & Fuels*, 22(3):1765–1773, 2008.
- [13] Isabelle Merdrignac, Anne-Agathe Quoineaud, and Thierry Gauthier. Evolution of asphaltene structure during hydroconversion conditions. *Energy & fuels*, 20(5):2028–2036, 2006.
- [14] Henning Groenzin and Oliver C Mullins. Molecular size and structure of asphaltenes from various sources. *Energy & Fuels*, 14(3):677–684, 2000.
- [15] S Zhao, LS Kotlyar, JR Woods, BD Sparks, K Hardacre, and KH Chung. Molecular transformation of athabasca bitumen end-cuts during coking and hydrocracking. *Fuel*, 80(8):1155–1163, 2001.
- [16] E Rogel and L Carbognani. Density estimation of asphaltenes using molecular dynamics simulations. *Energy & fuels*, 17(2):378–386, 2003.
- [17] Keith L Gawrys, P Matthew Spiecker, and Peter K Kilpatrick. The role of asphaltene solubility and chemical composition on asphaltene aggregation. *Petroleum science and technology*, 21(3-4):461–489, 2003.

- [18] Jeff M Sheremata, Murray R Gray, Heather D Dettman, and William C McCaffrey. Quantitative molecular representation and sequential optimization of athabasca asphaltenes. *Energy & Fuels*, 18(5):1377–1384, 2004.
- [19] P DuBois Murphy, BC Gerstein, Victoria L Weinberg, and TF Yen. Determination of chemical functionality in asphaltenes by high-resolution solid-state carbon-13 nuclear magnetic resonance spectrometry. *Analytical Chemistry*, 54(3):522–525, 1982.
- [20] Teh Fu Yen, Wen Hui Wu, and George V Chilingar. A study of the structure of petroleum asphaltenes and related substances by infrared spectroscopy. *Energy Sources*, 7(3):203–235, 1984.
- [21] Emmanuelle Durand, Martin Clemancey, Jean-Marc Lancelin, Jan Verstraete, Didier Espinat, and Anne-Agathe Quoineaud. Effect of chemical composition on asphaltenes aggregation. *Energy Fuels*, 24(2):1051–1062, 2010.
- [22] Teh Fu Yen, J Gordon Erdman, and Sidney S Pollack. Investigation of the structure of petroleum asphaltenes by x-ray diffraction. *Analytical Chemistry*, 33(11):1587–1594, 1961.
- [23] John W Shirokoff, Mohammad N Siddiqui, and Mohammad F Ali. Characterization of the structure of saudi crude asphaltenes by x-ray diffraction. *Energy & Fuels*, 11(3):561–565, 1997.
- [24] KH Altgelt and MM Boduszynski. composition and analysis of heavy petroleum fractions. 1994.
- [25] Henning Groenzin and Oliver C Mullins. Asphaltene molecular size and structure. *The Journal of Physical Chemistry A*, 103(50):11237–11245, 1999.
- [26] Rodrigo da Silva Souza, David E Nicodem, Simon J Garden, and Rodrigo J Corrêa. Study of the asphaltene aggregation structure by time-resolved fluorescence spectroscopy. *Energy & Fuels*, 24(2):1135–1138, 2009.
- [27] Yuan Hongfu, Chu Xiaoli, Li Haoran, and Xu Yupeng. Determination of multi-properties of residual oils using mid-infrared attenuated total reflection spectroscopy. *Fuel*, 85(12):1720–1728, 2006.

- [28] AA El-Bassoussi, MHM Ahmed, SM El Sayed, JS Basta, and E-SK Attia. Characterization of some local petroleum residues by spectroscopic techniques. *Petroleum Science and Technology*, 28(5):430–444, 2010.
- [29] R Rodrigues Coelho, I Hovell, E Lopez Moreno, A Lopes de Souza, and K Rajagopal. Characterization of functional groups of asphaltenes in vacuum residues using molecular modelling and ftir techniques. *Petroleum science and technology*, 25(1-2):41–54, 2007.
- [30] Jérémie Laxalde, Noémie Caillol, François Wahl, Cyril Ruckebusch, and Ludovic Duponchel. Combining near and mid infrared spectroscopy for heavy oil characterisation. *Fuel*, 133:310–316, 2014.
- [31] Sócrates Acevedo, Optalí Escobar, Lorenzo Echevarria, Luis B Gutiérrez, and Bernardo Méndez. Structural analysis of soluble and insoluble fractions of asphaltenes isolated using the pnp method. relation between asphaltene structure and solubility. *Energy & fuels*, 18(2):305–311, 2004.
- [32] Luis B Gutiérrez, María A Ranaudo, Bernardo Méndez, and Sócrates Acevedo. Fractionation of asphaltene by complex formation with p-nitrophenol. a method for structural studies and stability of asphaltene colloids. *Energy & Fuels*, 15(3):624–628, 2001.
- [33] Toshimasa Takanohashi, Shinya Sato, and Ryuzo Tanaka. Structural relaxation behaviors of three different asphaltenes using md calculations. *Petroleum science and technology*, 22(7-8):901–914, 2004.
- [34] ME Elyashberg, AJ Williams, and GE Martin. Computer-assisted structure verification and elucidation tools in nmr-based structure elucidation. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 53(1):1–104, 2008.
- [35] R Todeschini, P Gramatica, R Provenzani, and E Marengo. Weighted holistic invariant molecular descriptors. part 2. theory development and applications on modeling physico-chemical properties of polyaromatic hydrocarbons. *Chemometrics and Intelligent Laboratory Systems*, 27(2):221–229, 1995.
- [36] Jean-Loup Faulon. Stochastic generator of chemical structure. 2. using simulated annealing to search the space of constitutional isomers. *Journal of Chemical Information and Computer Sciences*, 36(4):731–740, 1996.

- [37] Hassan K Al Halwachi, Dmitry S Yakovlev, and Edo S Boek. Systematic optimization of asphaltene molecular structure and molecular weight using the quantitative molecular representation approach. *Energy & Fuels*, 26(10):6177–6185, 2012.
- [38] Luís Pereira de Oliveira, Jan J Verstraete, and Max Kolb. A monte carlo modeling methodology for the simulation of hydrotreating processes. *Chemical engineering journal*, 207:94–102, 2012.
- [39] Juan Murgich, Jesús Rodríguez, and Yosslen Aray. Molecular recognition and molecular mechanics of micelles of some model asphaltenes and resins. *Energy & Fuels*, 10(1):68–76, 1996.
- [40] Fernando Alvarez-Ramírez and Yosadara Ruiz-Morales. Island versus archipelago architecture for asphaltenes: polycyclic aromatic hydrocarbon dimer theoretical studies. *Energy & Fuels*, 27(4):1791–1808, 2013.
- [41] Zhenyu Huang, Hyun Su Lee, Michael Senra, and H Scott Fogler. A fundamental model of wax deposition in subsea oil pipelines. *AICHE Journal*, 57(11):2955–2964, 2011.
- [42] D Broseta, M Robin, T Savvidis, C Féjean, M Durandieu, H Zhou, et al. Detection of asphaltene deposition by capillary flow measurements. In *SPE/DOE Improved Oil Recovery Symposium*. Society of Petroleum Engineers, 2000.
- [43] Sunil Kokal, Jamal Al-Juraid, et al. Quantification of various factors affecting emulsion stability: watercut, temperature, shear, asphaltene content, demulsifier dosage and mixing different crudes. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 1999.
- [44] Erwin Schrödinger. La situation actuelle en mécanique quantique. *Physique quantique et représentation du monde*, page 138, 1992.
- [45] Gordon E Moore. Electronics magazine. *April 19th*, 1965.
- [46] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.

- [47] Carlo Adamo and Vincenzo Barone. Toward reliable density functional methods without adjustable parameters: The pbe0 model. *The Journal of chemical physics*, 110(13):6158–6170, 1999.
- [48] Carlo Adamo and Vincenzo Barone. Toward reliable adiabatic connection models free from adjustable parameters. *Chemical Physics Letters*, 274(1):242–250, 1997.
- [49] Axel D Becke. Density-functional thermochemistry. iii. the role of exact exchange. *The Journal of chemical physics*, 98(7):5648–5652, 1993.
- [50] Axel D Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical review A*, 38(6):3098, 1988.
- [51] Lee Chengteh, Yang Weitao, and RG Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical Review B (Condensed Matter)*, 37(2):785–789, 1988.
- [52] A Daniel Boese and Jan ML Martin. Development of density functionals for thermochemical kinetics. *The Journal of chemical physics*, 121(8):3405–3416, 2004.
- [53] Chr Møller and Milton S Plesset. Note on an approximation treatment for many-electron systems. *Physical Review*, 46(7):618, 1934.
- [54] Axel D Becke. Density-functional thermochemistry. v. systematic optimization of exchange-correlation functionals. *The Journal of chemical physics*, 107(20):8554–8560, 1997.
- [55] Fred A Hamprecht, Aron J Cohen, David J Tozer, and Nicholas C Handy. Development and assessment of new exchange-correlation functionals. *The Journal of chemical physics*, 109(15):6264–6271, 1998.
- [56] Philip J Wilson, Thomas J Bradley, and David J Tozer. Hybrid exchange-correlation functional determined from thermochemical data and ab initio potentials. *The Journal of Chemical Physics*, 115(20):9233–9242, 2001.
- [57] Thomas W Keal and David J Tozer. Semiempirical hybrid functional with improved performance in an extensive chemical assessment. *The Journal of chemical physics*, 123(12):121103, 2005.

- [58] Hisayoshi Iikura, Takao Tsuneda, Takeshi Yanai, and Kimihiko Hirao. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *The Journal of Chemical Physics*, 115(8):3540–3544, 2001.
- [59] Jeng-Da Chai and Martin Head-Gordon. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Physical Chemistry Chemical Physics*, 10(44):6615–6620, 2008.
- [60] PJ Krusic, E Wasserman, BA Parkinson, B Malone, ER Holler Jr, PN Keizer, JR Morton, and KF Preston. Electron spin resonance study of the radical reactivity of c60. *Journal of the American Chemical Society*, 113(16):6274–6275, 1991.
- [61] Pierre Barchewitz. *Spectroscopie atomique et moléculaire: Spectroscopie moléculaire*, volume 2. Masson, 1971.
- [62] E Bright Wilson Jr, John G Decius, Paul G Cross, and RT Lagemann. Molecular vibrations. *American Journal of Physics*, 23(8):550–550, 1955.
- [63] Claude Pouchan. *Approche de la fonction potentielle nucléaire à partir du calcul de fonction d'onde: application à l'étude des modes de vibrations moléculaires*. PhD thesis, 1978.
- [64] Khalil Zaki. *Etude des modes de vibration au-delà de l'approximation harmonique: Application à la résolution des problèmes vibrationnels de molécules de type X2CNX(X= F, H)*. PhD thesis, 1996.
- [65] Carl Eckart. Some studies concerning rotating axes and polyatomic molecules. *Physical Review*, 47(7):552, 1935.
- [66] Hendrik Brugt Gerhard Casimir. Rotation of a rigid body in quantum mechanics. 1931.
- [67] Peter Pulay, Geza Fogarasi, Frank Pang, and James E Boggs. Systematic ab initio gradient calculation of molecular geometries, force constants, and dipole moment derivatives. *Journal of the American Chemical Society*, 101(10):2550–2560, 1979.
- [68] Gary Simons, Robert G Parr, and J Michael Finlan. New alternative to the dunham potential for diatomic molecules. *The Journal of Chemical Physics*, 59(6):3229–3234, 1973.

- [69] Wilfried Meyer, Peter Botschwina, and Peter Burton. Ab initio calculation of near-equilibrium potential and multipole moment surfaces and vibrational frequencies of h+ 3 and its isotopomers. *The Journal of chemical physics*, 84(2):891–900, 1986.
- [70] Paul E Maslen, Dylan Jayatilaka, Susan M Colwell, Roger D Amos, and Nicholas C Handy. Higher analytic derivatives. ii. the fourth derivative of self-consistent-field energy. *The Journal of chemical physics*, 95(10):7409–7417, 1991.
- [71] AR Hoy, IM Mills, and G Strey. Anharmonic force constant calculations. *Molecular Physics*, 24(6):1265–1290, 1972.
- [72] Harald H Nielsen. The vibration-rotation energies of molecules. *Reviews of Modern Physics*, 23(2):90, 1951.
- [73] Peter Pulay. Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules: I. theory. *Molecular Physics*, 17(2):197–204, 1969.
- [74] Guillaume Fradet. Master 1, 2016.
- [75] Pierre Barchewitz. *Spectroscopie infrarouge: Vibrations moléculaires*, volume 1. Gauthier-Villars, 1961.
- [76] Philippe Carbonnière. *Calcul ab-initio de spectres vibrationnels anharmoniques.: Développement méthodologique pour le traitement de systèmes de grande taille*. PhD thesis, Pau, 2002.
- [77] Dusan Papousek and Mamed Ragimovich Aliev. Molecular vibrational-rotational spectra. 1982.
- [78] JH Van Vleck. On σ -type doubling and electron spin in the spectra of diatomic molecules. *Physical Review*, 33(4):467, 1929.
- [79] Takeshi Oka. Vibration—rotation interaction in symmetric-top molecules and the splitting between a1 and a2 levels. *The Journal of Chemical Physics*, 47(12):5410–5426, 1967.
- [80] MJ Frisch, GW Trucks, HB Schlegel, GE Scuseria, MA Robb, JR Cheeseman, G Scalmani, V Barone, B Mennucci, GA Petersson, et al. Gaussian 09, revision a. 02; gaussian, inc: Wallingford, ct, 2009. *There is no corresponding record for this reference*, 2015.

- [81] E Fermi. Über den ramaneffekt des kohlendioxyds. *Zeitschrift für Physik*, 71(3-4):250–259, 1931.
- [82] Byron T Darling and David M Dennison. The water vapor molecule. *Physical Review*, 57(2):128, 1940.
- [83] Jan ML Martin and Peter R Taylor. Accurate ab initio quartic force field for trans-hnnh and treatment of resonance polyads. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 53(8):1039–1050, 1997.
- [84] Joon O Jung and R Benny Gerber. Vibrational wave functions and spectroscopy of (H_2O) n , $n = 2, 3, 4, 5$: Vibrational self-consistent field with correlation corrections. *The Journal of chemical physics*, 105(23):10332–10348, 1996.
- [85] Lawrence S Norris, Mark A Ratner, Adrian E Roitberg, and RB Gerber. Mo/lle–plesset perturbation theory applied to vibrational problems. *The Journal of chemical physics*, 105(24):11261–11267, 1996.
- [86] Claude Pouchan, Khalil Zaki, et al. Ab-initio configuration-interaction determination of the overtone vibrations of methyleneimine in the region 2800–3200 cm⁻¹. *Journal of Chemical Physics*, 107(2):342–345, 1997.
- [87] Stuart Carter, Susan J Culik, and Joel M Bowman. Vibrational self-consistent field method for many-mode systems: A new approach and application to the vibrations of co adsorbed on Cu (100). *The Journal of chemical physics*, 107(24):10458–10469, 1997.
- [88] Jacek Koput, Stuart Carter, Nicholas C Handy, et al. Ab initio prediction of the vibrational-rotational energy levels of hydrogen peroxide and its isotopomers. *Journal of Chemical Physics*, 115(18):8345–8350, 2001.
- [89] Patrick Cassam-Chenaï and Jacques Liévin. Alternative perturbation method for the molecular vibration–rotation problem. *International journal of quantum chemistry*, 93(3):245–264, 2003.
- [90] Neil Gohaud, Didier Begue, Clovis Darrigan, and Claude Pouchan. New parallel software (p_anhar) for anharmonic vibrational calculations: Application to (CH₃Li) 2. *Journal of computational chemistry*, 26(7):743–754, 2005.

- [91] F Culot, F Laruelle, and J Liévin. A vibrational casscf study of stretch-bend interactions and their influence on infrared intensities in the water molecule. *Theoretica chimica acta*, 92(4):211–226, 1995.
- [92] O Coulaud, Marc Odulami, Vincent LeBris, Didier Bégué, and Isabelle Baraille. Notes pyravci : Python impélementation of the a-vci algorithm. Technical report, INRIA and IPREM, 2016.
- [93] Robert E Wyatt, Christophe Iung, and Claude Leforestier. Toward ab initio intramolecular dynamics. *Accounts of chemical research*, 28(10):423–429, 1995.
- [94] Didier Bégué, Neil Gohaud, Claude Pouchan, Patrick Cassam-Chenaï, and Jacques Liévin. A comparison of two methods for selecting vibrational configuration interaction spaces on a heptatomic system: Ethylene oxide. *The Journal of chemical physics*, 127(16):164115, 2007.
- [95] Romain Garnier, Marc Odunlami, Vincent Le Bris, Didier Bégué, Isabelle Baraille, and Olivier Coulaud. Adaptive vibrational configuration interaction (a-vci): a posteriori error estimation to efficiently compute anharmonic ir spectra. *The Journal of chemical physics*, 144(20):204123, 2016.
- [96] Joel M Bowman. Self-consistent field energies and wavefunctions for coupled oscillators. *The Journal of Chemical Physics*, 68(2):608–610, 1978.
- [97] Joel M Bowman. The self-consistent-field approach to polyatomic vibrations. *Accounts of Chemical Research*, 19(7):202–208, 1986.
- [98] RB Gerber and Mark A Ratner. Self-consistent-field methods for vibrational excitations in polyatomic systems. *Advances in Chemical Physics*, edited by I. Prigogine and SA Rice (Wiley, 1988), 70:97–132, 1988.
- [99] RB Gerber and Mark A Ratner. A semiclassical self-consistent field (sc scf) approximation for eigenvalues of coupled-vibration systems. *Chemical Physics Letters*, 68(1):195–198, 1979.

- [100] Stuart Carter, Nicholas C Handy, and Joel M Bowman. High torsional vibrational energies of h₂o₂ and ch₃oh studied by multimode with a large amplitude motion coupled to two effective contraction schemes. *Molecular Physics*, 107(8-12):727–737, 2009.
- [101] Manuel Sparta, Mikkel B Hansen, Eduard Matito, Daniele Toffoli, and Ove Christiansen. Using electronic energy derivative information in automated potential energy surface construction for vibrational calculations. *Journal of chemical theory and computation*, 6(10):3162–3175, 2010.
- [102] Kiyoshi Yagi, So Hirata, and Kimihiko Hirao. Vibrational quasi-degenerate perturbation theory: applications to fermi resonance in co 2, h 2 co, and c 6 h 6. *Physical Chemistry Chemical Physics*, 10(13):1781–1788, 2008.
- [103] B Huron, JP Malrieu, and P Rancurel. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. *The Journal of Chemical Physics*, 58(12):5745–5759, 1973.
- [104] Carlo Adamo, Elisa Rebolini, and Andreas Savin. Décrire la structure électronique avec des fonctionnelles de la densité. *l'actualité chimique*, (382-383), 2014.
- [105] Frank Neese. Prediction of molecular properties and molecular spectroscopy with density functional theory: from fundamental theory to exchange-coupling. *Coordination Chemistry Reviews*, 253(5):526–563, 2009.
- [106] Daniel Sánchez-Portal, Pablo Ordejon, Emilio Artacho, and Jose M Soler. Density-functional method for very large systems with lcao basis sets. *International Journal of Quantum Chemistry*, 65(5):453–461, 1997.
- [107] Mark P Waller, Arturo Robertazzi, James A Platts, David E Hibbs, and Peter A Williams. Hybrid density functional theory for π -stacking interactions: Application to benzenes, pyridines, and dna bases. *Journal of computational chemistry*, 27(4):491–504, 2006.
- [108] Stefan Grimme. Semiempirical gga-type density functional constructed with a long-range dispersion correction. *Journal of computational chemistry*, 27(15):1787–1799, 2006.

- [109] Stefan Grimme. Accurate description of van der waals complexes by density functional theory including empirical corrections. *Journal of computational chemistry*, 25(12):1463–1473, 2004.
- [110] Jinwoo Park, Byung Deok Yu, and Suklyun Hong. Ab initio calculations with van der waals corrections: Benzene-benzene intermolecular case and graphite. *J. Korean Phys. Soc.*, 59:196–19, 2011.
- [111] Christopher M Lee, Naseer MA Mohamed, Heath D Watts, James D Kubicki, and Seong H Kim. Sum-frequency-generation vibration spectroscopy and density functional theory calculations with dispersion corrections (dft-d2) for cellulose i α and i β . *The Journal of Physical Chemistry B*, 117(22):6681–6692, 2013.
- [112] Ilya G Kaplan. *Intermolecular interactions: physical picture, computational methods and model potentials*. John Wiley & Sons, 2006.
- [113] Werner Reckien, Florian Janetzko, Michael F Peintinger, and Thomas Bredow. Implementation of empirical dispersion corrections to density functional theory for periodic systems. *Journal of computational chemistry*, 33(25):2023–2031, 2012.
- [114] Alexandre Tkatchenko and Matthias Scheffler. Accurate molecular van der waals interactions from ground-state electron density and free-atom reference data. *Physical review letters*, 102(7):073005, 2009.
- [115] Leeor Kronik and Alexandre Tkatchenko. Understanding molecular crystals with dispersion-inclusive density functional theory: pairwise corrections and beyond. *Accounts of chemical research*, 47(11):3208–3216, 2014.
- [116] Tomáš Bučko, Sébastien Lebègue, János G Ángyán, and Jürgen Hafner. Extending the applicability of the tkatchenko-scheffler dispersion correction via iterative hirshfeld partitioning. *The Journal of chemical physics*, 141(3):034114, 2014.
- [117] Alexandre Tkatchenko, Robert A DiStasio Jr, Roberto Car, and Matthias Scheffler. Accurate and efficient method for many-body van der waals interactions. *Physical review letters*, 108(23):236402, 2012.

- [118] Tomáš Bučko, S Lebègue, Jürgen Hafner, and JG Ángyán. Tkatchenko-scheffler van der waals correction method with and without self-consistent screening applied to solids. *Physical Review B*, 87(6):064110, 2013.
- [119] Stefan Grimme. Do special noncovalent π - π stacking interactions really exist? *Angewandte Chemie International Edition*, 47(18):3430–3434, 2008.
- [120] Georgia B McGaughey, Marc Gagné, and Anthony K Rappé. π -stacking interactions alive and well in proteins. *Journal of Biological Chemistry*, 273(25):15458–15463, 1998.
- [121] John C Slater. Note on hartree's method. *Physical Review*, 35(2):210, 1930.
- [122] Ao D Buckingham. Permanent and induced molecular moments and long-range intermolecular forces. *Advances in Chemical Physics: Intermolecular Forces, Volume 12*, pages 107–142, 1967.
- [123] C David Sherrill. Counterpoise correction and basis set superposition error. *School of Chemistry and Biochemistry, Georgia Institute of Technology*, 2010.
- [124] S F_ Boys and Fiorenza de Bernardi. The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors. *Molecular Physics*, 19(4):553–566, 1970.
- [125] F London. Z phys chem b 1930, 11, 222; eisenschitz, r.; london, f. *Z Phys*, 60:491, 1930.
- [126] SC Wang. The mutual influence between hydrogen atoms. *Phys Z*, 28:663–666, 1927.
- [127] Keiji Morokuma. Why do molecules interact? the origin of electron donor-acceptor complexes, hydrogen bonding and proton affinity. *Accounts of Chemical research*, 10(8):294–300, 1977.
- [128] Walter J Stevens and William H Fink. Frozen fragment reduced variational space analysis of hydrogen bonding interactions. application to the water dimer. *Chemical physics letters*, 139(1):15–22, 1987.
- [129] Wei Chen and Mark S Gordon. Energy decomposition analyses for many-body interaction and applications to water complexes. *The Journal of Physical Chemistry*, 100(34):14316–14328, 1996.

- [130] Daniel M Chipman, JD Bowman, and JO Hirschfelder. Perturbation theories for the calculation of molecular interaction energies. i. general formalism. *The Journal of Chemical Physics*, 59(6):2830–2837, 1973.
- [131] Krzysztof Szalewicz and Bogumil Jeziorski. Symmetry-adapted double-perturbation analysis of intramolecular correlation effects in weak intermolecular interactions: the he-he interaction. *Molecular Physics*, 38(1):191–208, 1979.
- [132] HBG Casimir and D Polder. The influence of retardation on the london-van der waals forces. *Physical Review*, 73(4):360, 1948.
- [133] Andreas Heßelmann and Georg Jansen. First-order intermolecular interaction energies from kohn–sham orbitals. *Chemical physics letters*, 357(5):464–470, 2002.
- [134] Hayes L Williams and Cary F Chabalowski. Using kohn-sham orbitals in symmetry-adapted perturbation theory to investigate intermolecular interactions. *The Journal of Physical Chemistry A*, 105(3):646–659, 2001.
- [135] Alston J Misquitta and Krzysztof Szalewicz. Symmetry-adapted perturbation-theory calculations of intermolecular forces employing density-functional description of monomers. *The Journal of chemical physics*, 122(21):214109, 2005.
- [136] Georg Jansen and Andreas Hesselmann. Comment on “using kohn-sham orbitals in symmetry-adapted perturbation theory to investigate intermolecular interactions”. *The Journal of Physical Chemistry A*, 105(49):11156–11157, 2001.
- [137] Andreas Heßelmann and Georg Jansen. Intermolecular dispersion energies from time-dependent density functional theory. *Chemical physics letters*, 367(5):778–784, 2003.
- [138] EKU Gross, JF Dobson, and M Petersilka. Density functional theory of time-dependent phenomena. In *Density functional theory II*, pages 81–172. Springer, 1996.
- [139] Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang. Challenges for density functional theory. *Chemical Reviews*, 112(1):289–320, 2011.
- [140] Jan Rezac, Lucia Simova, and Pavel Hobza. Ccsd [t] describes noncovalent interactions better than the ccsd (t), ccsd (tq), and ccsdt methods.

- [141] Kevin E Riley, Michal Pitonak, Petr Jurecka, and Pavel Hobza. Stabilization and structure calculations for noncovalent interactions in extended molecular systems based on wave function and density functional theories. *Chemical reviews*, 110(9):5023–5063, 2010.
- [142] H Krause, B Ernstberger, and HJ Neusser. Binding energies of small benzene clusters. *Chemical physics letters*, 184(5-6):411–417, 1991.
- [143] JR Grover, EA Walters, and ET Hui. Dissociation energies of the benzene dimer and dimer cation. *Journal of Physical Chemistry*, 91(12):3233–3237, 1987.
- [144] Christopher A Hunter, Juswinder Singh, and Janet M Thornton. π - π interactions: The geometry and energetics of phenylalanine-phenylalanine interactions in proteins. *Journal of molecular biology*, 218(4):837–846, 1991.
- [145] Matthew CT Fyfe and J Fraser Stoddart. Synthetic supramolecular chemistry. *Accounts of chemical research*, 30(10):393–401, 1997.
- [146] Julius Rebek. Assembly and encapsulation with self-complementary molecules. *Chemical Society Reviews*, 25(4):255–264, 1996.
- [147] SK Burley and GA Petsko. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science*, 229(4708):23–28, 1985.
- [148] Kenneth C Janda, John C Hemminger, John S Winn, Stewart E Novick, Stephen J Harris, and William Klemperer. Benzene dimer: a polar molecule. *The Journal of Chemical Physics*, 63(4):1419–1421, 1975.
- [149] E Arunan and HS Gutowsky. The rotational spectrum, structure and dynamics of a benzene dimer. *The Journal of chemical physics*, 98(5):4294–4296, 1993.
- [150] BF Henson, GV Hartland, VA Venturo, and PM Felker. Raman-vibronic double-resonance spectroscopy of benzene dimer isotopomers. *The Journal of chemical physics*, 97(4):2189–2208, 1992.
- [151] Young Choon Park and Jae Shin Lee. Accurate ab initio binding energies of the benzene dimer. *The Journal of Physical Chemistry A*, 110(15):5091–5095, 2006.

- [152] Seiji Tsuzuki, Kazumasa Honda, Tadafumi Uchimaru, Masuhiro Mikami, and Kazutoshi Tanabe. Origin of attraction and directionality of the π/π interaction: model chemistry calculations of benzene dimer interaction. *Journal of the American Chemical Society*, 124(1):104–112, 2002.
- [153] Mutasem Omar Sinnokrot, Edward F Valeev, and C David Sherrill. Estimates of the ab initio limit for $\pi\text{-}\pi$ interactions: The benzene dimer. *Journal of the American Chemical Society*, 124(36):10887–10893, 2002.
- [154] Pavel Hobza, Heinrich L Selzle, and Edward W Schlag. Potential energy surface for the benzene dimer. results of ab initio ccSD (t) calculations show two nearly isoenergetic structures: T-shaped and parallel-displaced. *The Journal of Physical Chemistry*, 100(48):18790–18794, 1996.
- [155] Oliver C Mullins. The asphaltenes. *Annual Review of Analytical Chemistry*, 4:393–418, 2011.
- [156] Jeramie J Adams. Asphaltene adsorption, a literature review. *Energy & Fuels*, 28(5):2831–2856, 2014.
- [157] Lyes Benamsili, Jean-Pierre Korb, Gerald Hamon, Alain Louis-Joseph, Brice Bouyssiere, Honggang Zhou, and Robert G Bryant. Multi-dimensional nuclear magnetic resonance characterizations of dynamics and saturations of brine/crude oil/mud filtrate mixtures confined in rocks: The role of asphaltene. *Energy & Fuels*, 28(3):1629–1640, 2013.
- [158] Oliver C Mullins. The modified yen model. *Energy & Fuels*, 24(4):2179–2207, 2010.
- [159] Hassan Sabbah, Amy L Morrow, Andrew E Pomerantz, and Richard N Zare. Evidence for island structures as the dominant architecture of asphaltenes. *Energy & Fuels*, 25(4):1597–1604, 2011.
- [160] Joëlle Eyssautier, Pierre Levitz, Didier Espinat, Jacques Jestin, Jérémie Gummel, Isabelle Grillo, and Loïc Barré. Insight into asphaltene nanoaggregate structure inferred by small angle neutron and x-ray scattering. *The Journal of Physical Chemistry B*, 115(21):6827–6837, 2011.

- [161] L Barré, J Jestin, A Morisset, T Palermo, and S Simon. Relation between nanoscale structure of asphaltene aggregates and their macroscopic solution properties. *Oil & Gas Science and Technology-Revue de l'IFP*, 64(5):617–628, 2009.
- [162] Mohammad Sedghi, Lamia Goual, William Welch, and Jan Kubelka. Effect of asphaltene structure on association and aggregation using molecular dynamics. *The Journal of Physical Chemistry B*, 117(18):5765–5776, 2013.
- [163] Juan Liu, Yapu Zhao, and Sili Ren. Molecular dynamics simulation of self-aggregation of asphaltenes at an oil/water interface: formation and destruction of the asphaltene protective film. *Energy & Fuels*, 29(2):1233–1242, 2015.
- [164] Fengfeng Gao, Zhen Xu, Guokui Liu, and Shiling Yuan. Molecular dynamics simulation: the behavior of asphaltene in crude oil and at the oil/water interface. *Energy & Fuels*, 28(12):7368–7376, 2014.
- [165] Robel B Teklebrhan, Lingling Ge, Subir Bhattacharjee, Zhenghe Xu, and Johan Sjöblom. Initial partition and aggregation of uncharged polycyclic aromatic molecules at the oil–water interface: a molecular dynamics simulation study. *The Journal of Physical Chemistry B*, 118(4):1040–1051, 2014.
- [166] Bruno Schuler, Gerhard Meyer, Diego Peña, Oliver C Mullins, and Leo Gross. Unraveling the molecular structures of asphaltenes by atomic force microscopy. *Journal of the American Chemical Society*, 137(31):9870–9876, 2015.
- [167] Oliver C. Mullins, Hassan Sabbah, Joëlle Eyssautier, Andrew E. Pomerantz, Loïc Barré, A. Ballard Andrews, Yosadara Ruiz-Morales, Farshid Mostowfi, Richard McFarlane, Lamia Goual, Richard Lepkowicz, Thomas Cooper, Jhony Orbulescu, Roger M. Leblanc, John Edwards, and Richard N. Zare. Advances in asphaltene science and the yen–mullins model. *Energy & Fuels*, 26(7):3986–4003, 2012.
- [168] Sunghwan Kim, Ryan P Rodgers, Greg T Blakney, Christopher L Hendrickson, and Alan G Marshall. Automated electrospray ionization ft-icr mass spectrometry for petroleum analysis. *Journal of the American Society for Mass Spectrometry*, 20(2):263–268, 2009.

- [169] Lateefah A Stanford, Ryan P Rodgers, Alan G Marshall, Jan Czarnecki, and Xin A Wu. Compositional characterization of bitumen/water emulsion films by negative-and positive-ion electrospray ionization and field desorption/ionization fourier transform ion cyclotron resonance mass spectrometry. *Energy & fuels*, 21(2):963–972, 2007.
- [170] Edward Kendrick. A mass scale based on ch₂= 14.0000 for high resolution mass spectrometry of organic compounds. *Analytical Chemistry*, 35(13):2146–2154, 1963.
- [171] Christine A Hughey, Christopher L Hendrickson, Ryan P Rodgers, Alan G Marshall, and Kuangnan Qian. Kendrick mass defect spectrum: a compact visual analysis for ultrahigh-resolution broadband mass spectra. *Analytical Chemistry*, 73(19):4676–4681, 2001.
- [172] Quan Shi, Dujie Hou, Keng H Chung, Chunming Xu, Suoqi Zhao, and Yahe Zhang. Characterization of heteroatom compounds in a crude oil and its saturates, aromatics, resins, and asphaltenes (sara) and non-basic nitrogen fractions analyzed by negative-ion electrospray ionization fourier transform ion cyclotron resonance mass spectrometry. *Energy & Fuels*, 24(4):2545–2553, 2010.
- [173] Ana Carolina Rennó Sodero, Hugo Santos-Silva, Patricia Guevara Level, Brice Bouyssiere, Jean-Pierre Korb, Herve Carrier, Ahmad Alfarra, Didier Bégué, and Isabelle Baraille. Investigation of the effect of sulfur heteroatom on asphaltene aggregation. *Energy & Fuels*, 2016.
- [174] Francesco Frigerio and Daniele Molinari. A multiscale approach to the simulation of asphaltenes. *Computational and Theoretical Chemistry*, 975(1):76–82, 2011.
- [175] Yohei Mikami, Yunfeng Liang, Toshifumi Matsuoka, and Edo S Boek. Molecular dynamics simulations of asphaltenes at the oil–water interface: from nanoaggregation to thin-film formation. *Energy & Fuels*, 27(4):1838–1845, 2013.
- [176] Fan Yang, Plamen Tchoukov, Heather Dettman, Robel B Teklebrhan, Lan Liu, Tadeusz Dabros, Jan Czarnecki, Jacob Masliyah, and Zhenghe Xu. Asphaltene subfractions responsible for stabilizing water-in-crude oil emulsions. part 2: Molecular representations and molecular dynamics simulations. *Energy & Fuels*, 29(8):4783–4794, 2015.

- [177] Herman JC Berendsen, David van der Spoel, and Rudi van Drunen. Gromacs: a message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91(1):43–56, 1995.
- [178] Berk Hess, Carsten Kutzner, David Van Der Spoel, and Erik Lindahl. Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of chemical theory and computation*, 4(3):435–447, 2008.
- [179] David Van Der Spoel, Erik Lindahl, Berk Hess, Gerrit Groenhof, Alan E Mark, and Herman JC Berendsen. Gromacs: fast, flexible, and free. *Journal of computational chemistry*, 26(16):1701–1718, 2005.
- [180] Wilfred F van Gunsteren, Xavier Daura, and Alan E Mark. Gromos force field. *Encyclopedia of computational chemistry*, 1998.
- [181] Chris Oostenbrink, Alessandra Villa, Alan E Mark, and Wilfred F Van Gunsteren. A biomolecular force field based on the free enthalpy of hydration and solvation: the gromos force-field parameter sets 53a5 and 53a6. *Journal of computational chemistry*, 25(13):1656–1676, 2004.
- [182] Tetyana Kuznicki, Jacob H Masliyah, and Subir Bhattacharjee. Molecular dynamics study of model molecules resembling asphaltene-like structures in aqueous organic solvent systems. *Energy & Fuels*, 22(4):2379–2389, 2008.
- [183] Tetyana Kuznicki, Jacob H Masliyah, and Subir Bhattacharjee. Aggregation and partitioning of model asphaltenes at toluene- water interfaces: Molecular dynamics simulations. *Energy & Fuels*, 23(10):5027–5035, 2009.
- [184] Robel B Teklebrhan, Lingling Ge, Subir Bhattacharjee, Zhenghe Xu, and Johan Sjöblom. Probing structure–nanoaggregation relations of polyaromatic surfactants: a molecular dynamics simulation and dynamic light scattering study. *The Journal of Physical Chemistry B*, 116(20):5907–5918, 2012.
- [185] Alexander W Schüttelkopf and Daan MF Van Aalten. Prodrg: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallographica Section D: Biological Crystallography*, 60(8):1355–1363, 2004.

- [186] D. W.; Krekelberg W. P. Shen, V. K.; Siderius. Nist standard reference simulation website, nist standard reference database number 173, 2015.
- [187] Wei Huang, Andreas P Eichenberger, and Wilfred F van Gunsteren. Molecular dynamics simulation of thionated hen egg white lysozyme. *Protein Science*, 21(8):1153–1161, 2012.
- [188] Chengteh Lee, Weitao Yang, and Robert G Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical review B*, 37(2):785, 1988.
- [189] PJ Stephens, FJ Devlin, CFn Chabalowski, and Michael J Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *The Journal of Physical Chemistry*, 98(45):11623–11627, 1994.
- [190] Ansgar Schäfer, Hans Horn, and Reinhart Ahlrichs. Fully optimized contracted gaussian basis sets for atoms li to kr. *The Journal of Chemical Physics*, 97(4):2571–2577, 1992.
- [191] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: design and assessment of accuracy. *Physical Chemistry Chemical Physics*, 7(18):3297–3305, 2005.
- [192] Frank Neese. The orca program system. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(1):73–78, 2012.
- [193] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 126(1):014101, 2007.
- [194] Herman JC Berendsen, JPM van Postma, Wilfred F van Gunsteren, ARHJ DiNola, and JR Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.
- [195] William G Hoover. Canonical dynamics: equilibrium phase-space distributions. *Physical Review A*, 31(3):1695, 1985.
- [196] Shūichi Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molecular physics*, 52(2):255–268, 1984.

- [197] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.
- [198] Loup Verlet. Computer" experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Physical review*, 159(1):98, 1967.
- [199] Ulrich Essmann, Lalith Perera, Max L Berkowitz, Tom Darden, Hsing Lee, and Lee G Pedersen. A smooth particle mesh ewald method. *The Journal of chemical physics*, 103(19):8577–8593, 1995.
- [200] Berk Hess, Henk Bekker, Herman JC Berendsen, and Johannes GEM Fraaije. Lincs: a linear constraint solver for molecular simulations. *Journal of computational chemistry*, 18(12):1463–1472, 1997.
- [201] Nicholas E Jackson, Brett M Savoie, Kevin L Kohlstedt, Monica Olvera De La Cruz, George C Schatz, Lin X Chen, and Mark A Ratner. Controlling conformations of conjugated polymers and small molecules: The role of nonbonding interactions. *Journal of the American Chemical Society*, 135(28):10475–10483, 2013.
- [202] KH Michaelian, Q Wen, BE Billinghamurst, JM Shaw, and V Lastovka. Far-and mid-infrared photoacoustic spectra of tetracene, pentacene, perylene and pyrene. *Vibrational Spectroscopy*, 58:50–56, 2012.
- [203] Elisabetta Venuti, Raffaele Guido Della Valle, Luca Farina, Aldo Brillante, Matteo Masino, and Alberto Girlando. Phonons and structures of tetracene polymorphs at low temperature and high pressure. *Physical Review B*, 70(10):104106, 2004.
- [204] RB Campbell, J Monteath Robertson, and J Trotter. The crystal structure of hexacene, and a revision of the crystallographic data for tetracene. *Acta crystallographica*, 15(3):289–290, 1962.
- [205] Elisabetta Venuti, Raffaele Guido Della Valle, Aldo Brillante, Matteo Masino, and Alberto Girlando. Probing pentacene polymorphs by lattice dynamics calculations. *Journal of the American Chemical Society*, 124(10):2128–2129, 2002.
- [206] Aldo Brillante, Ivano Bilotti, Raffaele Guido Della Valle, Elisabetta Venuti, Matteo Masino, and Alberto Girlando. Characterization of phase purity in organic semicon-

- ductors by lattice-phonon confocal raman mapping: Application to pentacene. *Advanced Materials*, 17(21):2549–2553, 2005.
- [207] RB Campbell, J Monteath Robertson, and J Trotter. The crystal and molecular structure of pentacene. *Acta Crystallographica*, 14(7):705–711, 1961.
- [208] Daniel Holmes, Sriram Kumaraswamy, Adam J Matzger, and K Peter C Vollhardt. On the nature of nonplanarity in the [n] phenylenes. *Chemistry—A European Journal*, 5(11):3399–3412, 1999.
- [209] Theo Siegrist, Christian Kloc, Jan H Schön, Bertram Batlogg, Robert C Haddon, Steffen Berg, and Gordon A Thomas. Enhanced physical properties in a pentacene polymorph. *Angewandte Chemie International Edition*, 40(9):1732–1736, 2001.
- [210] Christine C Mattheus, Anne B Dros, Jacob Baas, Auke Meetsma, Jan L de Boer, and Thomas TM Palstra. Polymorphism in pentacene. *Acta Crystallographica Section C: Crystal Structure Communications*, 57(8):939–941, 2001.
- [211] Christine C Mattheus, Anne B Dros, Jacob Baas, Gert T Oostergetel, Auke Meetsma, Jan L de Boer, and Thomas TM Palstra. Identification of polymorphs of pentacene. *Synthetic metals*, 138(3):475–481, 2003.
- [212] Christine C Mattheus, Gilles A de Wijs, Robert A de Groot, and Thomas TM Palstra. Modeling the polymorphism of pentacene. *Journal of the American Chemical Society*, 125(20):6323–6330, 2003.
- [213] Charles W Bauschlicher. A comparison of the accuracy of different functionals. *Chemical Physics Letters*, 246(1):40–44, 1995.
- [214] Julian Tirado-Rives and William L Jorgensen. Performance of b3lyp density functional methods for a large set of organic molecules. *Journal of Chemical Theory and Computation*, 4(2):297–306, 2008.
- [215] Charles W Bauschlicher and Stephen R Langhoff. The calculation of accurate harmonic frequencies of large molecules: the polycyclic aromatic hydrocarbons, a case study. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 53(8):1225–1240, 1997.

- [216] Neil Gohaud, Didier Begue, and Claude Pouchan. Vibrational spectra of methylolithium and its aggregates: a new interpretation from ab initio anharmonic calculations. *Chemical physics*, 310(1):85–96, 2005.
- [217] Sándor Kristyán and Péter Pulay. Can (semi) local density functional theory account for the london dispersion forces? *Chemical physics letters*, 229(3):175–180, 1994.
- [218] Pavel Hobza, Tomáš Reschel, et al. Density functional theory and molecular clusters. *Journal of computational chemistry*, 16(11):1315–1325, 1995.
- [219] Didier Begue, Stéphane Elissalde, Eve Pere, Pierre Iratcabal, and Claude Pouchan. New theoretical and experimental infrared results on formaldehyde in solution. *The Journal of Physical Chemistry A*, 110(25):7793–7800, 2006.
- [220] Didier Begue, Greg GuangHua Qiao, and Curt Wentrup. Nitrile imines: matrix isolation, ir spectra, structures, and rearrangement to carbodiimides. *Journal of the American Chemical Society*, 134(11):5339–5350, 2012.
- [221] J_A Pople, R Krishnan, HB Schlegel, and J_S_ Binkley. Derivative studies in hartree-fock and möller-plesset theories. *International Journal of Quantum Chemistry*, 16(S13):225–241, 1979.
- [222] Grzegorz Chalasinski and Małgorzata M Szczesniak. State of the art and challenges of the ab initio theory of intermolecular interactions. *Chemical reviews*, 100(11):4227–4252, 2000.
- [223] Erin R Johnson and Gino A DiLabio. Structure and binding energies in van der waals dimers: Comparison between density functional theory and correlated ab initio methods. *Chemical physics letters*, 419(4):333–339, 2006.
- [224] Andrzej Eilmes. A theoretical study of small tetracene clusters: Stability and charge stabilization energies. *Computational and Theoretical Chemistry*, 982:25–29, 2012.
- [225] Jeng-Da Chai and Martin Head-Gordon. Systematic optimization of long-range corrected hybrid density functionals. *The Journal of chemical physics*, 128(8):084106, 2008.

- [226] Ulrike Salzner and Aykut Aydin. Improved prediction of properties of π -conjugated oligomers with range-separated hybrid density functionals. *Journal of chemical theory and computation*, 7(8):2568–2583, 2011.
- [227] RBJS Krishnan, J Stephen Binkley, Rolf Seeger, and John A Pople. Self-consistent molecular orbital methods. xx. a basis set for correlated wave functions. *The Journal of Chemical Physics*, 72(1):650–654, 1980.
- [228] Michael J Frisch, John A Pople, and J Stephen Binkley. Self-consistent molecular orbital methods 25. supplementary functions for gaussian basis sets. *The Journal of chemical physics*, 80(7):3265–3269, 1984.
- [229] Morihisa Saeki, Hiroshi Akagi, and Masaaki Fujii. Theoretical study on the structure and the frequency of isomers of the naphthalene dimer. *Journal of chemical theory and computation*, 2(4):1176–1183, 2006.
- [230] Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical review B*, 54(16):11169, 1996.
- [231] Stefan Grimme, Stephan Ehrlich, and Lars Goerigk. Effect of the damping function in dispersion corrected density functional theory. *Journal of computational chemistry*, 32(7):1456–1465, 2011.
- [232] Max Dion, Henrik Rydberg, Elsebeth Schröder, David C Langreth, and Bengt I Lundqvist. Van der waals density functional for general geometries. *Physical review letters*, 92(24):246401, 2004.
- [233] Roberto Dovesi, Roberto Orlando, Alessandro Erba, Claudio M Zicovich-Wilson, Bartolomeo Civalleri, Silvia Casassa, Lorenzo Maschio, Matteo Ferrabone, Marco De La Pierre, Philippe D’Arco, et al. Crystal14: A program for the ab initio investigation of crystalline solids. *International Journal of Quantum Chemistry*, 114(19):1287–1317, 2014.
- [234] MJEA Frisch, GW Trucks, Hs B Schlegel, GE Scuseria, MA Robb, JR Cheeseman, G Scalmani, V Barone, B Mennucci, GAea Petersson, et al. Gaussian 09, revision a. 02, gaussian. Inc., Wallingford, CT, 2009.

- [235] Charles W Bauschlicher and Harry Partridge. The sensitivity of b3lyp atomization energies to the basis set and a comparison of basis set requirements for ccSD (t) and b3lyp. *Chemical physics letters*, 240(5):533–540, 1995.
- [236] Joel M Bowman, Tucker Carrington, and Hans-Dieter Meyer. Variational quantum approaches for computing vibrational energies of polyatomic molecules. *Molecular Physics*, 106(16-18):2145–2182, 2008.
- [237] D Bégué, I Baraille, PA Garrain, A Dargelos, and T Tassaing. Calculation of ir frequencies and intensities in electrical and mechanical anharmonicity approximations: Application to small water clusters. *The Journal of chemical physics*, 133(3):034102, 2010.
- [238] I Baraille, C Larrieu, A Dargelos, and M Chaillet. Calculation of non-fundamental ir frequencies and intensities at the anharmonic level. i. the overtone, combination and difference bands of diazomethane, h 2 cn 2. *Chemical Physics*, 273(2):91–101, 2001.
- [239] G Mallocci, G Mulas, G Cappellini, and C Joblin. Time-dependent density functional study of the electronic spectra of oligoacenes in the charge states- 1, 0,+ 1, and+ 2. *Chemical Physics*, 340(1):43–58, 2007.
- [240] A Brillante, RG Della Valle, L Farina, A Girlando, M Masino, and E Venuti. Raman phonon spectra of pentacene polymorphs. *Chemical physics letters*, 357(1):32–36, 2002.
- [241] C Pratt Brock and JACK D Dunitz. Temperature dependence of thermal motion in crystalline naphthalene. *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry*, 38(8):2218–2228, 1982.
- [242] Julio C Facelli and David M Grant. Determination of molecular symmetry in crystalline naphthalene using solid-state nmr. 1993.
- [243] C Pratt Brock and JD Dunitz. Temperature dependence of thermal motion in crystalline anthracene. *Acta Crystallographica Section B: Structural Science*, 46(6):795–806, 1990.
- [244] M Hineno and H Yoshinaga. Far-ir spectra of naphthalene, anthracene, biphenyl, diphenyl ether, diphenylamine, diphenyl sulfoxide, and benzophenone at liquid helium temperature. *Spectrochimica Acta Part A: Molecular Spectroscopy*, 31(5):617–620, 1975.

- [245] Seiji Tsuzuki, Kazumasa Honda, Tadafumi Uchimaru, and Masuhiro Mikami. High-level ab initio computations of structures and interaction energies of naphthalene dimers: Origin of attraction and its directionality. *The Journal of chemical physics*, 120(2):647–659, 2004.
- [246] TR Walsh. An ab initio study of the low energy structures of the naphthalene dimer. *Chemical physics letters*, 363(1):45–51, 2002.
- [247] Miroslav Rubeš, Ota Bludský, and Petr Nachtigall. Investigation of the benzene–naphthalene and naphthalene–naphthalene potential energy surfaces: Dft/ccsd (t) correction scheme. *ChemPhysChem*, 9(12):1702–1708, 2008.
- [248] SF Boys and F Bernardi. The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors. *Molecular Physics*, 100(1):65–73, 2002.
- [249] Edward G Hohenstein and C David Sherrill. Density fitting and cholesky decomposition approximations in symmetry-adapted perturbation theory: Implementation and application to probe the nature of π - π interactions in linear acenes. *The Journal of Chemical Physics*, 132(18):184111, 2010.
- [250] Brant Billinghamurst, Tim May, Les Dallin, Ward Wurtz, Mark de Jong, and Kirk H Michaelian. Photoacoustic spectroscopy using coherent synchrotron radiation: application to α -lactose monohydrate. *Optics letters*, 35(18):3090–3092, 2010.
- [251] S Appalakondaiah, G Vaitheeswaran, and S Lebègue. Dispersion corrected structural properties and quasiparticle band gaps of several organic energetic solids. *The Journal of Physical Chemistry A*, 119(24):6574–6581, 2015.
- [252] Carlos Gonzalez and Edward C Lim. A quantum chemistry study of the van der waals dimers of benzene, naphthalene, and anthracene: Crossed (d 2 d) and parallel-displaced (c 2 h) dimers of very similar energies in the linear polyacenes. *The Journal of Physical Chemistry A*, 104(13):2953–2957, 2000.