



---

## THESE

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 :

---

---

**Vibrational spectra calculations, fingerprints allow to elucidate composition and reactivity of complex chemical systems: application to asphaltene systems**

---

Directeur de Thèse: **Pr Didier BÉGUÉ**

Date de soutenance prévue le 28 Novembre 2016

devant la commission d'examen composée de :

## JURY

Dr Joëlle MASCETTI	Université de Bordeaux 1, France	Rapporteur
Dr Benjamin HANOUNE	Université de Lille 1, France	Rapporteur



# Table of Contents

<b>Introduction</b>	<b>11</b>
<b>1 Généralités – Contexte de l'étude</b>	<b>16</b>
1.1 Généralités . . . . .	17
1.1.1 Composition du pétrole brut . . . . .	17
1.1.1.1 Définitions . . . . .	17
1.1.1.2 Classification des différentes fractions du pétrole brut . . . . .	18
1.1.2 Une fraction lourde du pétrole brut : les asphaltènes . . . . .	19
1.1.2.1 Définition . . . . .	19
1.1.2.2 Composition chimique des asphaltènes : une teneur importante en hétéroéléments . . . . .	20
1.1.2.3 Débat autour du poids moléculaire des asphaltènes . . . . .	23
1.2 État des lieux des connaissances sur les systèmes asphalténiques . . . . .	23
1.2.1 Revue des données expérimentales . . . . .	24
1.2.1.1 Analyses par spectroscopie RMN . . . . .	24
1.2.1.2 Analyses par diffraction des rayons X . . . . .	25
1.2.1.3 Analyses par spectroscopie de fluorescence résolue en temps . .	25
1.2.1.4 Analyses par spectroscopie infrarouge . . . . .	25
1.2.1.5 Conclusion : les asphaltènes, une coexistence de structures moléculaires ? . . . . .	26
1.2.2 La modélisation, outil indispensable dans l'élucidation des structures asphalténiques . . . . .	26
1.2.2.1 Des nombreux intérêts de la modélisation des asphaltènes . . . .	26
1.2.2.2 Construction de la structure moléculaire des asphaltènes . . . .	27
<b>2 Partie vibrationnelle</b>	<b>28</b>
2.1 Généralités . . . . .	30
2.2 Séparation des mouvements . . . . .	31
2.3 Energie cinétique de vibration . . . . .	33
2.4 Energie potentielle harmonique . . . . .	34
2.5 Équations de Lagrange . . . . .	34
2.5.1 Espace des coordonnées cartésiennes pondérées par les masses . . . .	34
2.5.2 Espace des coordonnées internes : résolution par la méthode de WILSON	35
2.5.3 Espace des coordonnées internes de symétrie . . . . .	36
2.6 Traitement quantique de la vibration . . . . .	36
2.7 La fonction potentielle anharmonique . . . . .	37
2.8 Représentation matricielle du Hamiltonien . . . . .	40

2.9	Méthodes perturbationnelles . . . . .	41
2.10	Méthodes variationnelles . . . . .	42
2.11	Les méthodes de variation-perturbation . . . . .	45
<b>3</b>	<b>Interactions à longue portée</b>	<b>47</b>
3.1	Théorie des perturbations . . . . .	50
3.1.1	L'approche en « supermolécules » . . . . .	51
3.1.2	Détermination théorique des coefficients du développement multipolaire de l'énergie d'interaction à grande distance : potentiel électrostatique .	52
3.1.2.1	Hamiltonien à l'ordre zéro pour le système $[a - b]$ . . . . .	53
3.1.2.2	Fonction d'onde moléculaire à l'ordre zéro . . . . .	54
3.1.2.3	Perturbation . . . . .	54
3.1.2.4	Le Hamiltonien d'interaction . . . . .	55
3.1.2.5	Développement à un centre du potentiel électrostatique dû à une distribution de charges discrètes. . . . .	57
3.1.2.6	Développement à deux centres de l'interaction coulombienne entre deux distributions de charges ne se recouvrant pas . . . . .	58
3.1.2.7	Détermination théorique des coefficients du développement multipolaire de l'énergie d'interaction à grande distance : énergie d'interaction . . . . .	59
3.1.2.8	Calcul de l'énergie électrostatique. . . . .	60
3.1.2.9	Calcul de l'énergie de polarisation . . . . .	61
3.1.3	Symmetry-Adapted Perturbation Theory (SAPT) . . . . .	64
3.2	De la nécessité de la DFT . . . . .	70
3.2.1	Les fondements de la DFT . . . . .	70
3.2.1.1	Modèle de THOMAS-FERMI . . . . .	70
3.2.1.2	Modèle de THOMAS-FERMI-DIRAC . . . . .	70
3.2.2	Modèle de SLATER . . . . .	71
3.2.3	Les théorèmes de HOHENBERG et KOHN . . . . .	71
3.2.3.1	Premier théorème : preuve d'existence . . . . .	71
3.2.3.2	Second théorème : théorème variationnel . . . . .	72
3.2.4	Approche Kohn-Sham . . . . .	72
3.2.5	Les différentes classes de fonctionnelles . . . . .	73
3.2.5.1	Local Density approximation (LDA) . . . . .	74
3.2.5.2	Generalized Gradient Approximation (GGA) . . . . .	74
3.2.5.3	Fonctionnelles hybrides . . . . .	75
3.2.5.4	Nécessité d'une fonctionnelle « longue portée » . . . . .	75
3.2.6	LC-DFT-D hybride : $\omega$ BXD . . . . .	75
3.2.6.1	B88 . . . . .	76
3.2.6.2	B97 . . . . .	76
3.2.6.3	$\omega$ B97 . . . . .	77
3.2.6.4	$\omega$ B97X . . . . .	78
3.2.6.5	$\omega$ B97X-D . . . . .	79
3.2.7	DFT+D . . . . .	80
3.2.7.1	DFT-D2 . . . . .	80
3.2.7.2	DFT-D3 . . . . .	81
3.2.7.3	La méthode de Tkatchenko-Scheffer (DFT-TS) . . . . .	82

---

3.2.7.4	Self-consistent screening de la méthode de Tkatchenko-Scheffer (DFT-TS-SCS) . . . . .	83
<b>4</b>	<b>Discerning inter- and intramolecular vibrations</b>	<b>85</b>
4.1	Introduction . . . . .	86
4.2	Experimental Spectra . . . . .	87
4.2.1	Raman spectra . . . . .	87
4.2.2	Far-infrared spectra . . . . .	88
4.2.3	Mid-infrared spectra . . . . .	88
4.3	Calculation method for thiophene-type molecules and their dimers . . . . .	88
4.3.1	Structural details . . . . .	88
4.4	Resolution of the vibrational Schrödinger equation (RVSE) . . . . .	89
4.5	Results and Discussion . . . . .	91
4.5.1	Monomers . . . . .	91
<b>5</b>	<b>Heteroelements contribution: a vibrational study</b>	<b>108</b>
5.1	Structural details . . . . .	109
5.2	Results and Discussion . . . . .	110
5.2.1	Global analysis . . . . .	110
5.2.2	Analyse . . . . .	115
5.3	Carbazole . . . . .	115
5.4	Dibenzothiophene . . . . .	122
<b>6</b>	<b>Phonon calculation</b>	<b>127</b>
6.1	Computational details . . . . .	128
6.2	A first case: Tetracene . . . . .	130
6.3	Impact of the chain-length parameter on the vibrational signature of the crystalline network. . . . .	134
6.4	Impact of both the chain-length and "widths" parameters on the vibrational signature of the crystalline network. . . . .	143
6.5	Impact of heteroatoms on the vibrational signature of the crystalline network. .	152
<b>7</b>	<b>Molecular dynamics study of the nano-aggregation in asphaltene</b>	<b>160</b>
7.1	Introduction . . . . .	162
7.2	Simulation Method . . . . .	165
7.3	Results and Discussion . . . . .	167
7.3.1	C <sub>5</sub> P <sub>e</sub> -derivatives systems . . . . .	167
7.3.2	Schuler's-derivatives systems . . . . .	180
7.3.2.1	Effect of the heteroatom . . . . .	180
7.3.2.2	Effect of chain-end . . . . .	182
7.3.2.3	Aggregation within heterogeneous boxes . . . . .	184
7.4	Conclusions . . . . .	186
<b>Conclusion Générale</b>		<b>187</b>
<b>Appendices</b>		<b>191</b>

<b>A Calcul de la polarisabilité</b>	<b>192</b>
A.1 Calcul de la polarisabilité en fonction de la fréquence imaginaire . . . . .	193
<b>B Calcul de Phonon</b>	<b>195</b>
B.1 Dynamique du réseau cristallin: Hypothèses fondamentales . . . . .	196
B.2 Quantification des ondes élastiques La notion de phonos . . . . .	198
B.3 Théorie classique du cristal harmonique . . . . .	199
B.3.1 Approximation adiabatique . . . . .	199
B.3.2 Cas du cristal harmonique . . . . .	201
B.3.3 Calcul du spectre de phonons au point $\Gamma$ . . . . .	202
<b>C Database</b>	<b>204</b>
C.1 Monomers . . . . .	205
C.2 Dimers . . . . .	227

# List of Figures

1.1	Structure moyenne du modèle continental des asphaltènes . . . . .	23
1.2	Molecule d'asphaltènes type archipel . . . . .	24
3.1	Structures du dimère de Benzène . . . . .	67
4.1	Molecular structures studied: from left to right: benzothiophene <b>1a</b> , dibenzothiophene <b>1b</b> , 4,6-dimethyldibenzothiophene <b>1c</b> and 4-methyldibenzothiophene <b>1d</b> . . . . .	87
4.2	Selection of normal modes which are stored in the expression of the PES : methodologies applied to realize the truncation of the vibrational basis a) truncation n°1 T'=X'+Y' (at the top); b) truncation n°2 T''=X''+Y'' (at the bottom). . . . .	91
4.3	Ensemble of results obtained using truncation n°1 for the first twelve modes of a benzothiophene molecule . . . . .	93
4.4	Ensemble of results using the truncation n°2 for the first twelve vibration modes of a benzothiophene molecule . . . . .	94
4.5	Far-infrared PA spectra of dibenzothiophene . . . . .	96
4.6	Configurations of benzothiophene dimers optimized at the $\omega$ B97X-D/6-311++G** level of theory. . . . .	98
4.7	Structural parameters of benzothiophene dimers . . . . .	98
4.8	Plots of the electronic reduced gradient density's against against the density $\rho$ times the sign of $\lambda_2$ . . . . .	103
4.9	Far-infrared spectra of dibenzothiophene <b>1b</b> , 4,6-dimethyldibenzothiophene <b>1c</b> and 4-methyldibenzothiophene <b>1d</b> . Data were acquired using coherent synchrotron radiation and a Bruker IFS 125HR FT-IR spectrometer. . . . .	107
5.1	Molecular structures studied: <b>a</b> Benzo, <b>b</b> Dibenzo, <b>c</b> Benzonaphtho, <b>d</b> Tribenzo, <b>e</b> methyldibenzo, <b>f</b> dimethyldibenzo. X= <b>1</b> C, <b>2</b> N, <b>3</b> O and <b>4</b> S. . . . .	111
5.2	Vibration modes heteroatoms contributions > 5% for studied molecules. . . . .	112
5.3	Vibration modes heteroatoms contributions > 5% for dimers configurations . . . . .	113
5.4	Calculated vibrational spectra for all molecules, motifs and heteroatoms comparisons. . . . .	114
5.5	Atomic percentage contributions to vibrational modes in Carbazole molecule. . . . .	116
5.6	Calculated vibrational spectra of Carbazole molecule, showing heteroatom contribution and selected modes. . . . .	117
5.7	Calculated vibrational spectra of Carbazole monomer and dimer, between 0- 800 $\text{cm}^{-1}$ , showing heteroatom contribution and selected modes. . . . .	118
5.8	Comparison between calculated vibrational spectra of monomer and dimer with IR experimental at 0- 250 $\text{cm}^{-1}$ of Carbazole . . . . .	118

5.9	Comparison of density of modes at low wavenumber region of Carbazole. . . . .	119
5.10	Comparison of calculated vibrational spectra and experimental data taking into account theoretical intensities for Carbazole. . . . .	120
5.11	Comparison between monomer and dimer calculated vibrational spectra for <b>a</b> Indole, <b>b</b> 4,5-iminophenanthrene, <b>c</b> 1-methylcarbazole and <b>d</b> 1,8-dimethylcarbazole at 0- 800 cm <sup>-1</sup> interval. . . . .	121
5.12	Variation of D and I modes between monomer- dimer of studied molecules . . .	121
5.13	Monomer and dimer comparison of calculated vibrational spetra of Dibenzothiophene . . . . .	122
5.14	Calculated vibrational spectra of Carbazole monomer and dimer, between 0- 800 cm <sup>-1</sup> , showing heteroatom contribution and selected modes. . . . .	123
5.15	Comparison between monomer, dimer calculated and experimental spectra, Densite of modes at the top and normal vibrational modes and IR intensities at the bottom. . . . .	123
5.16	Comparison between momoner and dimer calculated vibrational spectra for <b>a</b> Benzothiophene, <b>b</b> Tribenzothiophene, <b>c</b> 4-methyldibenzothiophene and <b>d</b> 4,6-dimethyldibenzothiophene at 0- 800 cm <sup>-1</sup> interval. . . . .	124
5.17	Comparison between momoner and dimer calculated vibrational spectra for <b>a</b> Benzofuran, <b>b</b> Dibenzofuran, <b>c</b> Benzonaphthofuran, <b>d</b> Tribenzofuran, <b>e</b> 4-methyldibenzofuran and <b>f</b> 4,6-dimethyldibenzofuran at 0- 800 cm <sup>-1</sup> interval. . .	125
5.18	Comparison between momoner and dimer calculated vibrational spectra for <b>a</b> Indene, <b>b</b> Fluorene, <b>c</b> Cyclopentaphenanthrene, <b>d</b> 1-methylfluorene and <b>e</b> 1,8-dimethylfluorene at 0- 800 cm <sup>-1</sup> interval. . . . .	126
6.1	Final assignment of the experimental infrared transitions for tetracene and pentacene . . . . .	129
6.2	The bc plane of Tetracene Crystal . . . . .	131
6.3	The ac plane of Tetracene Crystal . . . . .	131
6.4	The ab plane of Tetracene Crystal . . . . .	132
6.5	The ac and ab planes of Benzene crystal . . . . .	134
6.6	The ac and ab planes of Naphthalene crystal . . . . .	135
6.7	The ac and ab planes of Anthracene crystal . . . . .	136
6.8	The ac, ab planes of Pentacene polymorph C. . . . .	136
6.9	The bc plane of Pentacene polymorph H at the bottom . . . . .	137
6.10	The ac and ab planes of Phenanthrene crystal . . . . .	143
6.11	The ac and ab planes of Pyrene crystal . . . . .	144
6.12	The ac and ab planes of Chrysene crystal . . . . .	144
6.13	The ac and bc planes of Perylene crystal . . . . .	145
6.14	The ab and ac planes of Fluorene crystal . . . . .	146
6.15	The bc and ab planes of Carbazole crystal . . . . .	153
6.16	The ac and bc planes of Dibenzofuran crystal . . . . .	153
6.17	The ac plane and other view of molecular arrangement of Dibenzothiophene crystal	154
6.18	Final assignment of the experimental infrared transitions for carbazole. Experimental values are indicated for comparison, and represented with blue lines . . .	158
6.19	Molecular structures of studied molecules . . . . .	159
7.1	C5Pe molecular models of asphaltenes. . . . .	163

7.2	Two different asphaltene structures found in the crude petroleum: (a) PA3 and (b) CA22. $R_1$ and $R_2$ are the <i>n</i> -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. . . . .	164
7.3	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. . . . .	167
7.4	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. . . . .	168
7.5	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. . . . .	168
7.6	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. . . . .	169
7.7	Distances between monomers for “organized” starting point. . . . .	170
7.8	Distances between monomers for “random” starting point. . . . .	170
7.9	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.	172
7.10	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.	173
7.11	Distances between monomers for “organized” starting point. . . . .	174
7.12	Angle between monomers for “organized” starting point. . . . .	176
7.13	Distances between monomers for “random” starting point. . . . .	177
7.14	Angle between monomers for “random” starting point. . . . .	177
7.15	Normalized Radial Distribution Functions for the trimer systems. . . . .	178
7.16	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. . . . .	178
7.17	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. . . . .	179
7.18	Aggregation mechanism with $-CH_3$ 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 $\pi$ -stacking structure presented in Figure 7.3. . . . .	180
7.19	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. . . . .	181
7.20	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 $\pi$ -stacking structure presented in Figure 7.19. . . . .	183

7.21	60 ns-snapshots of 5 molecules of each asphaltene in toluene at 298K and 1 bar conditions: (a) A31 and (b) AAA. . . . .	183
7.22	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. . . . .	185
B.1	Réseau de Bravais et vecteur de déplacement pour une base monoatomique . . . . .	197
B.2	Chaque maille unitaire est repérée par un vecteur $\mathbf{R}_a$ . . . . .	201
C.1	Configurations of Indole dimers optimized at the $\omega\text{B97X-D}/6-311++\text{G}^{**}$ level of theory . . . . .	227
C.2	Configurations of Carbazole dimers optimized at the $\omega\text{B97x-D}/6-311+\text{G}^{**}$ level of theory. . . . .	229
C.3	Configurations of 4,5-iminophenanthrene dimers optimized at the $\omega\text{B97X-D}/6-311+\text{G}^{**}$ level of theory. . . . .	232
C.4	Configurations of 1-methylcarbazole dimers optimized at the $\omega\text{B97X-D}/6-311+\text{G}^{**}$ level of theory. . . . .	234
C.5	Configurations of 1,8-dimethylcarbazole dimers optimized at the $\omega\text{B97X-D}/6-311+\text{G}^{**}$ level of theory. . . . .	237
C.6	. . . . .	240
C.7	Configurations of dibenzofuran dimers optimized at the $\omega\text{B97x-D}/6-311+\text{G}^{**}$ level of theory. . . . .	242
C.8	Configurations of Tribenzofuran dimers optimized at the $\omega\text{B97x-D}/6-311+\text{G}^{**}$ level of theory. . . . .	244
C.9	Configurations of 4-methyldibenzofuran dimers optimized at the $\omega\text{B97x-D}/6-311+\text{G}^{**}$ level of theory. . . . .	247
C.10	Configurations of 4,6-dimethyldibenzofuran dimers optimized at the $\omega\text{B97x-D}/6-311+\text{G}^{**}$ level of theory. . . . .	248
C.11	Configurations of Indene dimers optimized at the $\omega\text{B97X-D}/6-311+\text{G}^{**}$ level of theory. . . . .	250
C.12	Configurations of Fluorene dimers optimized at the $\omega\text{B97x-D}/6-311+\text{G}^{**}$ level of theory. . . . .	252
C.13	Configurations of Cyclopentaphenanthrene dimers optimized at the $\omega\text{B97x-D}/6-311+\text{G}^{**}$ level of theory. . . . .	254
C.14	Configurations of 1-methylfluorene dimers optimized at the $\omega\text{B97x-D}/6-311+\text{G}^{**}$ level of theory. . . . .	256
C.15	Configurations of 1,8-dimethylfluorene dimers optimized at the $\omega\text{B97x-D}/6-311+\text{G}^{**}$ level of theory. . . . .	257

# List of Tables

1.1	Composition chimique du pétrole en termes de distribution élémentaire, selon les données recueillies par Speight <i>et al.</i>	18
1.2	Famille de composés soufrés caractérisés au sein du pétrole brut	20
1.3	Exemples de dérivés thiophéniques courants identifiés au sein des fractions lourdes : phénanthro-thiophènes et benzonaphto-thiophènes	21
1.4	Famille de composés azotés caractérisés au sein du pétrole brut	22
3.1	Energies d'interaction (en kcal/mol) du dimère de benzène calculées au niveau CCSD(T) et DFT-SAPT (PBE0)	68
4.1	Low wavenumber Raman and infrared spectra of benzothiophene	96
4.2	Low wavenumber Raman and infrared spectra of dibenzothiophene	97
4.3	Interaction energies of benzothiophene dimers	99
4.4	Interaction energies components in kcal/mol of benzothiophene dimers using DFT-SAPT(PBE0)/aug-cc pVTZ	102
4.5	Parameters used in the variation-perturbation process (values of X' and Y as defined in Figure 4.2 parameters and dimension of the variational problem <b>B''</b> ) for each intermolecular mode of the four most stable dimers <b>2a-2d</b>	106
4.6	calculated intermolecular vibrations modes of the four most stable dimers <b>2a-2d</b>	106
4.7	Experimental and calculated intermolecular vibration modes ( $\text{cm}^{-1}$ ) of dibenzothiophene dimer ( <b>B</b> )	107
5.1	Symbols and assigment of normal modes studied in vibrational spectra calculations	112
6.1	Lattice parameters of Tetracene molecule calculated in VASP code	130
6.2	Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE tetracene system).	133
6.3	Lattice parameters of Benzene molecule calculated in VASP code	135
6.4	Lattice parameters of Naphthalene molecule calculated in VASP code	135
6.5	Lattice parameters of Anthracene molecule calculated in VASP code	136
6.6	Lattice parameters of Pentacene molecule (Polymorphe H) calculated in VASP code	137
6.7	Lattice parameters of pentacene molecule (polymorphe C) calculated in VASP code	138
6.8	Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Benzene system).	139
6.9	Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Naphthalene system).	140

6.10 Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Anthracene system) . . . . .	141
6.11 Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Pentacene system) . . . . .	142
6.12 Lattice parameters of Phenanthrene molecule calculated in VASP code . . . . .	143
6.13 Lattice parameters of Pyrene molecule calculated in VASP code . . . . .	144
6.14 Lattice parameters of Chrysene molecule calculated in VASP code . . . . .	145
6.15 Lattice parameters of Perylene molecule calculated in VASP code . . . . .	145
6.16 Lattice parameters of Fluorene molecule calculated in VASP code . . . . .	146
6.17 Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Phenanthrene system) . . . . .	147
6.18 Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Pyrene system) . . . . .	148
6.19 Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Chrysene system) . . . . .	149
6.20 Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Perylene system) . . . . .	150
6.21 Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Fluorene system) . . . . .	151
6.22 Lattice parameters of Carbazole molecule calculated in VASP code . . . . .	153
6.23 Lattice parameters of Dibenzofuran molecule calculated in VASP code . . . . .	154
6.24 Lattice parameters of Dibenzothiophene molecule calculated in VASP code . . . . .	154
6.25 Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Carbazole system) . . . . .	155
6.26 Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Dibenzofuran system) . . . . .	156
6.27 Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Dibenzothiophene system) . . . . .	157
7.1 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. . . . .	165
7.2 Interaction energies ( $< E_i >$ ) in $\text{kcal} \cdot \text{mol}^{-1}$ calculated for PA3- and CA22-type dimer molecules. . . . .	182
C.1 Low wavenumber Raman ad PA infrared spectra of Indole ( $\text{Cs}$ ). . . . .	205
C.2 Raman and PA infrared spectra of Indole ( $\text{Cs}$ ), $700\text{--}2000\text{ cm}^{-1}$ . . . . .	205
C.3 Low wavenumber Raman ad PA infrared spectra of Carbazole ( $\text{C}_2\text{v}$ ). . . . .	206
C.4 Raman and PA infrared spectra of Carbazole ( $\text{C}_2\text{v}$ ), $700\text{--}2000\text{ cm}^{-1}$ . . . . .	207
C.5 Low wavenumber Raman ad PA infrared spectra of 4,5-iminophenanthrene ( $\text{C}_2\text{v}$ ). . . . .	208
C.6 Raman and PA infrared spectra of 4,5-iminophenanthrene ( $\text{C}_2\text{v}$ ), $700\text{--}2000\text{ cm}^{-1}$ . . . . .	208
C.7 Low wavenumber Raman and PA infrared spectra of 1-methylcarbazole ( $\text{Cs}$ ). . . . .	210
C.8 Raman and PA infrared spectra of 1-methylcarbazole ( $\text{Cs}$ ), $700\text{--}2000\text{ cm}^{-1}$ . . . . .	210
C.9 Low wavenumber Raman ad PA infrared spectra of 1,8-dimethylcarbazole ( $\text{C}_2\text{v}$ ). . . . .	212
C.10 Raman and PA infrared spectra of 1,8-dimethylcarbazole ( $\text{C}_2\text{v}$ ), $700\text{--}2000\text{ cm}^{-1}$	212

C.11 Low wavenumber Raman ad PA infrared spectra of Benzofuran (Cs) . . . . .	213
C.12 Raman and PA infrared spectra of Benzofuran (Cs), 700–2000 cm <sup>−1</sup> . . . . .	213
C.13 Low wavenumber Raman ad PA infrared spectra of Dibenzofuran (C <sub>2</sub> v) . . . . .	214
C.14 Raman and PA infrared spectra of Dibenzofuran (C <sub>2</sub> v), 700–2000 cm <sup>−1</sup> . . . . .	215
C.15 Low wavenumber Raman ad PA infrared spectra of Tribenzofuran (C <sub>2</sub> v) . . . . .	216
C.16 Raman ad PA infrared spectra of Tribenzofuran (C <sub>2</sub> v), 700 - 2000 cm <sup>−1</sup> . . . . .	216
C.17 Low wavenumber Raman ad PA infrared spectra of 4-methyldibenzofuran (Cs) . . . . .	217
C.18 Raman ad PA infrared spectra of 4-methyldibenzofuran (Cs), 700 - 2000 cm <sup>−1</sup> . . . . .	217
C.19 Low wavenumber Raman ad PA infrared spectra of 4,6-dimethyldibenzofuran (C <sub>2</sub> v) . . . . .	218
C.20 Raman ad PA infrared spectra of 4,6-dimethyldibenzofuran (C <sub>2</sub> v), 700 - 2000 cm <sup>−1</sup> . . . . .	218
C.21 Low wavenumber Raman ad PA infrared spectra of Indene (Cs) . . . . .	219
C.22 Raman and PA infrared spectra of Indene (Cs), 700–2000 cm <sup>−1</sup> . . . . .	219
C.23 Low wavenumber Raman ad PA infrared spectra of Fluorene (C <sub>2</sub> v) . . . . .	220
C.24 Raman and PA infrared spectra of Fluorene (C <sub>2</sub> v), 700–2000 cm <sup>−1</sup> . . . . .	221
C.25 Low wavenumber Raman ad PA infrared spectra of Cyclopentaphenanthrene (C <sub>2</sub> v) . . . . .	223
C.26 Raman ad PA infrared spectra of Cyclopentaphenanthrene (C <sub>2</sub> v), 700- 2000 cm <sup>−1</sup> . . . . .	223
C.27 Low wavenumber Raman ad PA infrared spectra of 1-methylfluorene (Cs) . . . . .	224
C.28 Raman ad PA infrared spectra of 1-methylfluorene (Cs), 700 - 2000 cm <sup>−1</sup> . . . . .	224
C.29 Low wavenumber Raman ad PA infrared spectra of 1,8-dimethylfluorene (C <sub>2</sub> v) . . . . .	225
C.30 Raman and PA infrared spectra of 1,8-dimethylfluorene (C <sub>2</sub> v), 700–2000 cm <sup>−1</sup> . . . . .	226
C.31 Interaction energies of Indole dimers (kcal/mol) in 8 different configurations using $\omega$ B97X-D/6-311++G**, and compared with SAPT-DFT . . . . .	227
C.32 Intermolecular distance between pyrrole rings center of Indole dimers (Å), selon figure 4.7 . . . . .	228
C.33 Angle between ring planes of Indole dimers (°) at the $\omega$ B97x-D/6-311++G** level of theory, according to figure 4.7 . . . . .	228
C.34 Calculated low wavenumber Raman ad PA infrared spectra of Indole Dimer . . . . .	228
C.35 Calculated Raman and PA infrared spectra of Indole Dimer, 700–2000 cm <sup>−1</sup> . . . . .	229
C.36 Interaction energies of Carbazole dimers (kcal/mol) in 3 different configurations using $\omega$ B97x-D/6-311++G** . . . . .	230
C.37 Intermolecular distance between pyrrole ring center (Å) and angles (°) according figure of Carbazole dimers using $\omega$ B97x-D/6-311++G** . . . . .	230
C.38 Calculated low wavenumber Raman ad PA infrared spectra of Carbazole Dimer . . . . .	230
C.39 Calculated Raman and PA infrared spectra of Carbazole Dimer, 700–2000 cm <sup>−1</sup> . . . . .	231
C.40 Interaction energies of 4,5-iminophenanthrene dimers (kcal/mol) using $\omega$ B97X-D/6-311++G** . . . . .	232
C.41 Intermolecular distance between pyrrole ring center (Å) and angles (°) according figure of 4,5-iminophenanthrene dimers using $\omega$ B97x-D/6-311++G** . . . . .	232
C.42 Calculated low wavenumber Raman ad PA infrared spectra of 4,5-iminophenanthrene Dimer . . . . .	232
C.43 Calculated Raman and PA infrared spectra of 4,5-iminophenanthrene Dimer, 700–2000 cm <sup>−1</sup> . . . . .	233
C.44 Interaction energies of 1-methylcarbazole dimers (kcal/mol) using $\omega$ B97X-D/6-311++G** . . . . .	234
C.45 Intermolecular distance between pyrrole ring center (Å) and plane angles (°) of 1-methylcarbazole dimers with the $\omega$ B97x-D/6-311++G** . . . . .	235

C.46 Calculated low wavenumber Raman ad PA infrared spectra of 1-methylcarbazole Dimer. . . . .	235
C.47 Raman ad PA infrared spectra of 1-methylcarbazole Dimer, 700- 200 cm <sup>-1</sup> . . . . .	236
C.48 Interaction energies of 1,8-dimethylcarbazole dimers (kcal/mol) using $\omega$ B97X-D/6-311++G**. . . . .	237
C.49 Intermolecular distance between pyrrole ring center (Å) and angles (°) according figure of 1,8-dimethylcarbazole dimers using $\omega$ B97x-D/6-311++G**. . . . .	237
C.50 Calculated low wavenumber Raman ad PA infrared spectra of 1,8-dimethylcarbazole Dimer. . . . .	237
C.51 Calculated Raman and PA infrared spectra of 1,8-dimethylcarbazole Dimer, 700–2000 cm <sup>-1</sup> . . . . .	238
C.52 Raman ad PA infrared spectra of 1,8-methylcarbazole Dimer, 700- 200 cm <sup>-1</sup> . . . . .	239
C.53 Interaction energies of benzofuran dimers (kcal/mol) in 8 different configurations using $\omega$ B97X-D/6-311++G**. and compared with other methods. . . . .	240
C.54 Intermolecular distance between furan rings center of benzofuran dimers (Å) . . . . .	240
C.55 Angle between ring planes of benzofuran dimers (°) at the $\omega$ B97x-D/6-311++G** level of theory. according to figure . . . . .	240
C.56 Calculated low wavenumber Raman ad PA infrared spectra of Benzofuran Dimer. . . . .	241
C.57 Raman ad PA infrared spectra of Benzofuran Dimer, 700- 2000 cm <sup>-1</sup> . . . . .	241
C.58 Interaction energies of dibenzofuran dimers (kcal/mol) in 3 different configurations using $\omega$ B97x-D/6-311++G**. . . . .	242
C.59 Intermolecular distance between furan ring center (Å) according figure of dibenzofuran dimers using $\omega$ B97x-D/6-311++G**. . . . .	242
C.60 Calculated low wavenumber Raman ad PA infrared spectra of Dibenzofuran Dimer. . . . .	243
C.61 Raman ad PA infrared spectra of Dibenzofuran Dimer, 700- 2000 cm <sup>-1</sup> . . . . .	243
C.62 Interaction energies of Tribenzofuran dimers (kcal/mol) in 3 different configurations using $\omega$ B97x-D/6-311++G**. . . . .	245
C.63 Intermolecular distance between furan ring center (Å) according figure of Tribenzofuran dimers using $\omega$ B97x-D/6-311++G**. . . . .	245
C.64 Calculated low wavenumber Raman ad PA infrared spectra of Tribenzofuran Dimer. . . . .	245
C.65 Raman ad PA infrared spectra of Dibenzofuran Dimer, 700- 2000 cm <sup>-1</sup> . . . . .	246
C.66 Interaction energies of 4-methyldibenzofuran dimers (kcal/mol) using $\omega$ B97X-D/6-311++G**. . . . .	247
C.67 Calculated low wavenumber Raman ad PA infrared spectra of 4-methyldibenzofuran Dimer. . . . .	247
C.68 Interaction energies of 4,6-dimethyldibenzofuran dimers (kcal/mol) in 3 different configurations using $\omega$ B97x-D/6-311++G**. . . . .	248
C.69 Calculated low wavenumber Raman ad PA infrared spectra of 4,6-dimethyldibenzofuran Dimer. . . . .	248
C.70 Raman ad PA infrared spectra of 4,6-dimethyldibenzofuran Dimer, 700- 2000 cm <sup>-1</sup> . . . . .	249
C.71 Interaction energies of Indene dimers (kcal/mol) in 8 different configurations using $\omega$ B97X-D/6-311++G**, and compared with SAPT-DFT. . . . .	250
C.72 Intermolecular distance between ciclopentadiene rings center of indene dimers (Å). . . . .	250
C.73 Calculated low wavenumber Raman ad PA infrared spectra of Indene Dimer. . . . .	250
C.74 Calculated Raman and PA infrared spectra of Indene Dimer, 700–2000 cm <sup>-1</sup> . . . . .	251

C.75 Calculated Raman and PA infrared spectra of Indene Dimer, 700–2000 cm <sup>-1</sup> . . . . .	252
C.76 Interaction energies of Fluorene dimers (kcal/mol) in 3 different configurations using $\omega$ B97x-D/6-311++G**. . . . .	252
C.77 Calculated low wavenumber Raman ad PA infrared spectra of Fluorene Dimer. . . . .	253
C.78 Raman ad PA infrared spectra of Fluorene Dimer, 700- 2000 cm <sup>-1</sup> . . . . .	253
C.79 Interaction energies of Cyclopentaphenanthrene dimers (kcal/mol) in 3 different configurations using $\omega$ B97x-D/6-311++G**. . . . .	255
C.80 Calculated low wavenumber Raman ad PA infrared spectra of Cyclopentaphen- anthrene Dimer. . . . .	255
C.81 Interaction energies of 1-methylfluorene dimers (kcal/mol) in 8 differents config- urations using $\omega$ B97X-D/6-311++G**. . . . .	256
C.82 Calculated low wavenumber Raman ad PA infrared spectra of 1-methylfluorene Dimer. . . . .	256
C.83 Interaction energies of 18-dimethylfluorene dimers (kcal/mol) in 3 different con- figurations using $\omega$ B97x-D/6-311++G**. . . . .	257
C.84 Calculated low wavenumber Raman ad PA infrared spectra of 18- dimethylfluorene Dimer. . . . .	257

# **Introduction**

Les asphaltènes représentent une fraction du pétrole brut, définie sur la seule base de ses propriétés de solubilité. Ainsi, appartiennent à la famille des asphaltènes les composés chimiques du pétrole solubles dans les solvants aromatiques, tels que le toluène, mais insolubles dans les *n*-alcanes, tels le *n*-pentane ou le *n*-hexane. Cette définition, limitée à une unique propriété physique, laisse comprendre implicitement que la nature chimique des asphaltènes, c'est-à-dire leur composition aussi bien que le type et la forme des différentes structures moléculaires qui les constituent, reste inconnue. En somme, ces systèmes chimiques sont pareils à un puzzle dont on ne possèderait que l'emballage. Les propriétés macroscopiques – la solubilité – correspondent à l'image du puzzle présentée sur la boîte, mais le nombre de pièces, de même que leurs formes et la manière dont elles s'assemblent, nous sont inconnus. Tel est le défi auquel font face depuis de nombreuses années les chercheurs, aidés de leur palette de techniques expérimentales et de méthodes théoriques. L'enjeu est de taille, tant cette famille de composés pose de problèmes, aux divers stades de l'extraction, de la production et du raffinage du pétrole. Dans l'industrie pétrolière, le seul mot “asphaltène” est invariablement associé aux phénomènes d'agrégation caractéristiques de ces systèmes, responsables entre autres de l'obstruction des pipelines.

Au fil des années, diverses techniques de caractérisation ont permis de lever le voile sur une partie des propriétés chimiques de ces systèmes. Il est désormais admis que les asphaltènes sont formés d'un ensemble de structures moléculaires composées de noyaux aromatiques poly-condensés, ramifiés par des chaînes alkyles. La majeure partie des propriétés physiques macroscopiques de ces systèmes, notamment leur propension à l'agrégation et leur comportement en solubilité, découlent toutefois de ce qu'ils présentent un taux significatif d'hétéroatomes (N, S et O) et de métaux (V et Ni), à l'origine d'interactions intermoléculaires caractéristiques, comme les liaisons hydrogène, les interactions acido-basiques, la formation de complexes métalliques, les phénomènes de “ $\pi - \pi$  stacking” et de “ $H - \pi$  stacking”. Comme il sera détaillé dans le **premier chapitre** du présent travail, l'ensemble des techniques de caractérisation potentiellement utilisables sur ces fractions les plus lourdes du pétrole brut ont été employées à maintes reprises, au gré des améliorations technologiques, pour tenter de percer à jour la structure chimique intime des asphaltènes. Bien que ces travaux aient pour certains offert des avancées remarquables, nombreuses sont les études – expérimentales comme théoriques – dont les conclusions *a priori* contradictoires suscitent la controverse. Si l'on a pu jusqu'à présent identifier quelques-uns des atomes à l'origine des interactions responsables des phénomènes parasites observés aux différentes étapes de la production du pétrole, l'état actuel des connaissances ne nous permet pas de donner une définition structurale précise pour ces systèmes.

Le présent travail de thèse prend naissance dans l'idée originale proposée en 2011 par J. Shaw et K. Michaelian, d'employer les spectroscopies infra-rouge (IR) et Raman – en tant que techniques de caractérisation courantes – à la base d'un “protocole” innovant de caractérisation moléculaire des asphaltènes. Dans le projet proposé par ces chercheurs, la spectroscopie IR sert de pivot, et sa simplicité de mise en œuvre, en même temps que les principes physiques sur lesquels elle repose, permettent de la coupler aisément à une large variété de techniques expérimentales.

En première étape à ce “protocole”, J. Shaw et K. Michaelian proposent d'employer des analyses IR et Raman à la constitution d'une large base de données, dont le but serait

de faciliter l'identification des motifs moléculaires constitutifs des asphaltènes. Dans une seconde étape, le couplage des techniques d'IR lointain et de photo-acoustique permettrait de caractériser la présence – ou non – d'interactions de type  $\pi - \pi$ . Cet approfondissement présenterait alors le double intérêt de vérifier la validité des familles de molécules précédemment identifiées, et d'associer à ces interactions une fréquence et une intensité, deux paramètres physiques caractéristiques.

Pour répondre à ces attentes, le soutien des méthodes de la chimie quantique pour la caractérisation des propriétés vibrationnelles apparaît crucial, et c'est dans ce cadre que s'inscrit ce travail de thèse. En aidant à l'identification des signatures caractéristiques des différentes familles de composés et des modes inter-moléculaires qui leurs sont associés, la chimie calculatoire offre un guide précieux aux expérimentateurs, et contribue à la construction des deux étapes du schéma proposé par J. Shaw et K. Michaelian.

Au vu de la diversité des asphaltènes, du point de vue de la nature chimique et de la taille, la première partie de ce travail de thèse a consisté en une réflexion en termes de stratégie calculatoire et méthodologique, le but étant d'accéder aux données vibrationnelles calculées dans les hypothèses harmonique et anharmonique. Ce travail préliminaire, véritable clé de voûte de notre investigation théorique, sera détaillé dans le **second chapitre** du présent manuscrit.

Au-delà de la stratégie calculatoire, la nature méconnue des asphaltènes et leur mélange au sein des fractions lourdes du pétrole posent de nombreux problèmes de mise en œuvre. Tout d'abord, la diversité des composés chimiques constitutifs de ces systèmes nous a contraint à faire le choix de molécules modèles judicieuses, aptes à représenter les différentes classes de composés postulées comme présentes au sein des asphaltènes. Par ailleurs, le comportement en mélange de ces différentes familles de composés induit, dans le cadre des modélisations, deux problématiques majeures, que sont :

- la prise en compte des effets de l'environnement chimique d'un système, de sorte à apporter une description théorique aussi proche que possible des conditions de mesure
- la traduction, en termes de description quantique, de l'ensemble des interactions identifiées entre ces systèmes, et notamment les interactions longue portée.

La taille des systèmes à l'étude a imposé le recours aux méthodes de la théorie de la fonctionnelle de la densité (DFT). Toutefois, il est bien connu que le talon d'Achille de ces méthodes demeure la description des forces de dispersion, pour laquelle les avancées méthodologiques sont encore lentes et parfois mises en défaut. La nécessité, dans la présente étude, de traduire les effets d'interactions de type longue portée, nous impose de présenter, dans le **troisième chapitre**, les solutions proposées jusqu'à ce jour pour palier certains des défauts de la DFT dans ce domaine. Cette partie abordera notamment la notion de fonctionnelle hybride et détaillera le principe des approches employées dans ce travail, qui ajoutent *ad hoc* à un calcul Kohn-Sham (KS) usuel des fonctions semi-empiriques. Ces approches sont reconnues aujourd'hui comme étant les plus aptes à rendre compte des effets de dispersion dans le cadre de l'utilisation des méthodes de la DFT.

À la suite de ces trois chapitres, indispensables pour établir la méthodologie suivie pour répondre à notre problématique, nous reviendrons dans une quatrième partie sur la mise en

œuvre pratique de notre stratégie calculatoire. Celle-ci présentera donc les calculs préliminaires menés sur des molécules test (à savoir des composés thiophéniques), pour lesquelles nous disposons d'une grande quantité d'informations moléculaires de référence, afin d'évaluer la fiabilité des approximations auxquelles nous avons eu recours. En effet, comme il aura été explicité plus avant, le traitement quantique de ces systèmes de grande dimension, qui repose sur la résolution de l'équation vibrationnelle de Schrödinger (REVS) dans la double hypothèse des anharmonicités électrique et mécanique, requiert de réduire la taille de l'espace actif des configurations vibrationnelles. De fait, dans l'objectif de parvenir à un ratio acceptable temps *versus* précision des calculs réalisés, nous nous sommes employés à la construction de bases réduites de modes normaux. L'emploi de ces bases réduites, déduites des couplages observées entre les modes inter-moléculaires et les modes de torsion angulaire et de stretching des liaisons C-H, sera donc corroboré et justifié par les résultats présentés au **chapitre 4**.

Les trois derniers chapitres de ce manuscrit présentent l'ensemble des résultats obtenus par le biais des stratégies calculatoires mises en place, en gardant à l'esprit que notre travail se veut traiter des différentes problématiques inhérentes aux asphaltènes, que sont : leur nature chimique, et particulièrement le rôle des hétéroéléments, l'influence de l'environnement et les interactions à l'origine de l'agrégation caractéristique de ces systèmes.

C'est pourquoi le **chapitre 5** se propose de décrire les résultats des calculs moléculaires réalisés dans le but d'élucider le rôle singulier des hétéroatomes dans les signatures vibrationnelles de ces systèmes. Ces calculs, qui emploient les outils méthodologiques développés pour la REVS, ont pour double objectif :

- de déterminer la signature propre de molécules modèles isolées
- d'analyser les modifications respectives de ces signatures sous l'effet des interactions inter-moléculaires, simulées sur des dimères.

Afin d'assurer une analyse complète, nous avons mené nos calculs sur une collection de six motifs moléculaires, pour les trois hétéroéléments (N, S et O), ainsi que sur des motifs identiques pris pour références, ne comportant que des atomes de carbone.

Le **chapitre 6** se concentre quant à lui sur l'effet de l'environnement, simulé par des calculs à l'état solide. Ce dernier constitue en effet l'état le plus proche des conditions réelles, pour lequel nous possédons des données expérimentales de référence obtenues en photo-acoustique sur des systèmes cristallins. Le passage de l'état moléculaire à l'état solide induit bien évidemment de profondes modifications dans la description des modes de vibration intermoléculaires. Nos calculs visent ainsi à caractériser explicitement ces modes en identifiant des singularités dans les signatures spectrales, avec pour objectif final de fournir aux expérimentateurs des caractéristiques précises leur permettant d'identifier ou non la présence de systèmes associés.

Enfin, le **chapitre 7** de ce manuscrit développe les conclusions de travaux menés dans le sens d'une meilleure compréhension des phénomènes d'agrégation des asphaltènes. Les calculs réalisés dans cet objectif se placent donc à une toute autre échelle et permettent d'inscrire le présent travail dans une étude complète de type "bottom-up" menée au laboratoire. Seront ainsi présentées dans cette partie des simulations en dynamique moléculaire, dont l'objectif est de comprendre le rôle des interactions intermoléculaires dans la nano-agrégation (les échelles supérieurs étant la micro agrégation et la flocculation toutes deux responsables des désagrégements

observés par les industriels à chaque fois que des asphaltènes sont présents dans les matrices de pétrole). Ces simulations seront de deux types, chacune présentant un objectif distinct :

- dans un premier temps, nous considérerons des systèmes modèles des asphaltènes, dans le but de comprendre l'influence de la position des hétéroéléments au sein de la molécule. En effet, suivant leur position, différentes sortes d'interactions spécifiques se créent (liaisons hydrogène, interactions  $\pi - \pi$  ou  $\pi - H$ , etc.)
- dans un second temps, nous nous baserons sur des systèmes moléculaires mis en évidence très récemment par le biais de la spectroscopie XPS, et pouvant être considérés *a priori* comme plus réalistes. L'intérêt est ici d'étudier le comportement des agrégats suivant les proportions dans lesquelles sont mélangés ces systèmes et suivant la nature et la teneur des hétéroatomes qui les constituent.

# Chapter 1

## Généralités – Contexte de l'étude

# Introduction

Les enjeux environnementaux mondiaux nous imposent depuis plusieurs années de repenser notre rapport aux énergies et, notamment, de revoir l'exploitation des énergies fossiles, première ressource mondiale devant la filière nucléaire, et ressource épuisable. En 2016, aucun scénario ni aucune projection ne peut encore, raisonnablement, se passer des ressources pétrolières, impliquant des réflexions et des recherches poussées pour faire évoluer les procédés industriels actuels d'extraction et de raffinage du pétrole. En premier lieu, la raréfaction des réserves de pétrole brut léger a progressivement conduit à l'exploitation de fractions plus lourdes du pétrole brut, *via* l'introduction d'étapes de désulfuration dans le processus industriel de distillation du pétrole. En second lieu, les enjeux liés à l'amélioration de ces procédés reposent tant sur l'amélioration de la qualité des produits de distillation (au travers de procédés de désulfuration plus adaptés) que sur la distillation de fractions toujours plus lourdes de pétrole brut, permettant une exploitation plus durable de l'énergie fossile.

Dans cette optique, il est indispensable de comprendre qu'une amélioration des procédés passe obligatoirement par une meilleure compréhension de la composition chimique des différents produits de distillation du pétrole brut. En effet, il ne s'agit pas simplement de dériver les méthodes d'extraction et de raffinage employées communément pour le traitement des fractions plus légères du pétrole brut ; le traitement des fractions lourdes nécessite de sonder plus précisément la nature chimique des molécules qui les constituent pour comprendre les propriétés physiques qui en découlent et être à même de créer des procédés innovants assurant un raffinage efficace d'un nombre croissant de fractions du pétrole brut. De façon logique, il apparaît indispensable d'étudier aussi bien la composition chimique des constituants "exploitables" de ces fractions lourdes que des constituants "non exploitables", dont le comportement physico-chimique induit de nombreux désagréments au cours des différentes phases de l'exploitation industrielle du pétrole. Les asphaltènes, objets de la présente étude, appartiennent à cette dernière catégorie. La méconnaissance de ces systèmes engendre de sérieuses limitations, qui vont de la baisse de rendement à l'obstruction des pipelines. La caractérisation des asphaltènes constitue donc un enjeu considérable.

Le présent chapitre se veut une présentation du cadre de ce travail de thèse, au travers d'une définition des asphaltènes et d'un état des lieux des connaissances établies sur ces systèmes. Après une description générale de la composition chimique du pétrole, et une description des systèmes asphalténiques au moyen d'un historique des études expérimentales publiées jusqu'à ce jour, nous verrons comment la modélisation s'impose comme un outil indispensable dans la caractérisation de ces composés complexes.

## 1.1 Généralités

### 1.1.1 Composition du pétrole brut

#### 1.1.1.1 Définitions

Le pétrole brut est un matériau provenant de la transformation physico-chimique de résidus animaux et végétaux sur de longues périodes géologiques. Les variations locales de température et de pression au sein des couches géologiques en font un matériau non uniforme, dont la com-

position chimique varie sensiblement suivant de multiples paramètres, tels que la localisation géographique, l'âge du filon exploité ou encore la profondeur de forage. Constitué en majeure partie de carbone et d'hydrogène, sa composition chimique fait néanmoins apparaître en proportions non négligeables des hétéroéléments (azote, oxygène, soufre), de même que des métaux (majoritairement du nickel et du vanadium). Speight [1] résume la composition chimique du pétrole en termes de distribution élémentaire :

Élément chimique	Proportions
Carbone	83.0–87.0%
Hydrogène	10.0–14.0%
Azote	0.1–2.0%
Oxygène	0.05–1.5%
Soufre	0.05–6.0%
Métaux (Nickel, Vanadium)	<1000 ppm

Table 1.1 – Composition chimique du pétrole en termes de distribution élémentaire, selon les données recueillies par Speight *et al.*

Les données rassemblées par Speight et son équipe font apparaître des fenêtres de distribution relativement étroites, qui engendrent pourtant de grandes variations dans les propriétés physiques observées (volatilité, viscosité, gravité spécifique, couleur, etc.). De fait, la définition des nombreuses fractions du pétrole découle souvent de leurs propriétés physiques, et particulièrement des propriétés de viscosité et de densité. Les définitions les plus communes font intervenir la densité API (baptisée ainsi car conçue par le “American Petroleum Institute” et exprimée en ° API), échelle spécifique au pétrole brut, calculée suivant la formule suivante :

$$d_{API} = \frac{141.5}{d_{60F}} - 131.5 \quad (1.1)$$

où  $d_{60F}$  représente la densité de la fraction de pétrole brut à 60 ° Farenheit.

Suivant cette formule, plus la fraction étudiée est légère - autrement dit plus sa densité est faible - plus la densité API est élevée.

De nombreux termes courants ont été définis suivant cette échelle spécifique. Ainsi, le terme "huile lourde" désigne une fraction dont la densité API n'excède pas 20 ° API. Faisant partie des huiles lourdes, et souvent considérés comme huiles "extra-lourdes", les sables bitumeux sont parfois définis comme des fractions de densité API inférieure à 12 ° API.

### 1.1.1.2 Classification des différentes fractions du pétrole brut

Comme évoqué dans le paragraphe précédent, le pétrole est un matériau non-homogène de composition chimique complexe. L'industrie pétrolière divise communément les constituants chimique du pétrole en trois catégories, que sont :

- les *paraffines*, hydrocarbures saturés de taille variable (de type *n*-alcanes et *iso*-alcanes) ;
- les *naphtènes*, constitués majoritairement d'hydrocarbures saturés cycliques (en majorité des cycloalcanes) ;

- les *aromatiques*, dont le benzène et ses dérivés.

Les résines et les asphaltènes appartiennent à cette dernière catégorie, en tant que molécules de grande taille comportant de nombreux cycles aromatiques ainsi que plusieurs hétéroatomes. Le paragraphe qui suit revient plus en détail sur les caractéristiques physico-chimiques des asphaltènes, objets de la présente étude.

## 1.1.2 Une fraction lourde du pétrole brut : les asphaltènes

### 1.1.2.1 Définition

Les asphaltènes sont des molécules de grande taille, généralement observées sous forme de poudre noire, hautement polaires, constituées de noyaux aromatiques condensés, substitués par des chaînes aliphatiques et naphténiques. Ces arrangements hétérocycliques comportent plusieurs hétéroéléments (azote, oxygène, soufre), en proportion variable suivant le procédé d'extraction employé, comme évoqué dans les paragraphes précédents [2].

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[3]. Dans ce contexte, les “asphaltènes” sont définis comme les constituants solides insolubles dans l'alcool mais solubles dans l'essence de térébenthine, par opposition aux “pétrolènes”, constituants volatils solubles dans l'éther. Désormais obsolète, cette définition des asphaltènes fait néanmoins appel aux propriétés de solubilité de ces systèmes, caractéristique que l'on emploie de nouveau aujourd'hui pour tenter de définir ces composés sur la base de leurs propriétés physiques. Ainsi, à l'heure actuelle, les asphaltènes sont définis comme une fraction du pétrole soluble dans les solvants aromatiques ”légers” (tels que le benzène) mais insolubles dans les solvants paraffiniques (tels que les *n*-alcanes). Une définition plus précise passe par une classification des asphaltènes suivant le type de solvant paraffinique apte à provoquer leur précipitation. L'étude menée par Speight *et al.* [4] tend à montrer la corrélation existant entre le caractère aromatique de l'asphaltène (indiqué par le rapport entre le nombre d'atomes d'hydrogène et le nombre d'atomes de carbone) et la masse molaire de la fraction d'asphaltène précipitée dans le solvant paraffinique.

Si l'on considère les propriétés physico-chimiques des asphaltènes et la nature des constituants chimiques qui composent les différentes fractions du pétrole, il apparaît rapidement qu'il existe des incompatibilités physiques, qui rendent problématique la présence des asphaltènes dans les différentes phases de distillation et de raffinage du pétrole. En effet, les asphaltènes ayant tendance à précipiter au contact de solvants de type paraffiniques, leur mélange aux autres constituants du pétrole déclenche un phénomène d'agrégation. Les asphaltènes forment des agrégats baptisés micelles, et le système constitué par les asphaltènes et les fractions hydrocarbonées du pétrole forme une solution colloïdale. Les procédés intervenant à tous les stades de la production industrielle du pétrole se doivent de tenir compte de ce système colloïdal, en retenant les micelles ou en les solvatant. Cette problématique est rendue un peu plus complexe encore par les nécessaires changements de température et/ou de pression aux différents stades de la distillation et du raffinage, paramètres qui influent notablement sur les propriétés de précipitation des asphaltènes.

### 1.1.2.2 Composition chimique des asphaltènes : une teneur importante en hétéroéléments

**Soufre** Au sein du pétrole, la composition élémentaire en soufre varie généralement entre 1% et 4%. La majorité (60 à 80%) des composés soufrés se retrouve dans la fraction asphalténique du pétrole, ce qui les rend *a priori* plus complexes à caractériser. On suppose néanmoins que les dérivés soufrés proprement caractérisés au sein des fractions légères du pétrole sont semblables aux composés soufrés présents dans les fractions asphalténiques, mais en proportions différentes. En règle générale, ces dérivés se divisent en cinq grandes classes chimiques : thiols, sulfures, disulfides, sulfoxydes et thiophènes.

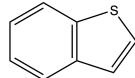
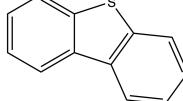
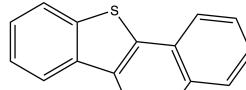
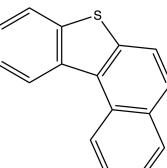
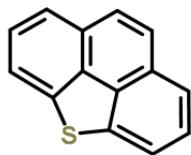
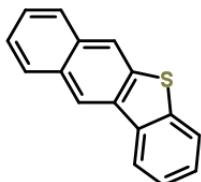
RSR'	Thiols
R(S=O)R'	Sulfoxydes
RSR'	Sulfures
	Sulfures cycliques
RSSR'	Disulfures
	Thiophène
	Benzothiophène
	Dibenzothiophène
	Naphtobenzothiophènes
	

Table 1.2 – Famille de composés soufrés caractérisés au sein du pétrole brut

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 thiophène (benzo, dibenzo, naphtobenzo-thiophènes, etc.). Au sein des fractions lourdes, la majorité des espèces soufrées sont des dérivés thiophéniques (cf. tableau 1.3), suivi par des dérivés sulfures (cycliques et acycliques). Enfin, des espèces de type sulfoxydes ont été mises en évidence, dans des proportions très variables, leur teneur variant entre 0.3% et 10.3% [5, 4].



Phénanthro-thiophènes



Benzonaphto-thiophènes

Table 1.3 – Exemples de dérivés thiophéniques courants identifiés au sein des fractions lourdes : phénanthro-thiophènes et benzonaphtho-thiophènes

**Azote** En dépit des faibles proportions dans lesquelles il est présent, les dérivés de l'azote sont particulièrement étudiés, puisqu'à l'origine de la destruction des catalyseurs employés lors du craquage. Comme leurs analogues soufrés, les composés azotés se retrouvent en majeure partie au sein des fractions lourdes et des résidus de distillation du pétrole.

On distingue communément deux classes de composés azotés, à savoir les dérivés dits "*basiques*", qui regroupent les homologues de la pyridine, et les dérivés dits "*neutres*", qui rassemblent des systèmes de type pyrrole, indole et carbazole. En règle générale, environ un tiers des composés azotés sont des dérivés "*basiques*", la part restante revenant donc aux composés dits "*neutres*", bien que ces proportions puissent varier sensiblement en fonction des paramètres géochimiques de provenance du pétrole.

Comme indiqué dans le tableau 1.4, les structures "*basiques*" identifiées sont des analogues de la pyridine, tels que la quinoléine ou l'iso-quinoléine, comprenant entre deux et quatre cycles aromatiques condensés avec différents degrés d'alkylation ; des benzo- et dibenzoquinoléines, tétrahydroquinoléines, ainsi que des azapyrènes, ont par exemple pu être caractérisés, qui forment autant de systèmes à haut encombrement stérique.

Les structures "*neutres*", majoritaires, sont essentiellement représentées par les dérivés carbazolés (benzo-carbazoles, dibenzo-carbazole, etc.). C'est également au sein de cette classe de composés que l'on trouve les porphyrines qui, bien que présentes en proportions plus anecdotiques, possèdent une stabilité remarquable. La brique élémentaire de ces systèmes de grande taille est la molécule de pyrrole, et la plus simple des porphyrines est constituée par l'assemblage de quatre molécules de pyrrole, reliées par des ponts méthylène =CH-. Le large système de résonance ainsi formé augmente considérablement la stabilité des structures ainsi formées. Par ailleurs, les porphyrines ont tendance à former des complexes métalliques avec les ions nickel et vanadium présents au sein du pétrole, par substitution des deux atomes d'hydrogène liés aux atomes d'azote de la structure porphyrinique [6]. La proportion de ces structures au sein du pétrole dépend une nouvelle fois de l'origine géochimique de ce dernier, et un nombre croissant d'études s'intéresse à la part de métaux présents dans les huiles brutes.

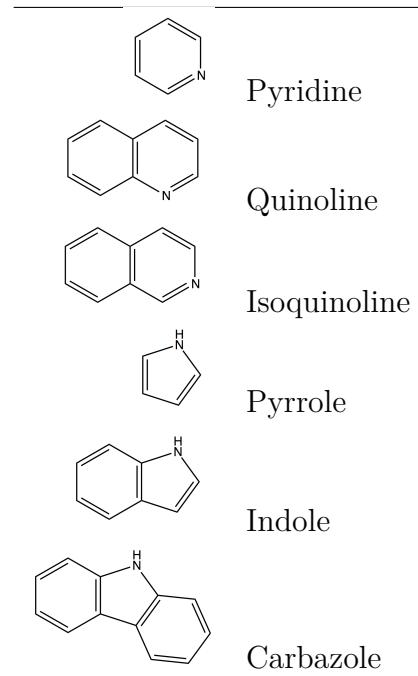


Table 1.4 – Famille de composés azotés caractérisés au sein du pétrole brut

On estime aujourd’hui que ces complexes porphyrine-métaux représentent entre 0.6 et 3.3% des fractions lourdes extraites[5, 4].

**Oxygène** Au sein des fractions lourdes qui nous intéressent, la teneur en composés oxygénés varie entre 0.3 et 4.9% [4]. Ces composés oxygénés figurent parmi les premières espèces caractérisées du pétrole. Les premiers travaux reportant la présence d’espèces acides au sein du pétrole remontent à 1874, et il faudra attendre quelques années de plus avant que celles-ci ne soient formellement identifiées comme étant des acides carboxyliques. Avec les dérivés du phénol, les acides carboxyliques représentent la majorité des espèces oxygénées constituant le pétrole, que l’on retrouve aussi bien dans les fractions lourdes qu’au sein de fractions plus légères. L’ensemble des études réalisées sur le sujet tend à établir que la structure de ces acides carboxyliques correspond à la structure de l’hydrocarbure prévalent. En d’autres termes, la structure de l’acide dépend du nombre d’atomes de carbone que compte le squelette carboné de l’hydrocarbure. Ainsi, lorsque l’acide comporte moins de huit atomes de carbone, ce qui correspond à un hydrocarbure de type paraffinique comme évoqué ci-avant, sa structure sera de type aliphatique. Suivant cette logique, des acides comportant un cycle commencent à apparaître à partir de six atomes de carbone et ces structures monocycliques prédominent jusqu’à quatorze atomes de carbone environ.

Une caractérisation plus fine des groupements carboxyles en présence révèle également la co-existence d'esters, de cétones, d'amides, d'anhydrides, d'éthers et de sulfoxydes. Néanmoins, à la différence des acides carboxyliques, plus aisés à caractériser puisque présent tant dans les fractions lourdes que dans les fractions légères, ces composés oxygénés se retrouvent majoritairement dans les fractions lourdes et leurs structures sont de fait plus complexes à identifier. Il est par ailleurs à noter que certains de ces composés sont supposés provenir de réactions entre les résidus de distillation et l'oxygène soufflé, qui intervient notamment dans les procédés de production de l'asphalte. De fait, il est difficile de conclure quant au fait que toutes ces espèces

sont ou non des constituants du pétrole brut.

### 1.1.2.3 Débat autour du 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<sup>1</sup> pour les fractions de faible poids moléculaire, et peut atteindre 10<sup>6</sup> Da pour les fractions de haut poids moléculaire [7]. 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 ambigus, qu'il serait hasardeux de généraliser. À 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 tel que la pyridine ou le nitrobenzène.

## 1.2 État des lieux des connaissances sur les systèmes asphalténiques

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 cœurs constitués de quatre à dix cycles aromatiques condensés, ramifiés par des chaînes alkyles courtes [8] (cf. figure 1.1). Ce modèle a été proposé par Zhao *et al.* [9] après caractérisation des asphaltènes obtenus par fractionnement avec du pentane supercritique, puis remployé par Rogel et Carbognani [10] 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).

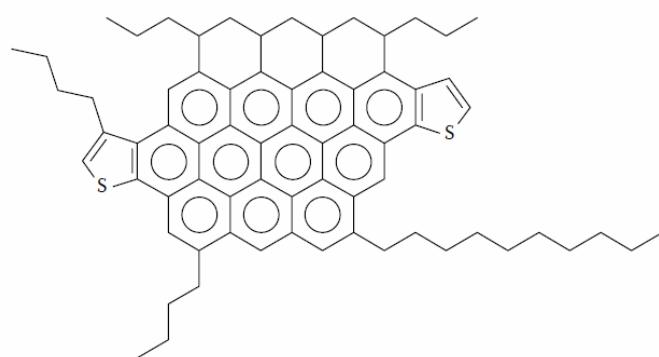


Figure 1.1 – Structure moyenne du modèle continental des asphaltènes (Zhao et col. [9])

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.* [11]. Sheremata

<sup>1</sup>Un dalton est défini comme égal à 1/12 de la masse d'un atome de carbone 12

*et al.*[12] ont notamment employé ce modèle pour la description des asphaltènes, comme le présente la figure reproduite ci-après (figure 1.2).

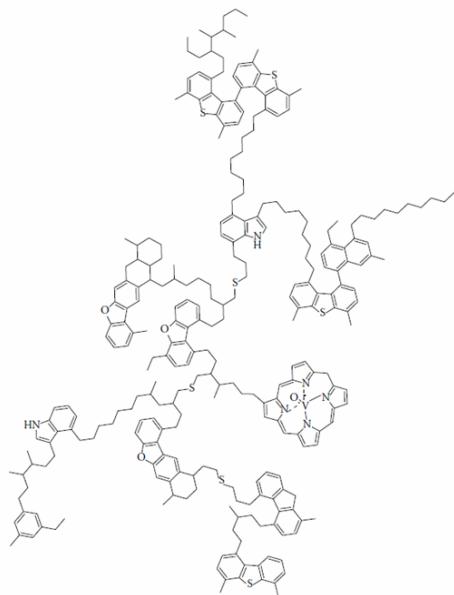


Figure 1.2 – Représentation d'asphaltènes à l'aide du modèle archipel, proposé par Sheremata *et al.*[12]

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'une ou l'autre de ces représentations. 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. À l'inverse, les chaînes alkyles du modèle archipel confèrent une toute autre géométrie à ces systèmes : les asphaltènes peuvent se courber par le biais d'interactions intermoléculaires pour former des macro-agrégats globulaires susceptibles de piéger les molécules de solvant.

Le paragraphe qui suit se veut résumer les nombreuses analyses expérimentales menées sur ces systèmes, analyses qui ont dans un premier temps permis l'émergence des deux modèles que nous venons d'évoquer, avant de travailler à les confirmer ou les infirmer.

## 1.2.1 Revue des données expérimentales

### 1.2.1.1 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.* [13] 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 [14]. 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 [15]. 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.

### 1.2.1.2 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 en diffraction des rayons X [16]. Leurs résultats, confortés par les travaux réalisés par Shirokoff *et al.* sur quatre échantillons d'asphaltènes issus d'huiles brutes provenant d'Arabie Saoudite [17], 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 [18] qu'il convient de prendre avec précaution les résultats provenant d'analyses de 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.

### 1.2.1.3 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 [19], tendent à montrer que les molécules d'asphaltènes ne seraient constituées que d'un seul groupe chromophore. 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 par spectrométrie de fluorescence 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 [20]. Leurs travaux corroborent par ailleurs les résultats de Groenzin *et al.* [19] quant à la structure de type "continental" des molécules d'asphaltènes, en cela qu'ils concluent à une structure primaire constituée d'un anneau polyaromatique de quatre cycles ou plus.

### 1.2.1.4 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. Toutefois, dans le cadre de l'étude des fractions lourdes du pétrole, la diversité chimique et la taille des espèces présentes rendent 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 [21]. Parmi les travaux notables, l'étude couplée en spectroscopies IR et UV-visible menée par El-Bassoussi *et al.* [22] 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. En 2007, Rodrigues Coelho *et al.* [23] 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.* [24] rapportent une interprétation, toutefois ajustée sur les structures modèles des asphaltènes, en repérant les vibrations de stretching des liaisons C-H (hors du plan) et C=C.

### 1.2.1.5 Conclusion : les asphaltènes, une coexistence de 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.* [25, 26] 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 (methylé-benzène) 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 expérimentales menées à ce jour sur les asphaltènes montre qu'aucun modèle structural définitif valide n'a encore émergé pour ce type de systèmes. De nombreux travaux cherchent encore à clarifier l'architecture moléculaire 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.

## 1.2.2 La modélisation, outil indispensable dans l'élucidation des structures asphalténiques

### 1.2.2.1 Des nombreux intérêts de la modélisation 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 le paragraphe précédent, la nature et la complexité des asphaltènes rendent impossible 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 solvants [27]. 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, représentée 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 [28].

### 1.2.2.2 Construction de la structure moléculaire des asphaltènes

La construction d'un modèle moléculaire utilisable en simulation se fait sur la base 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 [29].

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 alors par 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.* [30], qui ont développé un algorithme permettant la construction d'une structure sur la base de données RMN, que par Todeschini *et al.* [31] 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 [32, 33, 34].

# **Chapter 2**

## **Partie vibrationnelle**

## Introduction

La littérature consacrée aux applications des spectrométries vibrationnelles dans le domaine de la caractérisation des constituants présents dans les pétroles est relativement peu abondante, et ce bien que la spectrométrie infrarouge soit devenue une technique d'analyse de "routine" dans de très nombreux laboratoires, académiques comme industriels. La raison de cette faible abondance réside essentiellement dans la complexité des mélanges qui constituent un pétrole, un asphaltène, etc.. Pourtant, les champs d'application de cette technique 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 RMN<sup>1</sup> ou la spectrométrie de masse, la spectrométrie infrarouge à transformée de FOURIER (IRTF) a connu de nouvelles avancées technologiques, parmi lesquelles la spectrométrie photoacoustique que nous avons employée 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, ...

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 également à rappeler les fondements essentiels de la résolution de l'équation vibrationnelle de Schrödinger, qui soutiennent les principales méthodes – que nous développerons – actuellement disponibles pour tenter de répondre aux problématiques posées par les expérimentateurs.

Le domaine de la pétrochimie qui nous intéresse dans ce travail, et précisément la thématique visant à l'élucidation de la composition de mélanges complexes, composés de 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 de compréhension des processus d'interaction et d'agrégation de ces molécules. Dans ce domaine encore, la spectrométrie IRTF est certainement un des outils les plus efficaces 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 supposées présentes dans les asphaltènes.

---

<sup>1</sup>"Nuclear Magnetic Resonance Spectroscopy".

## 2.1 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 couramment utilisées pour répondre à ce problème, la spectroscopie vibrationnelle est indubitablement 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 comme le pétrole 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 Schrödinger (REVS), 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 qu'au manque d'implémentation d'approches de type anharmonique dans les grands codes de calculs commerciaux. 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 REVS. 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 (quelques 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 présents dans les asphaltènes. Ces activités s'inscrivent dans le prolongement et en 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 [35].

## 2.2 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 [36, 37] et repris dans de nombreuses thèses [38, 39].

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 [40]. 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  $\omega$ . Le mouvement de la molécule est défini par le trièdre mobile représentant à chaque instant la position  $\vec{r}_\alpha$  des noyaux  $\alpha$  par rapport à leur position d'équilibre  $\vec{a}_\alpha$ . Soit :

$$\vec{\rho}_\alpha = \vec{r}_\alpha - \vec{a}_\alpha \quad (2.1)$$

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

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

$$\vec{V}_\alpha = \vec{\dot{R}} + (\vec{\omega} \wedge \vec{r}_\alpha) + \vec{v}_\alpha \quad (2.2)$$

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

$$\begin{aligned}
2T = & \dot{R}^2 \sum_{\alpha} m_{\alpha} (\vec{\omega} \wedge \vec{r}_{\alpha})^2 + \sum_{\alpha} m_{\alpha} v_{\alpha}^2 \\
& + 2 \vec{\dot{R}} \sum_{\alpha} m_{\alpha} \vec{v}_{\alpha} \\
& + 2 \left( \vec{\dot{R}} \wedge \vec{\omega} \right) \sum_{\alpha} m_{\alpha} \vec{r}_{\alpha} \\
& + 2 \vec{\omega} \sum_{\alpha} m_{\alpha} (\vec{r}_{\alpha} \wedge \vec{v}_{\alpha})
\end{aligned} \tag{2.3}$$

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 :

$$\sum_{\alpha} m_{\alpha} \vec{v}_{\alpha} = 0 \tag{2.4}$$

et

$$\sum_{\alpha} m_{\alpha} \vec{r}_{\alpha} = 0 \tag{2.5}$$

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

$$\sum_{\alpha} m_{\alpha} (\vec{a}_{\alpha} \wedge \vec{v}_{\alpha}) = 0 \tag{2.6}$$

et donc

$$\sum_{\alpha} m_{\alpha} (\vec{r}_{\alpha} \wedge \vec{v}_{\alpha}) = \sum_{\alpha} m_{\alpha} (\vec{\rho}_{\alpha} \wedge \vec{v}_{\alpha}) \tag{2.7}$$

Les deux conditions ci-dessus portent le nom de conditions d'Eckart et simplifient l'expression 2.3 :

$$2T = \dot{R}^2 \sum_{\alpha} m_{\alpha} + \sum_{\alpha} m_{\alpha} (\vec{\omega} \wedge \vec{r}_{\alpha})^2 + \sum_{\alpha} m_{\alpha} v_{\alpha}^2 + 2 \vec{\omega} \sum_{\alpha} m_{\alpha} (\vec{r}_{\alpha} \wedge \vec{v}_{\alpha}) \tag{2.8}$$

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 [41], est généralement vérifiée pour les vibrations d'élongation, 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 :

$$2T_n = 2T_R + 2T_V \text{ en supposant } 2T_{VR} = 0 \tag{2.9}$$

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 :

$$\psi_{R_\alpha}^n = \psi_{R_\alpha}^R \bullet \psi_{R_\alpha}^V \quad (2.10)$$

$$E_n = E_V + E_R \quad (2.11)$$

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 :

$$(\hat{T}_V + \hat{V}_V) \psi_{R_\alpha}^V = E_V \psi_{R_\alpha}^V \quad (2.12)$$

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):

$$T_R \psi_{R_\alpha}^R = E_R \psi_{R_\alpha}^R \quad (2.13)$$

## 2.3 Energie cinétique de vibration

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

$$2T_V = \sum_{\alpha}^n m_{\alpha} (\dot{x}_{\alpha}^2 + \dot{y}_{\alpha}^2 + \dot{z}_{\alpha}^2) \quad (2.14)$$

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

En exprimant cette énergie dans le système des 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, fonction explicite de  $3n$  coordonnées :

$$2T_V = \sum_i^{3n} \dot{q}_i^2 \quad (2.15)$$

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

$$2T_V = [\dot{q}]^t [\dot{q}] \quad (2.16)$$

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 :

$$2T_V = [\dot{x}]^t [M_{\alpha}] [\dot{x}] \quad (2.17)$$

## 2.4 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-avant, elle prend la forme d'un polynôme caractéristique d'ordre  $n$ , dont nous limitons le développement à l'ordre 2 dans l'hypothèse harmonique :

$$V = V_{eq} + \sum_i^{3n} \left( \frac{\partial V}{\partial q_i} \right)_{eq} q_i + \frac{1}{2!} \sum_{i \leq j}^{3n, 3n} \left( \frac{\partial^2 V}{\partial q_i \partial q_j} \right)_{eq} q_i q_j \quad (2.18)$$

Les coefficients de ce polynôme représentent les dérivées  $n^{\text{ièmes}}$  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 :

$$2V = \sum_{i \leq j}^{3n, 3n} f_{ij}^{(q)} q_i q_j \quad (2.19)$$

soit, sous sa forme matricielle :

$$2V = [q]^t [f^q] [q] \quad (2.20)$$

## 2.5 Équations de Lagrange

### 2.5.1 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-paragraphe précédents :

$$\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_i} \right) + \frac{\partial V}{\partial q_i} = 0 \quad (2.21)$$

soit :

$$\ddot{q}_i + \sum_j^{3n} f_{ij}^{(q)} q_j \quad (2.22)$$

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 ces 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 mouvements de translation et aux trois mouvements de rotation de la molécule (deux mouvements de rotation si la molécule est linéaire) et  $3n - 6$  valeurs propres non nulles (5 dans le cas d'une molécule linéaire), correspondant aux mouvements de vibration de la molécule. Nous déduisons de ces valeurs propres les nombres d'ondes  $\varpi$  (exprimés en  $\text{cm}^{-1}$ ) et les fréquences de vibration  $\omega$  (en Hz) par la relation :

$$\lambda^{1/2} = 2\pi c\varpi = 2\pi\omega \quad (2.23)$$

avec  $c$  : célérité de la lumière dans le vide.

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  $3n - 6$  (ou  $3n - 5$ ) 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 du Hamiltonien vibrationnel dans le traitement quantique.

### 2.5.2 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 mouvements de translation et de rotation 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 complexe, 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 déplacement  $r$  :  $[r] = [B][x]$ . D'après l'équation 2.17, et dans l'hypothèse d'une transformation à coefficients constants, valable pour les petits mouvements, l'énergie cinétique s'écrit :

$$2T = [\dot{r}]^t [G^{-1}] [\dot{r}] \quad (2.24)$$

où :

$$[G^{-1}] = [B^{-1}]^t [M_\alpha] [B^{-1}] \quad (2.25)$$

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

$$2V = [r]^t [f^{(r)}] [r] \quad (2.26)$$

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

$$[G] [f^{(r)}] [L] = [\lambda] [L] \quad (2.27)$$

où  $[L]$  est la matrice des vecteurs propres.

### 2.5.3 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 propriété d'être bloc-symétriques. Une étude intéressante a été menée par PULAY [42] sur la construction de cet espace en fonction des différents groupements fonctionnels de composés organiques.

## 2.6 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 :

$$2T = \sum_{i=1}^{3n-6(5)} \dot{Q}_i^2 = [\dot{Q}]^t [\dot{Q}] \quad (2.28)$$

$$2V = \sum_{i=1}^{3n-6(5)} \lambda_i Q_i^2 = [Q]^t [\lambda] [Q] \quad (2.29)$$

et le Hamiltonien correspondant s'écrit :

$$\hat{H} = \hat{T} + \hat{V} = \frac{1}{2} \sum_{i=1}^{3n-6(5)} (\dot{Q}_i^2 + \lambda_i Q_i^2) \quad (2.30)$$

En associant à ces grandeurs leur opérateur correspondant,

$$\hat{Q} = Q \quad (2.31)$$

$$\hat{\dot{Q}} = \hat{P} = -i\hbar \frac{\partial}{\partial Q} \quad (2.32)$$

nous obtenons l'équation de Schrödinger vibrationnelle :

$$\sum_{i=1}^{3n-6(5)} \frac{\partial^2 \Psi}{\partial Q_i^2} + \frac{2}{\hbar^2} \left( E - \frac{1}{2} \sum_{i=1}^{3n-6(5)} \lambda_i Q_i^2 \right) \Psi = 0 \quad (2.33)$$

Cas de l'oscillateur non dégénéré :

Considérons les séparations de variables suivantes :

$$\Psi(Q_1, Q_2, \dots, Q_{3n-6}) = \Psi_1(Q_1)\Psi_2(Q_2)\dots\Psi_{3n-6}(Q_{3n-6}) \quad (2.34)$$

$$E = E_1 + E_2 + \dots + E_{3n-6} \quad (2.35)$$

L'équation 2.33 revient donc à résoudre  $(3n - 6)$  équations à une seule variable :

$$\frac{d^2\Psi_i(Q_i)}{dQ_i^2} + \frac{2}{\hbar^2} \left( E - \frac{\lambda_i}{2} Q_i^2 \right) \Psi_i(Q_i) = 0 \quad (2.36)$$

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 2.5.1) via l'application de la relation :

$$Q_i = \left( \frac{\hbar^2}{\lambda_i} \right)^{1/4} q_i \quad (2.37)$$

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

$$\frac{d^2\Psi_i(q_i)}{dq_i^2} + \left( \frac{2E_i}{\hbar\lambda_i^{1/2}} - q_i^2 \right) \Psi_i(q_i) = 0 \quad (2.38)$$

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

$$\Psi_i(q_i) = N_{v_i} H_{v_i}(q_i) e^{-\frac{q_i^2}{2}} \quad (2.39)$$

$$E_i = \int \Psi_i H_i \Psi_i dq_i = hc\varpi_i \left( v_i + \frac{1}{2} \right) \quad (2.40)$$

avec  $v_i$  : nombre quantique vibrationnel de la coordonnée  $q_i$ , entier positif,  
 $N_{v_i}$  : facteur de normation des fonctions d'état vibrationnelles (2.41),  
 $H_{v_i}(q_i)$  : un polynôme d'HERMITE (2.42).

$$N_{v_i} = \left( 2^{v_i} v_i! \sqrt{\pi} \right)^{-1/2} \quad (2.41)$$

$$H_{v_i}(q_i) = (-1)^{v_i} e^{q_i^2} \frac{d^{v_i}}{dq_i^{v_i}} \left( e^{-q_i^2} \right) \quad (2.42)$$

Notons de plus que d'après la partie 2.5.1, 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 appellerons 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.

## 2.7 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 fait, l'expression du potentiel développée dans la partie 2.4 n'est plus suffisante, et les termes d'ordres supérieurs à 2 de la fonction potentielle doivent être pris en compte pour modéliser plus exactement 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 :

$$\begin{aligned} V_v = V_{eq} + \sum_i^{3n-6} \left( \frac{\partial V}{\partial s_i} \right)_{eq} s_i + \frac{1}{2} \sum_{i \leq j}^{3n-6} \left( \frac{\partial^2 V}{\partial s_i \partial s_j} \right)_{eq} s_i s_j \\ + \frac{1}{3} \sum_{i \leq j \leq k}^{3n-6} \left( \frac{\partial^3 V}{\partial s_i \partial s_j \partial s_k} \right)_{eq} s_i s_j s_k \\ + \frac{1}{4} \sum_{i \leq j \leq k \leq l}^{3n-6} \left( \frac{\partial^4 V}{\partial s_i \partial s_j \partial s_k \partial s_l} \right)_{eq} s_i s_j s_k s_l \\ + \dots \end{aligned} \quad (2.43)$$

Les dérivées d'ordre 2, 3 et 4 sont appelées respectivement constantes de force quadratique, cubique et quartique. La troncature à l'ordre 4 de l'expression du potentiel est, selon MASLEN [45], 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 secondes 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 :

$$\frac{V_v}{hc} = \frac{1}{2!} \sum_i \varpi_i q_i^2 + \frac{1}{3!} \sum_{i,j,k} \phi_{ijk} q_i q_j q_k + \frac{1}{4!} \sum_{i,j,k,l} \phi_{ijkl} q_i q_j q_k q_l \quad (2.44)$$

où  $\phi_{ijk}$  et  $\phi_{ijkl}$  sont, respectivement, les constantes de force cubique et quartique, exprimées en  $\text{cm}^{-1}$ . Les relations entre les dérivées d'ordre 3 et 4 de l'équation 2.43 et les  $\phi$ , qui s'obtiennent par les termes de la matrice de passage  $[L]$ , sont détaillées dans la référence [46]. Lorsqu'il est nécessaire d'expliciter la dégénérescence des coordonnées, nous utiliserons la notation de Nielsen [47] dans la base des modes normaux sans dimension :

$$V_{pot} = \sum_i^{3n-6} \frac{\omega_i}{2} q_i^2 + \sum_{||S||_1=3}^S K_S \prod_{i=1}^{3n-6} q_i^{S_i} \quad (2.45)$$

avec  $\omega_i$  la fréquence harmonique (en  $\text{cm}^{-1}$ ) associée 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.

Les valeurs numériques des constantes de force peuvent être déterminées de plusieurs façons. 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

---

<sup>2</sup>Dans le cas d'oscillateurs fortement anharmoniques, des coordonnées de type SIMONS-PARR-FINLAN [43] ou des coordonnées de type MORSE [44] peuvent être aussi utilisées.

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 la grille de points ainsi obtenue. L'expression de cette fonction est déterminée par des procédés de régression linéaire. C'est cette approche qui est le plus couramment employée, car elle conduit à des résultats de précision satisfaisante à 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, contraignant l'utilisateur à “dégrader” 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 : ces approches, proposées par Peter Pulay [48], représentent un compromis entre les deux procédés de dérivation 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'atomes 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 calculs nécessaires pour traiter 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 à l'étude dans le présent travail (présentant toutes plus de 15 atomes), il nous est apparu très rapidement nécessaire de trouver un moyen pour : *(i)* ne pas avoir à calculer l'ensemble des termes de la PES jusqu'à l'ordre 4 ou *(ii)* 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. La première consiste à supprimer de la PES les monômes d'ordre 3 et 4 inférieurs à une constante de force pré-déterminée par l'utilisateur. Cette stratégie “arbitraire” a été employée avec succès dans le travail relatif à la famille des acènes [49] (chapitre Identification of network modes). La seconde piste, développée très récemment dans le cadre d'un stage de Master [50], 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 sont intimement liés 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, quant 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.

## 2.8 Représentation matricielle du Hamiltonien, calcul d'intégrales

Le premier terme de l'équation (2.45) 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 :

$$\psi_{j,v}^{(0)} = \prod_{k=1}^{nv} \Theta_{j,v}^{(0)}(q_k) \text{ avec } \langle \Theta_{j,v}^{(0)} | \Theta_{j,v}^{(0)} \rangle = \delta_{v_kv_{k'}} \quad (2.46)$$

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

$$|v\rangle_i = |v_1, v_2, \dots, v_{nv}\rangle_i = \sum_j C_j \psi_{j,v}^{(0)} \quad (2.47)$$

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 :

$$\hat{H}_{v,l} |v, l\rangle_i = E_{v,l} |v, l\rangle_i \quad (2.48)$$

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

Pour résoudre des systèmes d'équations tels que l'équation (2.48), 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. Les éléments matriciels  $H_{(v,l)(v',l')}$  sont définis par la relation :

$$H_{(v,l)(v',l')} = \int \Psi_{(v,l)}^{(0)} \hat{H} \Psi_{(v',l')}^{(0)} dq_{(i,j,\dots,nv)} \quad (2.49)$$

$$= \langle (v, l)^{(0)} | \hat{H} | (v', l')^{(0)} \rangle \quad (2.50)$$

dans laquelle  $\hat{H}$  est partitionné de la manière suivante :

$$\hat{H} = \hat{H}^0 + \hat{H}^1 + \hat{H}^2 \quad (2.51)$$

où  $\hat{H}^0$  se rapporte au Hamiltonien harmonique,  $\hat{H}^1$  et  $\hat{H}^2$  traduisent l'anharmonicité et font intervenir respectivement les opérateurs  $q^3$  et  $q^4$ , ces derniers pouvant être décomposés en un

produit d'opérateurs  $q$  [51] :  $q^3 = \prod_{i,j,k}^3 q_i q_j q_k$ ,  $q^4 = \prod_{i,j,k,l}^4 q_i q_j q_k q_l$ . Les variables étant séparables, chaque terme matriciel se décompose en un 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, notés ici  $\mathbf{H}$ , l'intégrale s'écrit :

$$\langle v | q | v' \rangle = N_v N_{v'} \int e^{-q^2} \mathbf{H}_v q \mathbf{H}_{v'} dq \quad (2.52)$$

On peut montrer qu'il existe des relations de récurrence entre les polynômes, ce qui conduit à [37] :

$$\langle v | q | v' \rangle = N_v N_{v'} \int e^{-q^2} \mathbf{H}_{v'} \left( v \mathbf{H}_{v-1} + \frac{1}{2} \mathbf{H}_{v+1} \right) dq \quad (2.53)$$

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

$$\langle v | q | v+1 \rangle = \langle v+1 | q | v \rangle = \sqrt{\frac{v+1}{2}} \quad (2.54)$$

$$\langle v | q | v-1 \rangle = \langle v-1 | q | v \rangle = \sqrt{\frac{v}{2}} \quad (2.55)$$

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 [52].

## 2.9 Méthodes perturbationnelles

Il existe deux méthodes de calcul perturbationnel, à savoir la méthode dite de “transformation de contact”, issue des travaux de Van Vleck [53, 54], et la méthode standard dite de Rayleigh-Schrödinger [55]. Cette dernière est relativement simple à mettre en place et conduit aux expressions bien connues de l'énergie :

$$E_{(v,l)}^{Tot} = E_{(v,l)}^{(0)} + H_{(v,l),(v,l)}^{(2)} + \sum_{v \neq v'} \frac{(H_{(v,l),(v',l')}^{(1)})^2}{E_{(v,l)}^{(0)} - E_{(v',l')}^{(0)}} \quad (2.56)$$

$E_v^{(0)}$  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 :

$$|v, l\rangle^{(1)} = |v, l\rangle^{(0)} + \sum_{v \neq v'} \frac{H_{(v,l),(v',l')}}{E_{(v,l)}^{(0)} - E_{(v',l')}^{(0)}} * |v', l'\rangle^{(0)} \quad (2.57)$$

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

$$\begin{aligned} E_{(v,l)} &= \sum_s \omega_s \left( v_s + \frac{1}{2} \right) + \sum_t \omega_t (v_t + 1) + \sum_{s \geq s'} x_{ss'} \left( v_s + \frac{1}{2} \right) \left( v_{s'} + \frac{1}{2} \right) \\ &\quad + \sum_{s,t} x_{st} \left( v_s + \frac{1}{2} \right) (v_t + 1) + \sum_{t \geq t'} (v_t + 1) (v_{t'} + 1) + \sum_{t \geq t'} g_{tt'} l_t l_{t'} + \dots \end{aligned} \quad (2.58)$$

soit encore<sup>3</sup> :

$$\begin{aligned} E_{(v,l)} &= \sum_s \omega_s \langle v_s | q_s^2 | v_s \rangle + \sum_t \omega_t \sum_l \langle v_t, l | q_t^2 | v_t, l \rangle \\ &\quad + \sum_{s \neq s'} k_{sss's'} \langle v_s | q_s^2 | v_s \rangle \langle v_{s'} | q_{s'}^2 | v_{s'} \rangle + \sum_{s=s'} k_{ssss} \langle v_s | q_s^4 | v_s \rangle \\ &\quad + \sum_{s,t} k_{sstt} \langle v_s | q_s^2 | v_s \rangle \left( \sum_l \langle v_t, l | q_{t+} q_{t-} | v_t, l \rangle \right) \\ &\quad + \sum_{t \neq t'} k_{ttt't'} \left( \sum_l \langle v_t, l | q_{t+} q_{t-} | v_t, l \rangle \right) \left( \sum_{l'} \langle v_{t'}, l' | q_{t'+} q_{t'-} | v_{t'}, l' \rangle \right) \\ &\quad + \sum_{t=t'} k_{tttt} \left( \sum_l \langle v_t, l | q_{t+}^2 q_{t-}^2 | v_t, l \rangle \right) \\ &\quad - \frac{1}{2} \sum_{t=t'} k_{tttt} l_t^2 + \Theta(q^3) \end{aligned} \quad (2.59)$$

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 cubique et quartique. 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 [56], il n'en demeure pas moins que son utilisation reste limitée puisqu'elle ne s'adresse, en toute rigueur, qu'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 [57] ou bien encore Darling-Dennison [58]) et au fait que, en 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 combinaison.

Des solutions ont 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 [59], la théorie de perturbation canonique de grand ordre de Van Vleck [47], ou encore l'approximation du champ vibrationnel auto-cohérent (VSCF, pour "Vibrational Self-Consistent Field") qui inclus 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" [60, 61].

## 2.10 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

---

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

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  $H_v$ , de sorte que :

- $|v, l\rangle$  soit normé :

$$\langle v, l | v, l \rangle = \sum_j^N C_j^2 \quad (2.60)$$

- l'énergie vibrationnelle :

$$E_{v,l} = \langle v, l | H_{v,l} | v, l \rangle = \sum_j^N C_j^2 \langle \psi_{j,v}^{(0)} | H_{v,l} | \psi_{j,v}^{(0)} \rangle + \sum_j^N \sum_{k \neq j}^N C_j C_k \langle \psi_{j,v}^{(0)} | H_{v,l} | \psi_{j,v}^{(0)} \rangle \quad (2.61)$$

$$E_{v,l} = \sum_j^N C_j^2 H_{jj} + 2 \sum_j^N \sum_{k \neq j}^N C_j C_k H_{jk} \quad (2.62)$$

soit minimale compte-tenu de la condition de normalisation.

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

$$dE_{v,l} = \sum_{j=1}^N \frac{\partial E_{v,l}}{\partial C_j} dC_j = 0 \quad (2.63)$$

où

$$\frac{\partial E_{v,l}}{\partial C_j} dC_j = 2C_j H_{jj} + 2 \sum_{k \neq j}^N C_k H_{jk} \quad (2.64)$$

Les variations ( $\partial C_j$ ) 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 ( $H$ ) :

$$(H)(C) = E(C) \quad (2.65)$$

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 [62, 63, 64, 65, 66, 67] 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 rapidement avec la taille de la molécule, i.e. proportionnellement au nombre de vibrateurs et au terme noté  $g$  qui représente le nombre de termes non nuls de la PES. Il est possible de montrer que cette croissance se comporte en  $O(gN)$  [68]. 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 ; toutefois, cette dernière étant forcément finie et limitée, le problème réside donc dans le choix des configurations  $N_s$  à diagonaliser. Ce point est crucial car de ce choix dépendent grandement la précision et la rapidité de convergence de la méthode variationnelle. La troncature faite sur l'espace  $N$  est ainsi devenue un axe de recherche majeur, qui a conduit à différentes méthodes possédant chacune ses 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 [69], nous avons, dans notre méthode, initialement commencé par implémenter une méthode dite P\_MWCI (Parallel Multiple Windows CI) basée sur la construction itérative et parallélisée ( $\gamma$  cycles) de  $\langle H_v \rangle$  de l'espace de configuration. Les détails sont reportés dans la référence [70]. 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 pour nos analyses (sans jamais avoir à diagonaliser la matrice globale du problème variationnel). Notre analyse est basée sur une méthode hiérarchique apparentée à la méthode de Rayleigh-Ritz variationnelle utilisant une nouvelle façon d'écrire l'erreur résiduelle commise entre deux itérations [71]. 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 raisonnée. L'utilisation *a posteriori* du résidu nous permet de contrôler la/les direction(s) la (les) plus pertinente(s), dans laquelle (lesquelles) 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 [72, 73, 74, 75], le développement du logiciel MULTIMODE [76] et à la très efficace approche VCC ("Vibrational Coupled Cluster") récemment développée par Christiansen [77]. 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 [78]. 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<sup>-1</sup>, 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.

## 2.11 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 [79], 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, et l'espace des fonctions générées à partir d'un sous espace initial  $S_0$  verra l'information qu'il contient traitée de manière plus approximative par une méthode de perturbation de type 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 tient 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 permettant 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  $S_0$  détermine donc 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  $H$  ( $H = H^{(0)} + H'$ )

$$H^{(0)}\psi_i^{(0)} = E_i^{(0)}\psi_i^{(0)} \quad (2.66)$$

avec

$$\psi_i^{(0)} = \prod_{k=1}^{nv} \Theta_{v_k}^{(0)}(q_k) \text{ avec } v_k = 0, 1, 2, \dots \quad (2.67)$$

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

$$\langle \psi_i^{(0)} | H' | \psi_k^{(0)} \rangle \text{ avec } \psi_k^{(0)} \notin S_0 \quad (2.68)$$

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'onde d'ordre 1, l'énergie d'ordre 0 et l'énergie corrigée à la  $(n)^{\text{ième}}$  itération. La diagonalisation du Hamiltonien dans le sous-espace  $S_m^{(n)}$  nous permet d'obtenir la fonction d'onde  $\psi_{0m}^{(n)}$  associée à l'état étudié, de valeurs propres associées  $E_{0m}^{(n)}$  :

$$\psi_{0m}^{(n)} = \sum_{i \in S_m^{(n)}} C_i \psi_i^{(0)} \quad (2.69)$$

$$\langle \psi_{0m}^{(n)} | H | \psi_{0m}^{(n)} \rangle \quad (2.70)$$

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

$$E_m^{(n)} = E_{0m}^{(n)} + \epsilon_m^{(n)} \quad (2.71)$$

$$\epsilon_m^{(n)} = \sum_{j \notin S_m^{(n)}} \frac{|\langle \psi_i^{(0)} | H' | \psi_j^{(0)} \rangle|^2}{E_{(i)}^{(0)} - E_{(j)}^{(0)}} \quad (2.72)$$

$\epsilon_m^{(n)}$  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 :

$$|\psi_{1m}^{(n)}\rangle = |\psi_{0m}^{(n)}\rangle + \sum_{j \notin S_m^{(n)}} \frac{|\langle \psi_i^{(0)} | H' | \psi_j^{(0)} \rangle|}{E_i^{(0)} - E_j^{(0)}} |\psi_j^{(0)}\rangle = |\psi_{0m}^{(n)}\rangle + \sum_{j \notin S_m^{(n)}} C_{ij}^{(n)} |\psi_j^{(0)}\rangle \quad (2.73)$$

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

$$|C_{ij}^{(n)}| > \eta_m^{(n)} \quad (2.74)$$

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 de recommencer un cycle, jusqu'à obtenir un résultat stable. Précisons toutefois que cette notion de "résultat stable" reste délicate à définir, comme l'on prouvé nos récents développements concernant la notion de résidu [71]. Néanmoins, de façon générale, la fonction d'onde vibrationnelle s'améliore au cours des itérations, cette amélioration étant obtenue 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 second ordre de chacune des fonctions qui le composent. Les fonctions du sous-espace génératrice doivent interagir faiblement avec celles de l'espace complémentaire, de sorte que la fonction d'onde  $\psi_{1m}^{(n)}$  soit normalisée, ce qui correspond à :

$$\sum_{j \notin S_m^{(n)}} |C_{ij}^{(n)}|^2 \cong 0 \quad (2.75)$$

# Chapter 3

Systèmes en interaction à longue portée - De la nécessité des approches DFT

## Introduction

Un des challenges qui nous ont été proposés initialement par les expérimentateurs était celui de la caractérisation des molécules et de leurs interactions au sein des mélanges d'asphaltènes. Comme nous l'avons rappelé dans le chapitre 1, très peu d'informations sont connues à ce sujet à l'heure actuelle. Aucune technique de caractérisation, quelle soit physique ou chimique, n'est véritablement à même de répondre précisément à cette question. J. Shaw a proposé, bien avant les autres acteurs du domaine, que ces systèmes " associés" soient caractérisés à l'aide des spectroscopies Infra-Rouge et Raman, en utilisant notamment la technique de la photo-acoustique, dans le but d'accéder aux informations spectrales de très bas nombres d'ondes (zone spectrale suspectée de contenir la signature des modes de libration associés aux interactions intermoléculaires). Au-delà des difficultés inhérentes au manque d'informations, aussi bien quant à la nature des molécules et des familles qui constituent ces mélanges complexes que concernant le rôle des hétéroéléments, une autre des difficultés qui s'est imposée à nous concernait le traitement quantique vibrationnel de systèmes de grande dimension et, qui plus est, en interaction. En admettant même que les deux problèmes vibrationnels majeurs, i.e. la détermination de la SEP (Surface d'Energie Potentiel) et la résolution de l'équation vibrationnelle de Schrödinger, soient absents et indépendants de la dimension du problème (ce qui n'a bien entendu aucun sens, comme nous l'avons montré dans le chapitre 2), la prise en compte dans nos modèles quantiques des interactions est à elle seule un des problèmes majeurs qu'il nous a fallu intégrer dès le départ, sous peine de ne pouvoir fournir aux expérimentateurs des données suffisamment fiables pour rendre compte des observations spectroscopiques.

Ainsi, pour ce qui concerne le traitement quantique des interactions intermoléculaires, qui fait l'objet principal de ce chapitre, il est important de rappeler que notre travail ne se situe pas dans un contexte de développement méthodologique mais plutôt dans un contexte de recherche des stratégies calculatoires les plus adaptées pour traiter de problèmes vibrationnels concernant des systèmes chimiques en interaction, de grandes dimensions, qu'une méthode fortement corrélée ne permet pas encore d'atteindre<sup>1</sup>. L'utilisation de la DFT s'est donc naturellement imposée à nous.

Force est de constater les progrès réalisés depuis une décennie dans la description du formalisme de la théorie de la fonctionnelle de la densité, qui rendent son utilisation quasi-systématique. Bien loin des effets de mode, on constate les réussites frappantes de la DFT dans la description de propriétés telles que les liaisons, les structures, la cohésion aussi bien au niveau moléculaire que dans les solides. Néanmoins, il est des domaines dans lesquels les progrès sont encore lents et où les méthodes sont encore mises en défaut, tels que : le traitement des atomes lourds, les systèmes possédant des électrons fortement corrélés, le traitement crucial en réactivité des états multi-déterminantaux, le problème de la correction de la self-itération et, bien entendu, le problème qui nous occupe directement de la description des forces de dispersion [80]. Cette dernière problématique n'est pas récente puisqu'elle émerge des premiers travaux de Gordon et Kim [81]. Elle s'avère particulièrement prégnante pour l'étude des systèmes fortement liés (tels certains oxydes métalliques) et dans les cas où les forces de van der Waals sont prédominantes (matière molle, complexes de van der Waals, biomolécules et autres matériaux

<sup>1</sup>malgré les progrès particulièrement significatifs réalisés au cours de ces dernières années, tels que l'emploi de bases cc-pVXZ-F12 (X=D,T,Q) développées pour l'utilisation de méthodes explicitement corrélées (citons notamment les bases dédiées à la résolution des approximations d'identité RI (en anglais « Resolution Identity »)), des approximations d'ajustement de la densité DF (en anglais « Density Fitting ») pour le calcul d'intégrales, la non-linéarité des facteurs de corrélation, etc.

lamellaires). L'origine de ce problème est désormais clairement identifiée : les effets de corrélation électronique des forces de dispersion sont purement non locaux et en aucun cas une approximation locale ou semi-locale ne pourra en rendre compte. Si la fonctionnelle d'échange-corrélation exacte pouvait être connue, il serait cependant possible de répondre à la question posée par la description des forces de van der Waals dans le formalisme KS. Cela fait plus d'une décennie que la recherche dans ce domaine est extrêmement active. L'idée qui prévaut aujourd'hui est celle de développer des méthodes ayant les moyens, pour un moindre coût calculatoire, d'incorporer des corrections au formalisme KS sans perdre de vue les avantages de la DFT, aussi bien au niveau moléculaire que dans le calcul en conditions périodiques.

Rappelons que l'utilisation pratique de la DFT dans son formalisme KS repose sur l'approximation de la fonctionnelle d'échange-corrélation. L'approximation originelle, i.e. l'Approximation de Densité Locale (LDA pour l'anglais “Local Density Approximation”), s'est révélée étonnamment efficace et difficile à améliorer de manière systématique. Il a été montré que la LDA, étant locale par construction, est particulièrement adaptée pour décrire les corrélations à très courte portée inter-électronique, mais échoue à décrire quantitativement les corrélations à longue portée électronique. Ceci reste vrai avec la plupart des améliorations post-LDA classiques (que nous n'aborderons pas dans ce travail), qui demeurent par essence des approximations de nature locale. Différentes solutions existent désormais pour traiter ce problème. Citons par exemple sans aucune volonté d'exhaustivité, avec différents niveaux d'empirisme, les travaux autour de paramétrisations spécifiques de nouvelles fonctionnelles GGA ou méta-GGA [82], les travaux de type DFT+D incluant des corrections de dispersion de Grimme pour ne citer que le plus connu [83, 84], les travaux de type Symmetry-Adapted Perturbation Theory (SAPT) ou SAPT(DFT) même si ces derniers ne s'attachent plus tout à fait à la recherche d'une fonctionnelle d'échange-corrélation, et les travaux qui vont finalement s'atteler à donner une forme explicite de la fonctionnelle d'échange-corrélation [85, 86], en recherchant son caractère non-local (notamment au travers de l'application du théorème de fluctuation-dissipation combiné à la connexion adiabatique, comme dans les méthodes AC-FDT, DRSLL [87, 88] et Bayesian error estimation functional BEEF [88]. Parmi ces dernières solutions non-empiriques, on peut citer notamment les modèles d'Andersson-Langreth-Lundqvist (ALL) ou ceux de type Random Phase Approximation (RPA) qui ont très récemment fait l'objet d'études approfondies de la part, par exemple, de S. Lebègue et coll [89, 90].

Pour résumer, on peut rappeler que, d'un point de vue théorique, l'énergie d'interaction intermoléculaire est une observable que l'on peut soit calculer directement (à l'aide d'une approche globale de type “supermolécule”), soit chercher à interpréter par différentes décompositions auxquelles on peut adjoindre, dans la mesure du possible, un sens physique. Nous avons fait le choix arbitraire, dans la première partie de ce chapitre, de présenter les termes d'interaction à partir de la théorie de perturbation développée jusqu'au second ordre. Cette présentation nous semblait justifiée par le fait qu'une fois les hypothèses de travail définies et les fonctions de réponses – nécessaires au calcul des termes d'interaction – établies, le détail des approches méthodologiques les plus avancées dans ce domaine s'en trouve facilité. De plus, parce qu'il est désormais acquis qu'un tel traitement conduit généralement à une mauvaise convergence de la série de perturbation, nous profiterons de cette présentation pour introduire l'approche SAPT (“Symmetry-Adapted Perturbation Theory”), qui corrige ces divergences et qui est considérée aujourd'hui comme l'une des plus précises pour le calcul des énergies d'interaction intermoléculaires. De façon formelle, toutes les approches présentées

dans cette première partie n'imposent pas l'utilisation des approches de la DFT. Néanmoins, de façon pratique, de par la dimension et la diversité des systèmes que l'on cherche à étudier, le recours aux approches DFT est indispensable. Ainsi, après avoir donné, dans la seconde partie de ce chapitre, quelques rappels sur les fondements théorique de la méthodologie DFT, nous présenterons la notion de fonctionnelle hybride puis présenterons le principe des approches employées dans ce travail, qui ajoutent *ad hoc* à un calcul KS usuel des fonctions semi-empiriques rendant compte des effets de dispersion.

### 3.1 Théorie des perturbations

Au-delà des liaisons hydrogène et des interactions de nature électrostatique (*e.g.* charge-charge, charge-dipôle, dipôle-dipôle), l'interaction non covalente parmi les plus étudiées à l'heure actuelle est probablement celle impliquant les systèmes aromatiques [91]. Ces interactions entre cycles aromatiques, dénommées “ $\pi$ -stacking” (peut-être à mauvais escient [92]), sont notamment considérées comme responsables de la stabilité d'un très grand nombre de structures remarquables [93], dont celles qui préfigurent certainement la structure des asphaltènes. L'étude expérimentale de ces interactions représente encore un défi car il est généralement complexe de séparer les interactions de type  $\pi$ -stacking des interactions secondaires ou des effets de l'environnement. Du fait de ces difficultés expérimentales, les études en chimie computationnelle apparaissent comme une alternative de choix pour comprendre aussi bien la nature fondamentale de ces interactions non covalentes, que leur influence sur les systèmes chimiques étudiés. Nous encourageons le lecteur à se reporter au travail de revue réalisé sur le sujet par Martinez et Iverson, dans lequel sont recensés les principales définitions et les principaux modèles employés pour décrire ces interactions particulières. Le “graal”, pour les expérimentateurs comme pour les théoriciens, serait d'arriver *in fine* à caractériser séparément la contribution de chacune de ces interactions non covalentes et de prévoir tout à la fois leur rôle et leur intensité.

À ce jour, nous savons que 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, et se situe donc dans 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.

D'un point de vue théorique l'énergie d'interaction intermoléculaire est une observable que l'on peut interpréter par différentes décompositions auxquelles on cherche à adjoindre, dans la mesure du possible, un sens physique. Buckingham [94] proposait déjà, en 1967, une décomposition de l'énergie d'interaction intermoléculaire en quatre grandes contributions : l'électrostatique  $E_{elec}$ , l'induction  $E_{ind}$ , l'échange-répulsion  $E_{rep}$ , et la dispersion  $E_{disp}$ .

$$\Delta E = E_{elec} + E_{ind} + E_{rep} + E_{disp} \quad (3.1)$$

L'interaction électrostatique est l'ensemble des interactions coulombiennes existant entre

deux densités de charges isolées. Elle 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 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. Cette énergie est non additive et toujours attractive.

Ces deux contributions sont très bien définies en physique classique, contrairement aux termes de dispersion et d'échange, qui sont liés à des effets quantiques.

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

Enfin, la dispersion n'a pas d'équivalent classique puisqu'il s'agit d'une interaction liée à la corrélation électronique de deux densités de charge en interaction (fluctuation quantique des distributions de charges). Elle est attractive et existe dans tous les complexes.

Il y a, de façon pratique, deux manières différentes de calculer une énergie d'interaction intermoléculaire. La première est la méthode dite « supermolécules », tandis que la seconde consiste en la construction de l'interaction intermoléculaire à partir de la connaissance des réponses des fonctions d'onde des monomères séparés (au sein du dimère) soumis à l'action de perturbations externes.

### 3.1.1 L'approche en « supermolécules »

Dans la méthode supermoléculaire, l'énergie d'interaction intermoléculaire est obtenue par différence entre l'énergie totale du dimère et la somme des énergies totales de chacun des monomères.

$$\Delta E = E_{AB} - E_A - E_B \quad (3.2)$$

où  $A$  et  $B$  représentent les monomères et  $AB$  le dimère. En procédant ainsi, une erreur subtile peut être induite autour de l'énergie d'interaction. Cette erreur est connue sous le nom de “Basis Set Superposition Error” (BSSE) [95]. En calculant les deux monomères dans leur base spécifique, puis le dimère dans l'ensemble des fonctions de base des monomères, on peut utiliser les orbitales virtuelles d'un monomère pour agrandir la base disponible pour la distribution de charge de l'autre monomère, et vice versa. Il en résulte une augmentation de la qualité de la base pour le dimère vis-à-vis des monomères, et par conséquent une surestimation de l'énergie d'interaction. Afin de corriger l'erreur de BSSE, une possibilité consiste à travailler dans une base complète ou saturée pour les monomères et le dimère. Toutefois, la méthode la plus couramment employée reste la méthode dite “counterpoise”, proposée par Boys et Bernardi [96], dans laquelle les énergies respectives des monomères sont calculées dans la base du dimère. Cette corrections implique donc le calcul, pour chaque distance intermoléculaire, de l'énergie totale du dimère et des monomères.

L'énergie sans correction donnée pour l'équation 3.2 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:

$$E_{BSSE}(A) = E_A^{AB}(A) - E_A^A(A) \quad (3.3)$$

$$E_{BSSE}(B) = E_B^{AB}(B) - E_B^B(B) \quad (3.4)$$

où l'exposant désigne la base utilisée, l'indice désigne la géométrie, et le symbole entre parenthèses est le système chimique considéré. L'énergie d'interaction corrigée est donc donnée par l'expression suivante :

$$E^{CP} = E_{A-B}^{A-B}(A-B) - E_A^{A-B}(A) - E_B^{A-B}(B) \quad (3.5)$$

Dans ces conditions, les électrons de chaque fragment du système bénéficient de la base d'orbitales des autres fragments. La correction de la BSSE est par définition la différence entre l'énergie d'interaction non corrigée et l'énergie corrigée :

$$\Delta E^{CP} = E - E^{CP} = (E_{A-B}^{A-B}(A-B) - E_A^A(A) - E_B^B(B)) - (E_{A-B}^{A-B}(A-B) - E_A^{A-B}(A) - E_B^{A-B}(B)) \quad (3.6)$$

$$\Delta E^{CP} = E_A^{A-B}(A) - E_B^{A-B}(B) - E_A^A(A) - E_B^B(B) \quad (3.7)$$

Les bases ayant une extension spatiale finie, cette erreur est d'autant moins importante que les molécules sont éloignées les une des autres, la BSSE étant nulle dans la limite des bases infinies. En pratique, le calcul de la BSSE doit donc être effectué pour chaque conformation d'énergie minimale déterminée pour chacune des distances intermoléculaires qui séparent les espèces A et B. Plus la base est petite, plus cette erreur est importante. À l'inverse, dans la limite d'une base complète, la BSSE s'annule. Les calculs HF et KS souffrent moins des effets de la BSSE [97], qui sont plus marqués dans les calculs post-HF (MPn, CCSD(T)), du fait en particulier du traitement de l'espace des virtuels, nécessaire pour ces approches.

Malheureusement, le calcul direct d'une énergie d'interaction par la méthode "supermolécules" est extrêmement coûteux et ne donne aucune information sur la nature des interactions mises en jeu.

### 3.1.2 Détermination théorique des coefficients du développement multipolaire de l'énergie d'interaction à grande distance : potentiel électrostatique

La première théorie proposée en mécanique quantique pour déterminer la nature des interactions intermoléculaires a été développée dès 1927 par Wang [98] puis reprise et très largement étendue à partir de 1930 par London *et al.* [99]. Nous avons fait le choix, dans les paragraphes qui vont suivre, de ne mener le recensement des termes d'interaction introduits par London qu'aux plus bas ordres de perturbation, ceux-ci contenant les termes porteurs du sens physique. Notamment, le second ordre de perturbation contient *a priori* les termes de distorsion

et de dispersion que l'on peut naturellement relier aux propriétés électroniques des fragments en interaction. Les choix du modèle et de la troncature sont eux aussi totalement arbitraires, puisqu'il existe aujourd'hui d'autres modèles tels ceux de Kitaura et Morokuma [100] ou de Stevens et Fink [101], pour ne citer qu'eux. Le modèle perturbatif est celui qui, selon nous, permet la meilleure compréhension de la physique qui se cache derrière les forces de dispersion. Il présente l'avantage de rendre compte des coefficients dits de van der Waals ainsi que des fonctions de réponse qui leur sont associées telles que les polarisabilités multipolaires, elles-mêmes reliées aux fonctions spectrales telles que les forces d'oscillateur et les énergies d'excitation [102]. Enfin, le choix de présenter les termes d'interaction à partir de la théorie de perturbation nous semblait justifié par le fait qu'une fois les hypothèses de travail définies et les fonctions de réponse établies, la présentation des approches de type SAPT (Symmetry-Adapted Perturbation Theory) et des méthodes basées sur la détermination des fonctions semi-empiriques rendant compte des effets de dispersion s'en trouve facilitée et justifiée.

Dans ce paragraphe consacré aux interactions à longues distances, nous allons donner les expressions générales du Hamiltonien d'interaction, et établir à l'aide de la méthode de perturbation les équations du problème.

Soient deux systèmes  $a$  et  $b$  qui, à l'état isolé, sont susceptibles de se trouver dans les états  $\phi_0^a$  ...  $\phi_p^a$  et  $\phi_0^b$  ...  $\phi_q^b$ , respectivement fonctions propres des Hamiltoniens  $H_0(a)$  et  $H_0(b)$ , auxquels correspondent les énergies  $E_0^a$  ...  $E_p^a$  et  $E_0^b$  ...  $E_q^b$ . Lorsque ces deux systèmes interagissent, on peut exprimer le Hamiltonien global  $H$  du système comme :

$$H = H_0(a) + H_0(b) + H' \quad (3.8)$$

L'association  $[a-b]$  de ces deux systèmes engendre au niveau de  $b$  un potentiel électrostatique  $\Phi_{Kb}$  avec lequel les charges de  $a$  interagissent, donnant naissance à une énergie d'interaction (et réciproquement au niveau de  $a$ ).

### 3.1.2.1 Hamiltonien à l'ordre zéro pour le système $[a - b]$

A distance suffisamment grande pour que les distributions de charges liées aux systèmes  $a$  et  $b$  ne se recouvrent pas (de telle sorte que l'on puisse les considérer comme séparées), le Hamiltonien à l'ordre zéro associé au système  $[a - b]$  est simplement la somme des Hamiltoniens propres du système  $a$  et du système  $b$ . Soit :

$$H_0 = H_0(a) + H_0(b) \quad (3.9)$$

avec :

$$H_0(a) = -\frac{1}{2} \sum_{k_a=1}^{n_a} \nabla^2 k_a - \sum_{k_a=1}^{n_a} \frac{Z_a}{r_{k_a}} + \sum_{k'_a > k_a} \frac{1}{r_{k_a, k'_a}} + H_{l.s}(a) \quad (3.10)$$

Où  $\nabla_{k_a}$ ,  $Z_a$ ,  $r_{k_a}$  et  $r_{k_a, k'_a}$  représentent respectivement l'opérateur énergie cinétique, la charge du noyau de  $a$ , les distances noyau-électron  $k_a$  et électron  $k_a$ -électron  $k'_a$ . L'expression de  $H_0(b)$  se déduit facilement de celle de  $H_0(a)$  en remplaçant les coordonnées des  $n_a$  électrons du système  $a$  par les coordonnées des  $n_b$  électrons du système  $b$ .  $H_{l.s}(a) = \sum_{k_a=1}^{n_a} \zeta_{k_a} \widehat{l}_{k_a} \widehat{S}_{k_a}$  est le hamiltonien de spin-orbite de l'atome  $a$ , avec  $\widehat{l}_{k_a}$  et  $\widehat{S}_{k_a}$  les opérateurs moment orbital et moment de spin de l'électron  $k_a$ , et  $\zeta_{k_a}$  une constante de couplage spin-orbite pour le même électron.

### 3.1.2.2 Fonction d'onde moléculaire à l'ordre zéro

La structure électronique d'un système moléculaire et les propriétés qui en découlent peuvent être déterminées à partir de la résolution de l'équation de Schrödinger. Pour un système constitué de  $N$  électrons se déplaçant dans le champ électrostatique créé par les noyaux, cette équation s'écrit :  $H\Psi = E\Psi$ . La résolution de l'équation intégro-différentielle de Schrödinger aux états stationnaires pour les systèmes pluriélectroniques n'est pas envisageable sans approximations. Les fonctions propres du Hamiltonien  $H$  décrivent les états du système dont l'énergie est égale à la valeur propre correspondante. A partir de l'expression 3.9 du Hamiltonien non perturbé, on peut exprimer la fonction d'onde  $\Psi_{00}$  décrivant l'interaction entre deux atomes  $a$  et  $b$  comme :

$$\Psi_{00} = \phi_0^a \cdot \phi_0^b \quad (3.11)$$

qui n'est autre que le produit non antisymétrisé des fonctions d'onde  $\phi_0^a$  et  $\phi_0^b$  des systèmes  $a$  et  $b$  non perturbés. En toute rigueur, le simple produit des fonctions  $\phi_0^a$  et  $\phi_0^b$  n'est pas correct du point de vue des lois de la mécanique quantique puisque le principe d'indiscernabilité n'est pas vérifié. Néanmoins, pour de grandes distances interatomiques – pour lesquelles les électrons ne peuvent s'échanger d'un atome à l'autre – on peut raisonnablement négliger le recouvrement entre les fonctions d'onde atomiques (et par conséquent ne pas tenir compte des termes d'échange).

Le dernier terme de l'expression 3.10 –  $H_{l.s}(a) = \sum_{k_a=1}^{n_a} \zeta_{k_a} \widehat{l_{k_a}} \widehat{S_{k_a}}$  – traduit le couplage spin-orbite engendré par l'interaction entre les deux dipôles magnétiques, ou plus précisément par le spin et le mouvement de chacun des électrons sur son orbite. Ce terme de couplage est ainsi de nature relativiste. Dans ce travail, nous négligerons les interactions des spins avec les orbites des autres spins, ainsi que les interactions spin-spin, approximations d'autant plus justifiées que le terme répulsif est faible. Cette approximation est dite "non couplée" [103, 104] et les fonctions d'onde moléculaires correspondantes représentant le système  $[a - b]$  ont pour expression générale :

$$\Psi_{00}^\nu = \sum_{k=1}^{n_\nu} \alpha_{\nu k} |S_k L_k M_{S_k} M_{L_k}\rangle_a |S'_k L'_k M'_{S_k} M'_{L_k}\rangle_b \quad (3.12)$$

où  $S_k$ ,  $L_k$ ,  $M_{S_k}$  et  $M_{L_k}$  représentent respectivement les nombres quantiques de spin, orbitalaire et magnétique de l'atome  $a$ . Les termes  $S'_k$ ,  $L'_k$ ,  $M'_{S_k}$  et  $M'_{L_k}$  représentent de même les nombres quantiques associés au système  $b$ .

### 3.1.2.3 Perturbation

Deux atomes ou molécules à couches fermées engendrent toujours une énergie d'interaction due aux forces de van der Waals. Cette énergie d'interaction peut être calculée à l'aide d'une méthode de perturbation. En effet, aux grandes distances interatomiques  $R \gg R_{eq}$ , l'énergie d'interaction est proche de l'énergie associée aux deux molécules placées à l'infini l'une de l'autre. Tout le problème consiste donc à évaluer le déplacement de l'état fondamental du à l'introduction de  $H'$ , et en particulier sa dépendance à  $R$ . L'interaction coulombienne  $V$ , entre les électrons et le noyau de  $a$  et les électrons et le noyau de  $b$  sera traitée comme une perturbation, en considérant que  $H'$  reste faible devant le Hamiltonien du système global isolé ( $H_0(a) + H_0(b)$ ).

$$V = - \sum_{k_a=1}^{n_a} \frac{Z_b}{r_{bk_a}} - \sum_{k_b=1}^{n_b} \frac{Z_a}{r_{ak_b}} + \sum_{k_a=1}^{n_a} \sum_{k_b=1}^{n_b} \frac{1}{r_{k_ak_b}} + \frac{Z_a Z_b}{R} \quad (3.13)$$

### 3.1.2.4 Le Hamiltonien d'interaction

Soit un système dans lequel les charges sont localisées. Le Hamiltonien d'interaction  $H'$  s'exprime en fonction du potentiel électrostatique ( $\Phi_{K_b} = \sum_{k_a=1}^{N_a} \frac{e_{k_a}}{r_{k_a}}$ ) induit par les charges du système  $a$  sur les charges du système  $b$  (ou réciproquement par son équivalent  $\Phi_{K_a}$  relatif au système  $a$ ).

$$H' = \sum_{k_b=1}^{N_b} e_{k_b} \Phi_{k_b} = \sum_{k_a=1}^{N_a} \sum_{k_b=1}^{N_b} \frac{e_{k_a} e_{k_b}}{r_{k_ak_b}} \quad (3.14)$$

En développant le potentiel  $\Phi_{K_b}$  en série de Taylor par rapport au centre de la distribution de charges de  $b$  (noté  $o$ ) :

$$\Phi_{K_b} = \Phi_o + (\nabla_\alpha \Phi)_o r_{K_b\alpha} + \frac{1}{2} (\nabla_\alpha \nabla_\beta \Phi)_o r_{K_b\alpha} r_{K_b\beta} + \dots \quad (3.15)$$

On obtient pour expression de  $H'$  :

$$H' = \sum_{K_b=1}^{N_b} e_{K_b} \left( \Phi_o + (\nabla_\alpha \phi)_o r_{K_b\alpha} + \frac{1}{2} (\nabla_\alpha \nabla_\beta \phi)_o r_{K_b\alpha} r_{K_b\beta} + \dots \right) \quad (3.16)$$

où  $\alpha, \beta = (x, y \text{ ou } z)$  et où l'indice  $k_b$  se rapporte indifféremment au noyau et aux électrons du système  $b$  ( $N_b$ ).  $\nabla_\alpha$  correspond à une des trois composantes de l'opérateur gradient ( $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$  ou  $\frac{\partial}{\partial z}$ ) et  $r_{K_b\alpha}$  représente la composante rayon du vecteur décrivant la particule  $k_b$ . En tenant compte des définitions des moments multipolaires centrés sur l'origine du système :

$$\sum_{k_b=1}^{N_b} e_{k_b} = q^b$$

charge électrique de la molécule  $b$

$$\sum_{k_b=1}^{N_b} e_{k_b} r_{k_b\alpha} = \mu_\alpha^b$$

composante du moment dipolaire de  $b$  à l'origine

$$\frac{1}{2} \sum_{k_b=1}^{N_b} e_{k_b} (3r_{k_b\alpha} r_{k_b\beta} - r_{k_b}^2 \delta_{\alpha\beta}) = \theta_{\alpha\beta}^b$$

élément  $\alpha\beta$  du tenseur moment quadripolaire de  $b$

Ainsi que des définitions du champ :

$$F_\alpha^b = - \left( \frac{\partial \phi_{k_b}}{\partial r_\alpha} \right)_o = -(\nabla_\alpha \phi_{k_b})_o \quad (3.17)$$

et du gradient du champ électrique en  $o$  :

$$F_{\alpha\beta}^b = - \left( \frac{\partial^2 \phi_{k_b}}{\partial r_\alpha \partial r_\beta} \right)_o = -(\nabla_\alpha \nabla_\beta \phi_{k_b})_o \quad (3.18)$$

on retrouve l'expression classique du Hamiltonien d'interaction :

$$H' = \Phi_o q^b - \sum_\alpha F_\alpha^b \mu_\alpha^b - \frac{1}{3} \sum_{\alpha\beta} F_{\alpha\beta}^b \theta_{\alpha\beta}^b + \dots \quad (3.19)$$

En posant le vecteur  $\vec{R}_k = \vec{R} - \vec{r}_k$  qui représente la distance de la particule  $k$  à l'origine de la distribution des charges, le potentiel  $\Phi_o$  créé en  $o$  peut s'écrire soit sous sa forme classique 3.20:

$$\Phi_o = \frac{q^a}{R} + \sum_\alpha \mu_\alpha^a \frac{R_\alpha}{R^3} + \sum_{\alpha\beta} \theta_{\alpha\beta}^a \frac{R_\alpha R_\beta}{R^4} + \dots \quad (3.20)$$

soit sous une forme plus simplifiée :

$$\Phi_o = q^a T - \sum_\alpha \mu_\alpha^a T_\alpha + \sum_{\alpha\beta} \theta_{\alpha\beta}^a T_{\alpha\beta} + \dots \quad (3.21)$$

dans laquelle, les quantités  $T$ ,  $T_\alpha$  et  $T_{\alpha\beta}$  sont posées égales à :

$$\begin{cases} T = R^{-1} \\ T_\alpha = \nabla_\alpha R^{-1} = -\frac{R_\alpha}{R^3} \\ T_{\alpha\beta} = \nabla_\alpha \nabla_\beta R^{-1} = \frac{3R_\alpha R_\beta - R^2 \delta_{\alpha\beta}}{R^5} \end{cases} \quad (3.22)$$

$F_\alpha^b$  3.17 et  $F_{\alpha\beta}^b$  3.18 sont exprimés en fonction des moments multipolaires [105] précédemment définis :

$$F_\alpha^b = -(\nabla_\alpha \phi_B)_o = -q^a T_\alpha + \mu_\alpha^a T_{\alpha\beta} - \frac{1}{3} \theta_{\beta\gamma}^a T_{\alpha\beta\lambda} + \dots \quad (3.23)$$

$$F_{\alpha\beta}^b = -(\nabla_\alpha \nabla_\beta \phi_B)_o = q^a T_{\alpha\beta} + \mu_\delta^a T_{\alpha\beta\lambda} - \frac{1}{3} \theta_{\gamma\delta}^a T_{\alpha\beta\gamma\delta} + \dots \quad (3.24)$$

Le Hamiltonien d'interaction devient alors :

$$H' = q^a q^b T + T_\alpha (q^a \mu_\alpha^a) + T_{\alpha\beta} \left( \frac{1}{3} q^a \delta_{\alpha\beta}^b + \frac{1}{3} q^b \delta_{\alpha\beta}^a - \mu_\alpha^a \mu_\beta^b \right) + \dots \quad (3.25)$$

et  $H'$  apparaît comme la somme d'une infinité de termes :

- contribution de la charge totale du système :  $\frac{q^a q^b}{R}$ . Ce terme représente le potentiel en  $(1/R)$  que créent les charges des systèmes  $a$  et  $b$ .
- contribution du moment dipolaire électrique du système : interaction charge de  $a$ -moment dipolaire de  $b$  et charge de  $b$ -moment dipolaire de  $a$ . Au total, on obtient un terme variant en  $T_\alpha$  soit en  $(1/R^2)$ .
- un terme d'interaction dipôle-dipôle et deux termes charge de  $a$ -quadrupole de  $b$  et charge de  $b$ -quadrupole de  $a$ , variant en  $(1/R^3)$ ...

La première contribution au potentiel d'interaction des moments multipolaires est généralement due au terme d'interaction dipôle–dipôle variant en  $(1/R^3)$ . Aux grandes distances interatomiques, la contribution des termes au développement 3.25 décroît très vite et seuls les premiers termes suffisent à exprimer le potentiel d'interaction. Nous verrons dans les applications les conséquences que cela entraîne sur le développement de l'énergie d'interaction, dont l'expression analytique sera obtenue à l'aide des termes de van der Waals ( $C_3, C_5, C_6, \dots$ ).

### 3.1.2.5 Développement à un centre du potentiel électrostatique dû à une distribution de charges discrètes.

Le système de coordonnées cartésiennes est le système le plus utilisé et le plus naturel puisqu'il s'appuie sur la tridimensionnalité de notre espace quotidien. Néanmoins, son utilisation devient très vite complexe lorsque l'on cherche à traiter des problématiques d'interaction. Il existe ainsi un système de référence plus adapté à l'étude de ce type de propriétés, qui s'appuie sur le principe simple de proportionnalité entre les harmoniques sphériques. Ce référentiel est le système de coordonnées sphériques. Nous étudierons dans ce chapitre les développements à un et deux centres du potentiel électrostatique dû à une distribution de charges discrètes.

Soit  $\Phi_\rho$ , le potentiel électrostatique créé par une distribution de charges centrée en  $a$ , pour un point  $\rho$  extérieur à cette distribution :

$$\Phi_\rho = \sum_{k_a=1}^{n_a} \frac{e_{k_a}}{r_{k_a\rho}} \quad (3.26)$$

où  $e_{k_a}$  représente la charge de la  $k_a^{\text{ème}}$  particule et  $r_{k_a\rho}$  la distance qui sépare cette charge du point  $\rho$  considéré. Le développement de  $\Phi_\rho$  sera effectué dans le système de coordonnées sphériques, en utilisant la formulation de Laplace pour le développement de la quantité  $\frac{1}{r_{k_a\rho}}$ .

$$\frac{1}{r_{k_a\rho}} = \sum_{i=0}^{\infty} \sum_{m=-i}^{+i} \frac{(i - |m|)!}{(i + |m|)!} \frac{r^i}{r^{i+1}} < P_m^i(\cos \theta_{k_a}) P_m^i(\cos \theta_\rho) e^{im(\phi_{k_a} - \phi_\rho)} \quad (3.27)$$

où  $r <$  correspond à  $\min\{r_{k_a}, r_\rho\} \equiv r_{k_a}$  et  $r >$  correspond à  $\max\{r_{k_a}, r_\rho\} \equiv r$  et où les termes  $P_m^i(\cos \theta_\rho)$  représentent les fonctions associées au polynôme de Legendre de première espèce. Ces fonctions s'identifient aux harmoniques sphériques  $Y_i^m(\theta, \phi)$  à un facteur près :

$$P_m^i(\cos \theta) = (-1)^{m-|m|} \frac{\sqrt{4\pi}}{\sqrt{2i+1}} \frac{\sqrt{(i+|m|)!}}{\sqrt{(i-|m|)!}} e^{-im\phi} Y_i^m(\theta, \phi) \quad (3.28)$$

D'après les équations 3.27 et 3.28, le développement de Laplace pour l'inverse de la distance  $r_{k_a\rho}$  entre la charge  $q_a$  appartenant à la distribution  $a$  et le point  $\rho(r_\rho, \theta_\rho, \phi_\rho)$  qui lui est extérieur, s'exprime comme :

$$\frac{1}{r_{k_a\rho}} = \sum_{i=0}^{\infty} \sum_{m=-i}^{+i} (-1)^{|m|} \frac{4\pi}{2i+1} \frac{r_{k_a}^i}{r_\rho^{i+1}} Y_i^m(\theta_{k_a}, \phi_{k_a}) Y_i^{-m}(\theta_\rho, \phi_\rho) \quad (3.29)$$

En remplaçant l'expression 3.29 dans la formulation du potentiel électrostatique 3.26, l'expression de  $\Phi_\rho$  obtenue dépend de  $\theta$  et  $\phi$  :

$$\Phi_\rho = \sum_{i=0}^{\infty} \sum_{m=-i}^{+i} (-1)^{|m|} \frac{4\pi}{2i+1} \frac{1}{r_\rho^{i+1}} Y_i^{-m}(\theta_\rho, \phi_\rho) \sum_{k_a} e_{k_a} r_{k_a}^i Y_i^m(\theta_{k_a}, \phi_{k_a}) \quad (3.30)$$

Dans un soucis de simplification, nous poserons la variable  $Q_m^i(a)$  égale à :

$$Q_m^i(a) = \frac{\sqrt{4\pi}}{\sqrt{2i+1}} \sum_{k_a} e_{k_a} r_{k_a}^i Y_i^m(\theta_{k_a}, \phi_{k_a}) \quad (3.31)$$

d'où il vient que :

$$\Phi_\rho = \sum_{i=0}^{\infty} \sum_{m=-i}^{+i} (-1)^{|m|} \frac{4\pi}{2i+1} \frac{1}{r_\rho^{i+1}} Y_i^{-m}(\theta_\rho, \phi_\rho) Q_m^i(a) \quad (3.32)$$

Physiquement, les quantités  $Q_m^i(a)$  représentent les combinaisons linéaires des composantes des opérateurs moment dipolaire du système  $a$ .

- i=0, m=0       $Q_0^0(a) = \sum_{k_a} e_{k_a} = q^a$       charge de  $a$
- i=1, m=0       $Q_0^1(a) = \sum_{k_a} e_{k_a} r_{k_a} P_1^0(\cos \theta_{k_a}) = \mu_z^a$       composante z de l'opérateur
- i=1, m=1       $Q_1^1(a) = \sum_{k_a} \frac{1}{\sqrt{2}} e_{k_a} r_{k_a} P_1^1(\cos \theta_{k_a}) e^{i\phi_{k_a}}$       moment dipolaire de l'atome  $a$
- i=1, m=-1       $Q_{-1}^1(a) = - \sum_{k_a} \frac{1}{\sqrt{2}} e_{k_a} r_{k_a} P_{-1}^1(\cos \theta_{k_a}) e^{-i\phi_{k_a}}$

Si l'on considère que l'axe  $z$  est l'axe de plus haute symétrie, les formulations des composantes perpendiculaires  $x$  et  $y$  du moment dipolaire électrique du système s'expriment de la façon suivante :

$$\mu_x^a = \frac{1}{\sqrt{2}} (Q_1^1(a) - Q_{-1}^1(a)) \quad (3.33)$$

$$\mu_y^a = \frac{1}{\sqrt{2i}} (Q_1^1(a) + Q_{-1}^1(a)) \quad (3.34)$$

Les autres composantes du tenseur d'ordre deux seront obtenues en remplaçant  $m$  par ses valeurs permises ( $m = \pm 1, m = \pm 2$ ).

Les propriétés de ces opérateurs sont indépendantes du système étudié et satisfont à un ensemble de règles de sélection sur lesquelles nous ne reviendrons pas dans ce travail [106].

- i=2, m=0       $Q_0^2(a) = \theta_{zz}^a$       composante z de l'opérateur moment quadripolaire

### 3.1.2.6 Développement à deux centres de l'interaction coulombienne entre deux distributions de charges ne se recouvrant pas

De nombreux travaux [107, 108, 109, 110] ont montré que l'inverse de la distance  $r_{ij}$  existant entre une charge  $i$  et une charge notée  $j$  peut s'écrire sous forme d'un développement bipolaire dans le système de coordonnées de deux distributions (système de coordonnées bipolaires).

$$\frac{1}{r_{k_a k_b}} = \sum_{i,j=0}^{\infty} \sum_{l<}^{l>} B_{ij}^{|m|}(r_{k_a}, r_{k_b}; R) P_i^m(\cos \theta_{k_b}) P_j^m(\cos \theta_{k_a}) e^{im(\phi_{k_a} - \phi_{k_b})} \quad (3.35)$$

où  $l <= \inf(i, j)$

Dans l'hypothèse où les deux distributions de charges ne se recouvrent pas, on peut considérer que  $R > r_i + r_j$  et poser :

$$B_{ij}^{|m|}(r_{k_a}, r_{k_b}; R) = \frac{(1-)^{j+|m|}(i+j)!}{(i+|m|)!(j+|m|)!} r_{k_a}^i r_{k_b}^j \frac{1}{R^{i+j+1}} \quad (3.36)$$

Le potentiel d'interaction entre les deux distributions de charges s'écrit :

$$V = \sum_{k_a=1}^{N_a} \sum_{k_b=1}^{N_b} \frac{e_{k_a} e_{k_b}}{r_{k_a} r_{k_b}} \quad (3.37)$$

Le polynôme de Legendre associé s'exprime en fonction des harmoniques sphériques :

$$V = \sum_{i,j=0}^{\infty} \frac{1}{R^{i+j+1}} \sum_{-l<}^{l>} \frac{(-1)^j (i+j)!}{\sqrt{(i+m)!(i-m)!(j-m)!(j+m)!}} \quad (3.38)$$

$$\frac{\sqrt{4\pi}}{\sqrt{2i+1}} \sum_{k_a=1}^{N_a} e_{k_a} r_{k_a}^i Y_i^m(\theta_{k_a}, \phi_{k_a}) \frac{\sqrt{4\pi}}{\sqrt{2j+1}} \sum_{k_b=1}^{N_b} e_{k_b} r_{k_b}^j Y_j^{-m}(\theta_{k_b}, \phi_{k_b}) \quad (3.39)$$

où  $V$  se développe en fonction des opérateurs moments multipolaires de  $a$  et de  $b$ , définis par la relation 3.31.

En posant la quantité  $\frac{(-1)^j (i+j)!}{\sqrt{(i+m)!(i-m)!(j-m)!(j+m)!}} = d_m(i, j)$  on obtient pour expression simplifiée du potentiel d'interaction l'équation générale suivante :

$$V = \sum_{i,j=0}^{\infty} \frac{1}{R^{i+j+1}} \sum_{-l<}^{l>} d_m(i, j) Q_m^i(a) Q_{-m}^j(b) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{V_{ij}(a, b)}{R^{i+j+1}} \quad (3.40)$$

Les opérateurs multipolaires  $Q_m^{i,j}$  (charge, dipôle, quadripôle,...) contiennent toute l'information sur la distribution de charges du système considéré. Le potentiel  $V(a, b)$  contient donc chacune de ces informations pour chacun des deux systèmes  $a$  et  $b$  pris isolément. En prenant l'ensemble des valeurs de  $i$  et de  $j$  permises, les différentes interactions intervenant entre les multipôles (induits ou permanents) des deux systèmes sont :

- $i = j = 0$  :  $V_{00}$  représente l'interaction entre la charge de  $a$  et la charge de  $b$ . Ce terme est nul pour les systèmes neutres.
- $i = 0$  et  $j = 1$  :  $V_{01}$  représente l'interaction entre la charge de  $a$  et le dipôle de  $b$ .
- $i = j = 1$  :  $V_{11}$  représente l'interaction entre le dipôle de  $a$  et le dipôle de  $b$ .

$i$  et  $j$  prennent des valeurs théoriquement infinies, mais nous verrons que l'expression 3.40 peut être tronquée pour des valeurs de  $i$  et  $j$  faibles. De plus, suivant la nature des espèces en présence, toutes les combinaisons des paramètres  $i$  et  $j$  ne sont pas permises.

### 3.1.2.7 Détermination théorique des coefficients du développement multipolaire de l'énergie d'interaction à grande distance : énergie d'interaction

Cette partie est consacrée au recensement de l'ensemble des termes d'interaction obtenus à partir d'un calcul au premier et au second ordre de perturbation. Nous présenterons également les termes prépondérants d'ordre 6 et 8 que nous pouvons estimer [111].

### 3.1.2.8 Calcul de l'énergie électrostatique.

#### Expression des éléments de matrice de l'énergie électrostatique

$$E_1 = \langle \psi_{00}^\nu | V | \psi_{00}^\lambda \rangle \quad (3.41)$$

En remplaçant l'opérateur  $V_{ij}$  de perturbation par son expression 3.40, l'énergie électrostatique d'un système  $[a - b]$  s'exprime sous la forme d'une suite infinie :

$$E = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{E_{ij}(\nu, \lambda)}{R^{i+j+1}} = \sum_m \frac{C_m}{R^m} \quad (3.42)$$

dans laquelle les coefficients  $C_m$  représentent directement les éléments de matrice  $E_{ij}$  de l'énergie d'interaction électrostatique entre deux états  $\nu$  et  $\lambda$  :

$$E_{ij}(\nu, \lambda) = \langle \psi_{00}^\nu | V_{ij}(a, b) | \psi_{00}^\lambda \rangle = C_m \quad (3.43)$$

avec  $m = i + j + 1$

En reprenant les expressions des fonctions d'onde :

$$F_l^m(\theta, \phi) = \lambda(l) Y_l^m(\theta, \phi) \quad (3.44)$$

Les éléments de matrice  $E_{ij}$  s'expriment en fonction du terme d'interaction induit par les multipôles électrostatiques de  $a$  et de  $b$  :

$$E_{ij}(\nu, \lambda) = \sum_{k=1}^{n_\nu} \sum_{l=1}^{n_\lambda} \alpha_{\nu k} \alpha_{\lambda l}$$

$$\langle S_k L_k M_{S_k} M_{L_k} | \langle S'_k L'_k M'_{S_k} M'_{L_k} | V_{ij}(a, b) | S_l L_l M_{S_l} M_{L_l} \rangle | S'_l L'_l M'_{S_l} M'_{L_l} \rangle \quad (3.45)$$

avec :

$$V_{ij}(a, b) = \sum_{m=-l<}^{l>} d_m(i, j) Q_m^i(a) Q_{-m}^j(b) \quad (3.46)$$

Le terme  $V_{ij}$  apparaît comme le produit d'un opérateur tensoriel irréductible de rang  $i$ , agissant sur les seules coordonnées de  $a$ , par l'opérateur tensoriel irréductible de rang  $j$  n'agissant que sur les coordonnées de  $b$ . Dans le cadre de cette propriété, les éléments de matrice de l'énergie électrostatique se décomposent en deux parties : une partie relative au système  $a$  et l'autre relative au système  $b$ .

$$E_{ij}(\nu, \lambda) = \sum_{k=1}^{n_\nu} \sum_{l=1}^{n_\lambda} \alpha_{\nu k} \alpha_{\lambda l} \sum_{m=-l<}^{l>} d_m(i, j)$$

$$\langle S_k L_k M_{S_k} M_{L_k} | Q_m^i(a) | S_l L_l M_{S_l} M_{L_l} \rangle \langle S'_k L'_k M'_{S_k} M'_{L_k} | Q_{-m}^j(b) | S'_l L'_l M'_{S_l} M'_{L_l} \rangle \quad (3.47)$$

**Règles de sélection pour la détermination des coefficients de dispersion  $C_m$**  Les règles de sélection nous ont conduits aux relations :

$$\begin{cases} L_k + i + L_l = \text{entier pair} \\ |L_k - L_l| \leq i \leq L_k + L_l \\ L'_k + j + L'_l = \text{entier pair} \\ |L'_k - L'_l| \leq j \leq L'_k + L'_l \end{cases} \quad (3.48)$$

- Si  $L_k + L_l$  est pair,  $i$  est nécessairement pair (idem pour  $j$ ). Il s'ensuit que  $i + j$  est pair.
- Si  $L_k + L_l$  est impair,  $i$  est nécessairement impair (idem pour  $j$ ). Il en découle que  $i + j$  est pair.

Ainsi, quelle que soit les valeurs des moments orbitaux,  $i + j$  est toujours pair. L'énergie électrostatique :

$$E(l) = \sum_{n=0}^{\infty} \frac{C_{i+j+1}}{R^{i+j+1}} \quad (3.49)$$

sera donc toujours représentée par un développement en puissance impaire ( $i + j + 1$ ) de  $R$ .

- pour  $L_k = L_l = L'_k = L'_l = 0$ , nous aurons donc  $i = j = 0$  ; interaction charge-charge ( $Q_0^0 - Q_0^0$ )
- pour  $L_k = L_l = 0$ ,  $L'_k = L'_l = 1$ , les sommes  $i + 1$  et  $j + 1$  sont entières et paires. Soit  $i = j = 1$ , alors l'énergie électrostatique possède un terme de résonance non nul d'ordre trois :  $C_3/R^3$ .
- pour  $L_k = L_l = L'_k = L'_l = 1$ , les sommes  $i + 2$  et  $j + 2$  sont entières et paires. Ces règles impliquent que  $i = j = 2$ . L'énergie électrostatique possède un terme d'ordre cinq non nul :  $C_5/R^5$ .

### 3.1.2.9 Calcul de l'énergie de polarisation

L'énergie de polarisation représente l'énergie au second ordre des perturbations. Elle se compose de deux termes : un terme de dispersion impliquant simultanément les états excités de  $a$  et de  $b$ , et un terme d'induction ne faisant intervenir qu'un seul ensemble d'états excités issu de l'un des deux systèmes. En reprenant l'expression 3.40 de l'opérateur de perturbation, on obtient pour expression de l'énergie au second ordre :

$$E(2) = \sum_{i,j=0}^{\infty} \sum_{i',j'=0}^{\infty} \frac{E_{\text{disp}}^{ii'jj'}(\nu, \lambda) + E_{\text{induc}}^{ii'jj'}(\nu, \lambda)}{R^{i+i'+j+j'+2}} = \sum_{i+i'+j+j'+2=2}^{\infty} \frac{C_{i+i'+j+j'+2}}{R^{i+i'+j+j'+2}} \quad (3.50)$$

Le domaine de validité d'un tel développement est limité pour les très grandes distances par les effets de retardement. Pour les faibles distances où le recouvrement orbitalaire ne peut plus être négligé, l'équation 3.50 n'est plus valable et l'on doit prendre en compte l'énergie de répulsion. Pour des distances comprises entre ces deux domaines, où il n'existe généralement que très peu d'informations expérimentales, les méthodes dites asymptotiques peuvent apporter de précieux résultats.

**Energie de dispersion** L'énergie de dispersion fait intervenir simultanément les états excités de  $a$  et de  $b$ , que nous noterons respectivement  $\Psi_r(a)$  et  $\Psi_t(b)$ .

$$E_{disp}^{ii'jj'}(\nu, \lambda) = - \sum_{r0}^{\infty} \sum_{t0}^{\infty} \frac{\langle \Psi_0^\nu(a) \Psi_0^\nu(b) | V_{ij} | \Psi_r(a) \Psi_t(b) \rangle \langle \Psi_r(a) \Psi_t(b) | V_{i'j'} | \Psi_0^\lambda(a) \Psi_0^\lambda(b) \rangle}{(E_r(a) - E_0(a)) + (E_t(b) - E_0(b))} \quad (3.51)$$

En reprenant la forme générale des fonctions d'onde dépendantes des nombres quantiques orbitaux, magnétiques et de spin de chacun des systèmes pris isolément, l'équation 3.51 peut s'écrire :

$$\begin{aligned} E_{disp}^{ii'jj'}(\nu, \lambda) &= - \sum_{k=1}^{n_\nu} \sum_{l=1}^{n_\lambda} \alpha_{\nu,k} \alpha_{\lambda,l} \times \\ \sum_{w''_a w''_b} &\frac{\langle S_k L_k M_{S_K} M_{L_K}(a) | \langle S'_k L'_k M'_{S_K} M'_{L_K}(b) | V_{ij}(ab) | S''_a L''_a M''_{S_a} M''_{L_a}(a) \rangle | S''_b L''_b M''_{S_b} M''_{L_b}(b) \rangle}{(E_{n_a''} - E_0(a)) + (E_{n_b''} - E_0(b))} \times \\ &\frac{\langle S''_a L''_a M''_{S_a} M''_{L_a}(a) | \langle S''_b L''_b M''_{S_b} M''_{L_b}(b) | V_{i'j'}(a,b) | S_l L_l M_{S_l} M_{L_l}(a) \rangle | S'_l L'_l M'_{S_l} M'_{L_l}(b) \rangle}{(E_{n_a''} - E_0(a)) + (E_{n_b''} - E_0(b))} \end{aligned} \quad (3.52)$$

avec  $w''_a = (n''_a, l''_a, S''_a, L''_a, M''_{S_a}, M''_{L_a})$ . L'indice “” sera réservé aux états excités.

Comme pour l'énergie électrostatique, la démarche adoptée consiste à séparer les parties relatives au système  $a$  de celles relatives au système  $b$ .

Par application des règles de sélection sur les termes de Wigner qui composent la partie purement radiale, on obtient les valeurs de  $i, j, i'$  et  $j'$  permises :

$$\left\{ \begin{array}{l} l_k + i + l''_a = \text{entier pair} \\ |l_k - l''_a| \leq i \leq l_k + l''_a \\ l''_a + i' + l_l = \text{entier pair} \\ |l''_a - l_l| \leq i' \leq l''_a + l_l \\ l'_k + j + l''_b = \text{entier pair} \\ |l'_k - l''_b| \leq j \leq l'_k + l''_b \\ l'_l + j' + l''_b = \text{entier pair} \\ |l'_l - l''_b| \leq j' \leq l'_l + l''_b \end{array} \right. \quad (3.53)$$

Par exemple, pour  $l_k = l_l = l'_k = l'_l = 0$ ,  $i = i' = l''_a$  et  $j = j' = l''_b$ . L'ensemble des combinaisons  $(i, i', j, j')$  permises sont telles que  $i = i'$  et  $j = j'$ . En tenant compte de la définition de l'énergie de dispersion 3.50,

$$E_{disp}(2) = \sum_{i+i'+j+j'+2}^{\infty} \frac{C_{i+i'+j+j'+2}}{R^{i+i'+j+j'+2}} \quad (3.54)$$

il est possible de dénombrer et connaître les différentes contributions des coefficients de dispersion (encore appelés coefficients de van der Waals) à la correction énergétique au second ordre sur l'énergie du système  $[a - b]$ .

La combinaison (1,1,1,1) représente la contribution du coefficients  $C_{1+1+1+1+2} = C_6$ .

Les combinaisons (1,1,2,2) et (2,2,1,1) représentent les deux contributions au coefficient  $C_8$ , etc..

**Energie d'induction** Contrairement à l'énergie de dispersion, l'énergie d'induction ne fait intervenir que les états excités de l'un ou l'autre des systèmes  $a$  et  $b$  :

$$E_{ind}(2) = \sum_{i,j=0}^{\infty} \sum_{i',j'=0}^{\infty} \frac{E_{ind}^{ii'jj'}(a; \nu, \lambda) + E_{ind}^{ii'jj'}(b; \nu, \lambda)}{R^{i+i'+j+j'+2}} = \sum_{i+j+j'+i'+1=2}^{\infty} \frac{C_{i+i'+j+j'+2}}{R^{i+i'+j+j'+2}} \quad (3.55)$$

Soit par exemple l'énergie d'induction du système  $a$  :

$$E_{ind}^{ii'jj'}(a; \nu, \lambda) = \sum_{r \neq 0}^{\infty} \sum_{t \neq 0}^{\infty} \frac{\langle \Psi_0^{\nu}(a) \Psi_0^{\nu}(b) | V_{ij}(a, b) | \Psi_r(a) \Psi_0(b) \rangle \langle \Psi_r(a) \Psi_0(b) | V_{i'j'}(a, b) | \Psi_0^{\lambda}(a) \Psi_0^{\lambda}(b) \rangle}{(E_r(a) - E_0(a))} \quad (3.56)$$

avec  $\Psi_0^{\nu}(a) = |S_k L_k M_{S_k} M_{L_k}\rangle$  et  $\Psi_r(a) = |S''_a L''_a M''_a M''_{L_a}\rangle$ . On pourra trouver de même une formule similaire  $E_{ind}^{ii'jj'}(b; \nu, \lambda)$  pour le système  $b$ . Chaque terme est calculé séparément.

Les règles de sélection concernant le système  $a$  sont identiques à celles énoncées pour l'énergie de dispersion. Les règles de sélection propres au système  $b$  sont :

$$\begin{cases} L_b + j + L'_k = \text{entier pair} \\ |L'_k - L_b| \leq j \leq L'_k + L_b \\ L_b + j' + L'_l = \text{entier pair} \\ |L'_l - L_b| \leq j' \leq L'_l + L_b \end{cases} \quad (3.57)$$

- pour  $l_k = l_l = 0$  et  $L'_k = L'_l = 0$ , on a  $i = i' = l''_a$  et  $j = j' = 0$ . Les combinaisons  $(i, i', j, j')$  permises seront celles pour lesquelles  $i = i'$ ;  $(i, i, 0, 0)$ . La somme  $i + i'$  est paire. Les termes d'induction sont donc pairs d'ordre 4. Ces termes représentent les interactions dipôle de  $a$ -charge de  $b$ , quadrupôle de  $a$ -charge de  $b$ , etc., et sont toujours nuls dans le cas de systèmes neutres. Il est évident que le résultat demeure identique en considérant les états excités de  $b$  plutôt que ceux de  $a$ .
- pour  $l_k = l_l = 1$  et  $L'_k = L'_l = 0$ , les règles de sélection conduisent pour  $b$  aux mêmes conclusions que dans le cas précédent ( $j = j' = 0$ ). Ce terme est toujours nul dans le cas des atomes neutres.
- pour  $l_k = l_l = 0$  et  $L'_k = L'_l = 1$ , les règles de sélection conduisent pour  $b$  à  $j = 0, 2$  et  $j' = 0, 2$  et pour  $a$  à  $i = i'$ . Dans ce cas, les combinaisons permises sont  $(1, 1, 2, 2)$  pour le terme d'ordre 8,  $(2, 2, 2, 2)$  pour le terme d'ordre 10, etc. Notons que la valeur 0 est interdite pour les systèmes neutres. Le premier terme qui contribue à la valeur de l'énergie d'induction est donc le terme d'ordre 8 ( $C_8$ ) et cette énergie s'exprime comme :

$$E_{ind}^{ii'22}(a; \nu, \lambda) = \sum_{i=i'=1}^{\infty} \frac{C_{i+i'+6}}{R^{i+i'+6}} \quad \text{avec } j = 2 \quad (3.58)$$

...

### 3.1.3 Symmetry-Adapted Perturbation Theory (SAPT)

Comme nous l'avons déjà mentionné, les interactions de van der Waals ne se distinguent clairement des autres types d'interactions qu'à longue distance. Dans la région intermédiaire, qui inclus la région du minimum, il est même particulièrement difficile de décomposer de manière univoque l'énergie d'interaction totale en différentes contributions. Plusieurs schémas de décomposition ont néanmoins été proposés au fil du temps (comme celui présenté ci-avant). Le principe de ces dernières approches est d'évaluer par morceaux les différentes contributions, telles que l'électrostatique, l'induction ou encore la dispersion. Cependant, ces méthodes possèdent le défaut majeur de conduire à des processus non convergés (ou mal convergés) de la série perturbationnelle, tant du fait d'une troncature à des ordres souvent trop bas que du fait de leur non complétude. C'est pour palier ce dernier défaut que la théorie SAPT (Symmetry-Adapted Perturbation Theory) a été proposée. Dans cette approche, chacune des composantes précédentes inclus une correction d'échange afin de rendre compte de la nature anti-symétrique de la fonction d'onde du complexe formé. Cette théorie a par ailleurs été développée en partant du constat que le terme de polarisation se révèle en grande partie responsable de la divergence de la série perturbative, ce que vise à corriger l'introduction d'une correction d'échange.

L'approche SAPT est en quelque sorte une théorie de double perturbation : l'une concerne le potentiel intermoléculaire, tandis que l'autre traite du potentiel de corrélation intra-moléculaire. Afin de comprendre cette remarque, il est nécessaire de développer *a minima* les principes mathématiques sur lesquels repose la théorie SAPT.

Un moyen simple de présenter cette méthode est de considérer une partition classique du Hamiltonien d'un complexe moléculaire constitué de deux monomères A et B :

$$H = H^A + H^B + V \quad (3.59)$$

dans laquelle le terme  $V$  représente la perturbation intermoléculaire.

Le Hamiltonien total s'écrit encore :

$$H = H_0^A + W^A + H_0^B + W^B + V \quad (3.60)$$

dans lequel les termes  $H_0^{A/B}$  sont les opérateurs de Fock des monomères A et B,  $W^{A/B}$  représentent les termes de perturbation intramoléculaire des monomères et traduisent donc la corrélation "normale" de chaque fragment (*i.e.* des opérateurs de Möller-Plesset des monomères A et B), et où  $V$  représente toujours la perturbation intermoléculaire (*i.e.* l'interaction électrostatique entre A et B).

Le premier niveau de perturbation s'obtient en ne tenant pas compte, dans un premier temps, de la corrélation intramoléculaire, soit en considérant l'interaction classique intramoléculaire. Dans ce cas, et en introduisant un terme noté  $\zeta$  visant à moduler la perturbation (due à la partie intermoléculaire),  $H = H_0 + \zeta.V$ , l'équation de Schrödinger s'écrit :

$$(H_0 + \xi V)\psi(\xi) = (E_{int} + E_0)\psi(\xi) \quad (3.61)$$

ou encore :

$$(H_0 - E_0)\psi(\xi) = (E_{int} - \xi V)\psi(\xi) \quad (3.62)$$

La théorie des perturbations de Rayleigh-Schrödinger nous fournit une série de corrections pour l'énergie [112], de formes identiques à celles déjà fournies pour le seul ordre 2 dans le paragraphe précédent (voir formules 3.51 et 3.55). Par soucis de simplification, nous noterons ces termes :

$$E_{pol}^{(n)} = \langle \phi_0 | V | \phi_{pol}^{(n)} \rangle \quad (3.63)$$

dans lesquels  $n$  représente l'ordre de la première perturbation. Le terme “*pol*” rappelle que l'on se place dans l'approximation de polarisation, sans effet lié à l'échange d'électron entre les monomères, violant ainsi le principe d'antisymétrie. Les fonctions issues d'un tel développement ne sont ni orthogonalisées, ni relaxées.

Il est malheureusement courant de constater que la série perturbationnelle ne converge pas (il arrive même qu'elle diverge), ce même en réalisant un développement aux ordres supérieurs à 2. Pour assurer la convergence de la fonction et de l'énergie, on emploie un terme d'atténuation asymptotique :

$$E_{int} = \sum_{n=1}^N E_{pol}^{(n)} + O(1/R^{2(N+1)}) \quad (3.64)$$

Afin de rétablir le principe d'antisymétrie, un opérateur d'antisymétrisation classique est initialement appliqué sur la fonction d'onde  $\mathcal{A}\phi_{pol}^{(n)}$ , permettant ainsi d'établir l'expression de l'énergie SRS (Symmetrized Rayleigh-Schrödinger), d'où découle le nom de la théorie SAPT, corrigée des différents ordres  $n$  de perturbation :

$$E_{RS}^{(n)} = \frac{1}{\langle \phi_0 | \mathcal{A}\phi_0 \rangle} \left[ \langle \phi_0 | V | \mathcal{A}\phi_{pol}^{(n-1)} \rangle - \sum_{\kappa=1}^{n-1} E_{RS}^{\kappa} \langle \phi_0 | \mathcal{A}\phi_{pol}^{(n-\kappa)} \rangle \right] \quad (3.65)$$

soit :

$$E_{SRS}^{(1)} = E_{pol}^{(1)} + E_{exch}^{(1)} \quad (3.66)$$

$$E_{SRS}^{(2)} = E_{pol}^{(2)} + E_{exch}^{(2)} \quad (3.67)$$

jusqu'à un ordre de perturbation  $n$  déterminé et suffisamment poussé pour contenir toute l'information à longue portée (rappelons que  $pol = disp$  (variant à minima en  $R^{-6}$ ) +  $ind$ ).

Suivant la méthodologie SAPT, le second niveau de perturbation s'obtient en tenant compte de la corrélation intramoléculaire (à présent qu'a été incluse la perturbation intermoléculaire). Ainsi, sur chaque ordre de perturbation intermoléculaire  $k$ , on peut ajouter une série de perturbations intramoléculaires propres au monomère A (d'ordre  $m$  représentatif pour développer l'opérateur  $W^A$ ) et au monomère B (d'ordre  $p$  représentatif pour développer l'opérateur  $W^B$ ). La somme sur tous les ordres,  $m$  et  $p$ , conduit aux formules correctives suivantes :

$$E_{pol}^{(k)} = \sum_{m,p} E_{pol}^{(kmp)} \quad (3.68)$$

$$E_{exch}^{(k)} = \sum_{m,p} E_{exch}^{(kmp)} \quad (3.69)$$

soit une énergie d'interaction corrigée globale :

$$E_{int} = \sum_{k=1}^N \left( \sum_{m,p} E_{pol}^{(kmp)} + \sum_{m,p} E_{exch}^{(kmp)} \right) + O(1/R^{2(N+1)}) \quad (3.70)$$

dans laquelle  $E_{pol}^{(kmp)}$  représente l'énergie de polarisation, identique aux corrections obtenues via la théorie des perturbations de Rayleigh-Schrödinger. Les corrections d'échange  $E_{exch}^{(kmp)}$  proviennent de l'utilisation de l'opérateur d'antisymétrisation permettant l'échange des électrons entre les monomères lorsqu'ils sont suffisamment proches les uns des autres. Le défaut de la théorie des perturbations de Rayleigh-Schrödinger classique est donc corrigé. A une distance interatomique R suffisamment grande pour négliger les termes d'échange, la théorie SAPT est similaire à la théorie des perturbations de Rayleigh-Schrödinger.

En pratique, dans une approche SAPT, quelques termes seulement sont calculés de manière exacte, les autres termes (non tronqués dans les séries perturbationnelles) étant calculés à partir des fonctions d'ondes corrélées sur les monomères. Jeziorski et coll [113, 114] rapportent dans leur travail l'ensemble des termes non nuls issus des séries perturbationnelles développées et montrent que l'on peut aisément séparer la contribution Hartree-Fock à l'énergie d'interaction, des contributions calculées, au moyen d'approches post-HF nécessaires à l'estimation de l'énergie de corrélation pour chaque monomère.

$$E_{int}^{HF} = E_{pol}^{(100)} + E_{exch}^{(100)} + E_{ind}^{(200)} + E_{exch-ind}^{(200)} + \delta E_{int}^{HF} \quad (3.71)$$

Le terme  $\delta E_{int}^{HF}$  contient ainsi tous les termes d'induction et d'échange d'ordres supérieurs à deux.

En s'arrêtant à l'ordre deux de perturbation intermoléculaire, il est possible de calculer explicitement la correction de la partie intermoléculaire sur l'énergie d'interaction :

$$\Delta E_{int}^{inter} = E_{disp}^{(200)} + E_{exch-disp}^{(200)} \quad (3.72)$$

Il est de même possible d'accéder à la correction de la partie intramoléculaire sur l'énergie d'interaction, permettant ainsi d'atteindre une décomposition énergétique globale. À terme pour le présent travail, il sera alors intéressant de relier cette décomposition de l'énergie aux phénomènes physiques spécifiques, tels que la nature des interactions.

Jeziorski et coll [113, 114] relèvent dans leurs travaux que le développement en série du terme d'échange n'est pas unique et a conduit à plusieurs variantes de méthodes SAPT. Ainsi, il est évident que si l'expression des termes  $E_{exch}^{(kmp)}$  est propre à chacune de ces méthodes, la convergence n'est pas équivalente entre les différentes versions de la théorie SAPT. Il est bon de noter que, pour l'étude de systèmes non polaires, dominés principalement par des forces de van der Waals, les différences reportées ci-dessus sont minimes puisque ce sont alors les termes du second ordre (pratiquement identiques dans chaque approche) qui dominent la série perturbationnelle. Le choix de l'approche SAPT n'a dans ce cas pratiquement aucune incidence et il n'est pas nécessaire de pousser l'ordre de perturbation au delà du second ordre. Cette remarque n'est pas sans conséquence, lorsque l'on connaît le coût, en termes de temps de calcul, du recours aux approches SAPT pour l'étude de systèmes moléculaires. Il est à

noter également que l'approche Symmetrized Rayleigh-Schrödinger (SRS), présentée et utilisée dans ce travail, est actuellement considérée comme la méthode la plus adaptée à l'étude des systèmes à faible interaction intermoléculaire, comme les interactions  $\pi - \pi$  pour lesquelles il n'existe, à notre avis, que très peu d'informations.

En pratique, l'énergie d'interaction SAPT peut être calculée à différents niveaux de corrélation intramonomère. Ainsi, d'autres variantes de la méthode SAPT existent. Le recensement exhaustif de ses approches figure dans le travail de Jeziorski et coll. Mais nous citerons toutefois l'approche SAPT(DFT) principalement développée par Misquitta *et al.* [115], Heßelmann et Jansen [116], initiée à partir les idées de Williams et Chabalowski [117] et basée sur le développement de la corrélation électronique des monomères à partir d'un formalisme KS (principalement PBE0), incluant les énergies dispersives calculées au niveau TDDFT. *A priori*, seuls les termes intermoléculaires restent donc à évaluer par le biais du formalisme SAPT.

Enfin, et toujours de façon pratique, un calcul SAPT utilise en principe les orbitales des monomères non relaxées. Aussi, le terme de transfert de charge est partiellement absent. Pour éviter cet écueil, nous avons réalisés, dans nos approches SAPT, pour chaque situation géométrique considérée, une ré-optimisation de chaque paramètre structural (à l'exception du paramètre fixant la distance entre deux monomères). L'importance de ces termes s'est révélée pratiquement négligeable dans nos études.

A titre d'exemple, nous reportons dans ce chapitre un calcul SAPT(DFT) réalisé sur une molécule de référence, choisie tant pour ce qu'elle présente les interactions de type  $\pi - \pi$  stacking qui ont motivé notre travail, que pour l'ensemble de calculs et de mesures expérimentales dont regorge la littérature à son sujet : le dimère du benzène.

Krause *et al.* [118] ont reporté une estimation de l'énergie de liaison du dimère, obtenue à partir de la mesure des énergies de dissociation de l'ion et des potentiels d'ionisation du dimère et du monomère :  $D_0 = 1.6 \pm 0.2$  kcal mol<sup>-1</sup>. Cette valeur est toutefois éloignée de celle reportée initialement par Grover *et al.* [119], qui avançaient  $2.4 \pm 0.4$  kcal mol<sup>-1</sup>. L'ensemble des études expérimentales (et théoriques) disponibles traitent toutes de trois configurations, comme reportées dans la figure ci-dessous 3.1 .

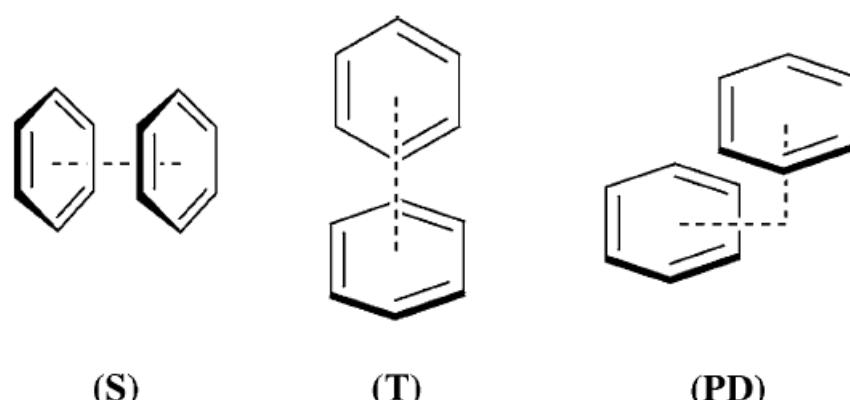


Figure 3.1 – Structures possibles des dimères de benzène : (S) Sandwich, (T) En forme de T et (PD) Parallèle Déplacé

L’analyse de ces travaux montre que le débat visant à déterminer la conformation la plus stable est loin d’être clos. La majorité des travaux annoncent les formes T et PD comme les formes les plus stables. Toutefois, intuitivement, il n’est pas illogique de penser que la configuration S, dite “en sandwich”, qui correspond à une superposition maximale des deux monomères et de fait à une maximisation des interactions dispersives, puisse apparaître comme stable. La configuration parallèle déplacée est, par exemple, souvent observée dans les expériences réalisées à l’état cristallin sur des composés purement aromatiques [120, 121, 122] ou dans des études visant à caractériser les interactions des chaînes latérales aromatiques de protéines [120, 123]. Klemperer *et al.* [124] ont, quant à eux, rapporté que la configuration en T était prédominante à l’état gazeux. D’autres travaux, par Arunan et Gutowsky [125], issus de l’analyse des spectres rotationnels et en accord avec les études Raman de Henson *et al.* [126], rapportent que les structures favorables du dimère de benzène s’approchent de la conformation T sans pour autant la confirmer.

Table 3.1 – Energies d’interaction (en kcal/mol) du dimère de benzène calculées au niveau CCSD(T) et DFT-SAPT (PBE0)

	S		T		PD	
	R		R		R <sub>1</sub>	R <sub>2</sub>
Park et Lee <sup>1</sup>			-2,67	5,0	-3,03	3,5
Tsuzuki et al <sup>2</sup>	-1,48	3,8	-2,46	5,0	-2,48	3,5
Sinnokrot et al <sup>3</sup>	-1,81	3,8	-2,74	5,0	-2,78	3,4
Hobza et al <sup>4</sup>	-1,12	4,1	-2,17	5,1	-2,01	3,6
Sherrill et al <sup>5</sup>	-1,65	3,9	-2,69	5,0	-2,67	3,5
Rocca <sup>6</sup>	-1,06	3,9	-2,15	5,0	-1,82	3,4
DFT-SAPT (PBE0) <sup>7</sup>	-1,47	3,9	-2,44	5,0	-2,35	3,5
						1,7

<sup>1</sup>ref [127], CCSD(T)/CBS, <sup>2</sup>ref [128], CCSD(T)/CBS, <sup>3</sup>ref [129], CCSD(T)/CBS, <sup>4</sup>ref [130], CCSD(T)/aug-cc-pVDZ <sup>5</sup>ref [131], CCSD(T)/CBS, <sup>6</sup>ref [132], RPA,  
<sup>7</sup>SAPT(PBE0)/aug-cc-pVTZ (notre travail)

Il existe dans la littérature une très grande variété d’approches théoriques concernant l’étude des dimères du benzène. Nous faisons le choix de ne reporter, dans le tableau ci-dessous, que les modélisations conduisant aux résultats faisant référence, jugés *a priori* comme les plus précis du fait des méthodes employées. Les travaux de Park et Lee [127], de Tsuzuki *et al.* [128] et de Sinnokrot *et al.* [130] ont tous été obtenus à partir de calculs CCSD(T) avec extrapolation CBS de base (en anglais “Complete Basis Set extrapolation”). Cette méthode est une technique de calcul empirique basée sur un minimum de trois calculs séparés utilisant des bases de plus en plus complètes, telles que les bases de type cc-pVXZ (X=T, Q, 5 . . .). Ceci revient à obtenir une estimation approchée du résultat que l’on obtiendrait en utilisant une base de dimensions infinies. Elle permet donc d’atteindre la limite de précision de la méthode de calcul la plus performante. Nous avons reporté dans le tableau 3.1 les résultats que nous avons obtenu, entre autre, sur la forme PD, cette forme nous intéressant *a priori* plus spécifiquement puisqu’elle a pu être caractérisée expérimentalement à l’état solide (état dans lequel seront les structures à modéliser dans le présent travail, après extraction des données IR obtenues en photo

acoustique). On peut remarquer, en accord avec toutes les données actuellement disponibles par le biais des approches SAPT, que notre valeur d'énergie d'interaction calculée au niveau SAPT(PBE0)/aug-cc-pVTZ est conforme à la valeur attendue. Sans surprise, l'approche SAPT permet donc de reproduire avec une excellente vélocité les énergies d'interaction de systèmes en interaction  $\pi - \pi$  et d'atteindre la précision nécessaire pour reproduire des variations aussi faibles que 1 à 2 kcal/mol.

## 3.2 De la nécessité de la DFT

Dans ce chapitre, nous rappellerons dans un premier temps les fondements de la théorie de la fonctionnelle de la densité, abrégée en 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 sous-partie l'approche KS qui contourne ce problème et légitime certaines approximations. Ces dernières donnant naissance à différents types de fonctionnelles. Après une présentation succincte, nous introduirons la notion de fonctionnelle hybride puis détaillerons le principe des approches qui ajoutent *ad hoc* à un calcul KS usuel des fonctions semi-empiriques rendant compte des effets de dispersion.

### 3.2.1 Les fondements de la DFT

Contrairement aux méthodes Hartree-Fock (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 :

$$\begin{aligned}\rho_{(\vec{r})} &= \int \Psi_{(\vec{r})}^* \Psi_{(\vec{r})} \\ &= \int |\Psi_{(\vec{r})}|^2\end{aligned}\tag{3.73}$$

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  $E_0 = E_{(\rho)}$ .

#### 3.2.1.1 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:

$$\hat{T}_{TF}[\rho] = \frac{3}{10}(3\pi)^{2/3} \int \rho_{(\vec{r})}^{5/3} . d\vec{r}\tag{3.74}$$

Cette fonctionnelle est alors combinée aux expressions classiques des interactions électron-noyau et électron-électron, exprimées elles aussi en fonction de la densité électronique :

$$E_{TH}[\rho] = T_{TH}[\rho] + V_{Ne}[\rho] + V_{ee}[\rho]\tag{3.75}$$

#### 3.2.1.2 Modèle de Thomas-Fermi-Dirac

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

$$\begin{aligned} K[\rho] = E_x[\rho] &= \int \rho_{\vec{r}} \epsilon_x[\rho].d\vec{r} \\ &= -\frac{3}{4} \left(\frac{3}{\pi}\right)^{1/3} \int \rho_{(\vec{r})}^{4/3}.d\vec{r} \end{aligned} \quad (3.76)$$

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 3.74 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.

### 3.2.2 Modèle de Slater

Partant d'une approche basée sur la méthode HF, Slater propose 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 être généralisé en introduisant le paramètre  $\alpha$  :

$$E_x[\rho] = -\frac{9\alpha}{8} \left(\frac{3}{\pi}\right)^{1/3} \int \rho_{(\vec{r})}^{4/3}.d\vec{r} \quad (3.77)$$

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

### 3.2.3 Les théorèmes de Hohenberg et Kohn

Tous ces modèles, bien qu'ils constituent les fondements de la DFT, ne démontrent pas formellement que seule la connaissance de la densité est importante pour accéder à 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.

#### 3.2.3.1 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 :

$$E[\rho] = F_{HK}[\rho] + \int \rho_{(\vec{r})} \nu_{ext}.d\vec{r} \quad (3.78)$$

où :

$$\begin{aligned} F_{HK}[\rho] &= T_e[\rho] + V_{ee}[\rho] \\ \nu_{ext} &= V_{Ne}[\rho] \end{aligned} \quad (3.79)$$

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  $\nu_{ext}$ . L'énergie de l'état fondamental est *a priori* accessible de manière exacte car cette fonctionnelle ne repose sur aucune approximation.

### 3.2.3.2 Second théorème : théorème variationnel

Basé sur l'équation 3.78, Hohenberg et Kohn ont ensuite établi un principe variationnel pour déterminer la densité électronique de l'état fondamental :

$$E[\rho] \geq E[\rho_0] \quad (3.80)$$

avec  $\rho_0$  : densité électronique de l'état fondamental,  
 $\rho$  : densité électronique quelconque.

Dans cette équation, à une densité d'essai  $\rho$  correspond une seule énergie potentielle  $\int \rho(\vec{r})\nu_{ext}.d\vec{r}$  et une seule fonction d'onde  $\Psi_\rho$ . La méthode de double minimisation, *i.e.* sous contrainte de Levy, permet de différencier la fonction d'onde  $\Psi_{\rho_0}$ , correspondant à l'état fondamental, parmi le jeu infini des fonctions d'ondes  $\Psi_\rho$  donnant la même densité. Ainsi, nous pouvons déterminer, parmi toutes les densités, celle qui minimisera l'énergie par la relation suivante :

$$E[\rho_0] = \min_{\rho} \left( \min_{\Psi \rightarrow \rho} (F[\rho] + \int \rho(\vec{r})\nu(\vec{r}).d\vec{r}) \right) \quad (3.81)$$

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

### 3.2.4 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 HF. L'opérateur monoélectronique de Kohn-Sham  $\hat{K}_{KS}$  s'exprime ainsi :

$$\hat{H}_{KS} = -\frac{1}{2}\nabla^2 + \nu_H[\rho] + \nu_{xc}[\rho] + \nu_{ext}[\rho] \quad (3.82)$$

où :

$$\nu_H[\rho] = \int \frac{\rho(\vec{r}) - \rho(\vec{r}')}{|\vec{r} - \vec{r}'|}.d\vec{r}' \quad (3.83)$$

$$\nu_{xc}[\rho] = \frac{\partial E_{xc}[\rho(\vec{r})]}{\partial \rho(\vec{r})} \quad (3.84)$$

$\nu_H[\rho]$  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  $\rho$  :

$$\nu_{eff}(\vec{r}) = \nu_H[\rho] + \nu_{xc}[\rho] + \nu_{ext}[\rho] \quad (3.85)$$

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

$$\hat{H}_{KS}\psi_{i(\vec{r})} = \epsilon_i\psi_{i(\vec{r})} \quad (3.86)$$

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 :

$$E_{KS}^{tot}[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho] + \int V_{ext}(\vec{r})\rho(\vec{r}).d\vec{r} \quad (3.87)$$

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 :

$$E_{reel}^{tot}[\rho] = T[\rho] + V_{ee}[\rho] + \int V_{ext}(\vec{r})\rho(\vec{r}).d\vec{r} \quad (3.88)$$

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

$$E_{xc}[\rho] = T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho] \quad (3.89)$$

La théorie de la fonctionnelle de la densité de Kohn-Sham (KS) a connu un grand succès parmi les méthodes de calcul appliquées aux systèmes de grande dimension, du fait de 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.

### 3.2.5 Les différentes classes de fonctionnelles

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 :

$$\int F_{xc}[\rho(\vec{r})].d\vec{r} = E_{xc}[\rho(\vec{r})] \quad (3.90)$$

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 –  $E_c[\rho]$  :

$$E_{xc}[\rho(\vec{r})] = E_x[\rho(\vec{r})] + E_c[\rho(\vec{r})] \quad (3.91)$$

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 élaboré par Perdew et couramment appelé “échelle de Perdew”. Il a proposé de classer les fonctionnelles en

fonction du degré d'informations non locales contenues 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, dites GGA (“Generalized Gradient Approximation”), dans lesquelles la non-localité est introduite grâce à leur dépendance vis-à-vis du 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.

### 3.2.5.1 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 :

$$E_{xc}^{LDA}[\rho(\vec{r})] = \int \rho(\vec{r}) \epsilon_{xc}[\rho(\vec{r})].d\vec{r} \quad (3.92)$$

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

$$\nu_{xc(\vec{r})}^{LDA} = \epsilon_{xc}[\rho(\vec{r})] + \rho(\vec{r}) \frac{\partial \epsilon_{xc}[\rho(\vec{r})]}{\partial \rho} \quad (3.93)$$

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, dans laquelle il n'est pas tenu compte de l'inhomogénéité de la densité électronique.

### 3.2.5.2 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  $\rho$  à son gradient  $\nabla\rho$ . L'expression générale des fonctionnelles de type GGA est la suivante :

$$E_{xc}^{GGA}[\rho(\vec{r})] = A_x \int \rho(\vec{r})^{4/3} E^{GGA}(s).d\vec{r}^3 \quad (3.94)$$

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

$$s = \frac{|\nabla \rho(\vec{r})|}{2k_F \rho(\vec{r})} \quad (3.95)$$

avec  $k_F = (3\pi^2\rho(\vec{r}))^{1/3}$ . Ainsi, on fait apparaître par le biais du terme  $s$  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 Ernzerhof, notée PBE [133].

### 3.2.5.3 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 :

$$E_{xc}^{hybride}[\rho(\vec{r})] = (1 - \alpha)E_{xc}^{GGA}[\rho(\vec{r})] + \alpha E_{xc}^{HF}[\rho(\vec{r})] \quad (3.96)$$

où le coefficient de la combinaison  $\alpha$  donne le rapport HF/DFT.

La fonctionnelle PBE0 [134], par exemple, présente 25% d'échange HF [135] dans une fonctionnelle GGA de type PBE [133] :

$$E_{xc}^{PBE0} = E_{xc}^{PBE} + \frac{1}{4}(E_x^{HF} - E_x^{PBE}) \quad (3.97)$$

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, tant au niveau du calcul des structures moléculaires et électroniques que pour le calcul de propriétés spectroscopiques [134].

Dans cette catégorie figure également la fonctionnelle la plus utilisée pour traiter des systèmes moléculaires, à savoir la fonctionnelle B3LYP [136]. Comme son nom l'indique, elle inclus trois paramètres et se base sur les fonctionnelles GGA d'échange et de corrélation de Becke (B) [137] et Lee, Yang et Parr (LYP) [138], suivant l'expression :

$$E_{xc}^{B3LYP} = (1 - a)E_x^{LSDA} + aE_x^{HF} + b\Delta E_x^B + (1 - c)E_c^{LSDA} + cE_c^{LYP} \quad (3.98)$$

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

### 3.2.5.4 Nécessité d'une fonctionnelle « longue portée »

Bien que la théorie de la fonctionnelle de la densité connaisse un large succès, elle conduit toujours à des écueils dès lors que l'on tente de l'employer pour l'étude de 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 reste problématique. Se pose alors la question de savoir comment modéliser ces interactions de façon idiomatique. Comme nous allons l'expliciter, les fonctionnelles hybrides dites “à longue portée” permettent de répondre, au premier degré de l'échelle de précision, à cette problématique.

### 3.2.6 Construction d'une LC-DFT-D hybride : cas de la $\omega$ BXD

Les fonctionnelles 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 comparaison aux méthodes DFT hybrides.

### 3.2.6.1 B88

Comme nous l'avons vu dans le cadre des approximations de la fonctionnelle de la densité, notées DFAs<sup>2</sup>, la décroissance exponentielle du potentiel d'échange-corrélation engendre une mauvaise représentation des interactions à longue distance, qui varient quant à elles en  $1/r$ . Cette erreur, baptisée erreur d'auto-interaction (ou SIE, pour “self-interaction error”), est liée au fait que ces approximations se basent sur la densité de spin locale (LSDA, pour “local spin density approximation”). De fait, elles 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 3.2.4), Becke proposa en 1988 d'incorporer dans sa fonctionnelle d'échange B88 [137] une fraction d'échange calculée de façon exacte Hartree-Fock.

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

$$E_{xc} = c_x E_x^{HF} + E_{xc}^{DFA} \quad (3.99)$$

où  $c_x$  prend généralement des valeurs comprises entre 0,2 et 0,25 [136] pour traduire au mieux les propriétés thermodynamiques, et entre 0,4 et 0,6 [139] pour traduire au mieux les données cinétiques.

Basée sur ce modèle, la désormais bien connue DFT hybride B3LYP [136] (équation 3.98) donne des résultats comparables à ceux obtenus à partir de la théorie perturbative Møller-Plesset à l'ordre 2 [140] (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  $E_{xc}[\rho]$ .

À ce stade, il me paraît essentiel d'attirer l'attention du lecteur sur le caractère quelque peu “erratique” des résultats de nombre de travaux menés en DFT, caractère qui tient au choix de la fonctionnelle d'échange et/ou de corrélation. Lorsque, pour un système donné, une fonctionnelle d'échange-corrélation se révèle capable de traduire des paramètres physiques avec une précision satisfaisante autour de la position d'équilibre, rien ne garantit qu'elle soit en mesure de donner la même précision pour un autre (type de) système(s). Il en va de même pour l'étude de dimères (atomiques ou moléculaires). Dans le présent travail, une attention toute particulière a donc été portée au choix de la fonctionnelle d'échange-corrélation, s'agissant d'une étude portant sur une très grande variété de systèmes, *i.e.* non seulement des familles de molécules comportant un ou plusieurs hétéroatome(s) mais les homo-dimères correspondants. La fonctionnelle  $\omega b97X$  a retenu notre attention dans ce travail.

### 3.2.6.2 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 LCAO<sup>3</sup>, Becke 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 [141]. La base de données contient notamment des valeurs relatives à l'interaction entre systèmes conjugués.

<sup>2</sup>“Density Functional Approximations”

<sup>3</sup>“Linear Combination of Atomic Orbitals”.

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 [142] (1998), la B97-2 [143] (2001) et la B97-3 [144] (2005). Il s’agissait alors de réoptimisations des coefficients linéaires par rapport à de nouvelles bases de données expérimentales, plus complètes.

Mais cette nouvelle paramétrisation 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 3.99) pourrait résoudre ce problème mais serait incompatible avec le terme de corrélation DFA  $E_c^{DFA}$ , engendrant une mauvaise compensation des erreurs respectives.

### 3.2.6.3 $\omega$ 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 sur le plan mathématique. Le principe est de traiter séparément, à l’aide d’une fonction erreur ( $erf$ ), les interactions à courte et à longue distance. Les interactions à courte distance sont représentées par une fonctionnelle de la densité, tandis que les interactions à longue distance sont traduites 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 à courte portée.

Iikura *et al.* [145] ont été parmi les premiers à proposer ce type de solutions. Leur méthode consiste à traiter la partie d’échange (LR) par la théorie HF, tandis 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 :

$$E_{xc}^{LC-DFA} = E_x^{LR-HF} + E_x^{SR-DFA} + E_c^{DFA} \quad (3.100)$$

Ce schéma de séparation de portée a l’avantage de conduire à des temps de calcul très proches des DFT hybrides. Reste toutefois à développer une fonctionnelle d’échange SR précise ainsi qu’une fonctionnelle de corrélation compatible avec la fonctionnelle d’échange.

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

$$\frac{1}{r} = \frac{erf(\omega r_{12})}{r_{12}} + \frac{erfc(\omega r_{12})}{r_{12}} \quad (3.101)$$

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,  
 $\omega$  : paramètre contrôlant la séparation.

Dans cette équation, l’introduction du paramètre  $\omega$ , qui s’exprime comme l’inverse d’une distance, permet de donner un sens physique à la fonction d’erreur, puisqu’il est étroitement lié à une longueur caractéristique de la séparation. Naturellement, il existe différents types de fonctions erreur ( $erf$ ), qui facilitent son intégration mathématique dans les codes de calculs. Dans le cas de la fonctionnelle  $\omega$ B97 [146] et de ses améliorations (à savoir les fonctionnelles  $\omega$ B97X et  $\omega$ 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  $E_x^{SR-LSDA}$  obtenue par intégration du carré de la matrice densité LSDA :

$$E_x^{LR-HF} = -\frac{1}{2} \sum_{\sigma} \sum_{ij}^{occ.} \iint \phi_{i\sigma}^*(r_1) \phi_{j\sigma}^*(r_1) \frac{erf(\omega r_{12})}{r_{12}} \phi_{i\sigma}(r_2) \phi_{j\sigma}(r_2) dr_1 dr_2, \quad (3.102)$$

$$E_x^{SR-LSDA} = \sum_{\sigma} \int \underbrace{-\frac{3}{2} \left( \frac{3}{4\pi} \right)^{1/3} \rho_{\sigma}^{4/3}(r) F(a_{\sigma})}_{e_{x\sigma}^{SR-LSDA}(\rho_{\sigma}).dr} dr \quad (3.103)$$

où :

$$\begin{aligned} k_{F\sigma} &= (6\pi^2 \rho_{sigma}(r))^{1/3}, \\ F(a_{\sigma}) &= 1 - \frac{8}{3} a_{\sigma} \left[ \sqrt{\pi} erf\left(\frac{1}{2a_{\sigma}}\right) - 3a_{\sigma} + 4a_{\sigma}^3 + (2a_{\sigma} - 4a_{\sigma}^3) \exp\left(-\frac{1}{4a_{\sigma}^2}\right) \right], \\ a_{\sigma} &= \frac{\omega}{2k_{F\sigma}}. \end{aligned}$$

avec  $k_{F\sigma}$  : vecteur d'onde local de Fermi,  
 $F(a_{\sigma})$  : fonction d'atténuation,  
 $a_{\sigma}$  : 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 pour l'anglais “Range Separated Hybrid eXchange” [147] s'écrit :

$$E_{xc}^{RSXLDA} = E_x^{LR-HF} + E_x^{SR-LSDA} + E_c^{LSDA} \quad (3.104)$$

La fonctionnelle  $\omega B97$ [146] s'écrit alors :

$$E_{xc}^{\omega B97} = E_x^{LR-HF} + E_x^{SR-B97} + E_c^{B97} \quad (3.105)$$

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.

En dépit de nombreuses études visant à optimiser la valeur du paramètre  $\omega$ , la précision calculatoire reste insuffisante du point de vue de la thermochimie. En effet, comme nous l'avons souligné ci-dessus, une valeur trop élevée du paramètre  $\omega$  engendrerait une incompatibilité entre le terme d'échange non-local,  $E_x^{LR-HF}$ , et le terme local de corrélation,  $E_c^{LSDA}$ . De plus, d'après l'équation 3.101, plus  $\omega$  est petit, plus la contribution du terme d'échange (SR),  $E_x^{SR-LSDA}$ , sera importante. Attribuer à  $\omega$  une valeur trop faible reviendrait alors à traiter le problème dans un cadre très proche de la LDA classique qui, comme nous l'avons vu dans la partie 3.2.5.1, est incapable de traduire correctement le terme d'échange à courte portée.

### 3.2.6.4 $\omega B97X$

Afin de remédier à ce problème, une partie d'échange (SR) HF, notée  $E_x^{SR-HF}$ , est ajoutée à  $E_x^{SR-LSDA}$  dans une proportion d'environ 16%, de la même manière que Becke dans la fonctionnelle B88. Ceci présente l'avantage de ne pas perturber la partie LR, qui est quant à elle

correcte. La nouvelle fonctionnelle comporte donc désormais un paramètre, noté  $c_x$ , contrôlant la proportion d'échange exact HF à courte distance :

$$E_{xc}^{LC-DFA} = E_x^{LR-HF} + c_x E_x^{SR-HF} + E_x^{SR-DFA} + E_c^{DFA} \quad (3.106)$$

où :

$$E_x^{SR-HF} = -\frac{1}{2} \sum_{\sigma} \sum_{ij}^{occ.} \iint \phi_{i\sigma}^*(r_1) \phi_{j\sigma}^*(r_1) \frac{\operatorname{erfc}(\omega r_{12})}{r_{12}} \phi_{i\sigma}(r_2) \phi_{j\sigma}(r_2) dr_1 dr_2, \quad (3.107)$$

C'est ainsi que la fonctionnelle  $\omega$ B97X[146] se décompose de la façon suivante :

$$E_{xc}^{\omega B97X} = E_x^{LR-HF} + c_x E_x^{SR-HF} + E_x^{SR-B97} + E_c^{B97} \quad (3.108)$$

La valeur de  $\omega$  et les valeurs des coefficients de développement linéaire et de développement à l'ordre  $m$  des fonctionnelles  $\omega$ B97 et  $\omega$ B97X, 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 (*i.e.* les forces de London). Comme nous allons le voir dans le cas de la fonctionnelle  $\omega$ B97X-D, ces écarts peuvent être corrigés par une prise en compte empirique des effets de dispersion.

### 3.2.6.5 $\omega$ B97X-D

Cette dernière correction pourrait naturellement passer par le calcul de l'énergie de dispersion entre chaque atome, mais ceci engendrerait une augmentation démesurée du coût calculatoire. C'est pourquoi Jeng-Da Chai et Martin Head-Gordon ont fait le choix d'introduire cette correction de façon empirique au travers de l'ajout d'un terme, noté  $E_{disp}$ , à la fonctionnelle KS-DFT (dans le cadre du présent travail, la fonctionnelle  $\omega$ B97X). L'expression de l'énergie de la fonctionnelle  $\omega$ B97X-D [146] ainsi obtenue devient alors :

$$E_{DFT-D} = E_{\omega B97X} + E_{disp} \quad (3.109)$$

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

$$E_{disp} = - \sum_{i=1}^{N_{at}-1} \sum_{j=i+1}^{N_{at}} \frac{C_6^{ij}}{R_{ij}^6} f_{damp}(R_{ij}) \quad (3.110)$$

où :

$$f_{damp}(R_{ij}) = \frac{1}{1 + a(\frac{R_{ij}}{R_r})^{-12}} \quad (3.111)$$

En conclusion, les travaux de Jeng-Da Chai et Martin Head-Gordon ont conduit à la fonctionnelle  $\omega$ B97X-D, de type LC-DFT-D hybride, dans laquelle la totalité de l'échange exact HF est prise 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. La partie empirique a été paramétrée par rapport à la base de données précédemment employée pour les fonctionnelles  $\omega$ B97 et  $\omega$ BX97.

### 3.2.7 DFT+D

Les fonctionnelles DFT-D – *i.e.* BLYP-D, B3LYP-D, B97-D, PBE-D, M05-2x et M06-2x – sont traditionnellement incluses dans la famille des fonctionnelles dites “semi-empirical dispersion-corrected functionals”. En effet, pour ces fonctionnelles, les effets de corrélation à longue distance sont exclusivement traités par une correction empirique de dispersion selon trois niveaux d’empirisme notés respectivement DFT-D1, DFT-D2 et DFT-D3.

#### 3.2.7.1 DFT-D2

S. Grimme [83] propose en 2006 une première amélioration de sa correction empirique initiale de la dispersion [148], appelée D2, dans le but de corriger trois défauts principalement constatés :

- les coefficients  $C_6$  n’étaient disponibles, dans la version originale, que pour les atomes les plus légers de la classification périodique
- l’utilisation des données tabulées pour les atomes de la troisième période conduisaient à des erreurs systématiques
- les résultats obtenus ne permettaient pas de retrouver certains principes usuels de la thermochimie.

Conformément à ce que nous avons reporté dans le paragraphe précédent, l’expression de l’énergie de la fonctionnelle s’exprime toujours comme :

$$E_{DFT-D2} = E_{KS-DFT} + E_{disp}^{(2)} \quad (3.112)$$

L’énergie de dispersion  $E_{disp}$  est quant à elle de nouveau définie par rapport à une fonction d’amortissement  $f_{damp}$  :

$$E_{disp}^{(2)} = -s_6 \sum_{i=1}^{N_{at}-1} \sum_{j=i+1}^{N_{at}} \frac{C_6^{ij}}{R_{ij}^6} f_{d,6}(R_{ij}) \quad (3.113)$$

où :

$$f_{d,6}(R_{ij}) = \frac{s_6}{1 + \exp^{(-d(\frac{R_{ij}}{s_R R_{Oij}} - 1))}} \quad (3.114)$$

dans laquelle le paramètre  $s_6$  est optimisé suivant la fonctionnelle d’échange-corrélation utilisée. Les coefficients de van der Waals d’ordre 6 et les rayons de van der Waals des espèces en interaction sont calculés à partir des données tabulées sur les atomes à partir des expressions suivantes :

$$C_{6ij} = \sqrt{C_{6ii} C_{6jj}} \quad (3.115)$$

$$C_{6ii} = 0.05 N I_p i \alpha_i \quad (3.116)$$

et

$$R_{Oij} = R_{Oi} + R_{Oj} \quad (3.117)$$

Les termes  $I_p i$  et  $\alpha_i$  représentent respectivement le potentiel d'ionisation atomique et la polarisabilité statique de l'élément  $i$ .

La correction D2 a été principalement testée par Grimme [83] pour l'étude des systèmes de référence que sont les gaz diatomiques, le benzène ainsi que de petites molécules aromatiques telles que l'anthracène. Quelques années plus tard, Park *et al.* [149] ont étudié des systèmes graphitiques et ont rapporté d'excellentes corrélations avec les données expérimentales, notamment en ce qui concerne la description des paramètres de maille. D'autres auteurs tels que Lee *et al.* [150] ont commencé à publier, dès 2013, des travaux dont le but était de calculer des fréquences de vibration à l'aide du logiciel VASP.

### 3.2.7.2 DFT-D3

Quatre ans plus tard, en 2010, S. Grimme [83] propose une nouvelle amélioration de sa correction empirique initiale de la dispersion [148], baptisée D3. Dans cette version, la majorité des termes empiriques sont calculés au niveau KS-DFT.

L'énergie de la fonctionnelle s'exprime désormais comme :

$$E_{DFT-D3} = E_{KS-DFT} + E_{disp}^{(2)} + E_{disp}^{(3)} \quad (3.118)$$

où :

$$E_{disp}^{(2)} = - \sum_{i>j} \sum_{n=6,8,10,\dots} s_n \frac{C_n^{ij}}{R_{ij}^n} f_{damp}^n(R_{ij}) \quad (3.119)$$

$$E_{disp}^{(3)} = - \sum_{i>j>k}^{N_{at}} \frac{C_9^{ijk}(3 \cos \theta_I \cos \theta_J \cos \theta_K + 1)}{(R_{ij} R_{jk} R_{ki})^3} f_{damp}^{(3)}(\bar{R}_{ijk}) \quad (3.120)$$

Dans cette dernière expression, les termes  $\theta_I$ ,  $\theta_J$  et  $\theta_K$  représentent les angles internes du triangle  $ijk$ . Deux fonctions d'amortissement complètent l'expression de l'énergie :

$$f_{damp}^n(R_{ij}) = \frac{1}{1 + 6(R_{ij}/(s_{r,n} R_O^{ij}))^{-\alpha_n}} \quad (3.121)$$

$$f_{damp}^{(3)}(\bar{R}_{ijk}) = \frac{1}{1 + 6(\bar{R}_{ijk}/(4R_O^{ijk}/3))^{-16}} \quad (3.122)$$

Dans cette équation :

- les rayons de coupure  $R_O^{ij}$  et  $R_O^{ijk}$  sont évalués de façon semi-empirique,
- les coefficients  $s_n$  ( $n = 8, 10, \dots$ ) sont ajustés à partir de références dont les valeurs sont ajustées selon la fonction d'échange-corrélation utilisée
- les coefficients de van der Waals sont évalués au moyen de calculs TDKS et à l'aide de relations récursives,
- le terme  $C_6^{ij}$ , jusque-là calculé au moyen de formules d'interpolation dérivées de manière empirique est, dans la version D3, évalué à partir de l'expression de Casimir et Polder selon la procédure rappelée en annexe de ce travail:

$$C_6^{ij} = \frac{3}{\pi} \int_0^\infty \alpha^i(i\omega) \alpha^j(i\omega) d\omega \quad (3.123)$$

Les coefficients d'ordre supérieur sont quant à eux évalués à l'aide des formules de récursivité établies par Starkschall et Gordon [151] :

$$C_8^{ij} = 3C_6^{ij} \sqrt{Q_i Q_j} \quad (3.124)$$

avec

$$Q_i = s \sqrt{Z^i} \frac{\langle r^4 \rangle r^i}{\langle r^2 \rangle r^i} \quad (3.125)$$

$$C_{10}^{ij} = \frac{49}{40} \frac{(C_8^{ij})^2}{C_6^{ij}} \quad (3.126)$$

et

$$C_{n+4}^{ij} = C_{n-2}^{ij} \left( \frac{C_{n+2}^{ij}}{C_n^{ij}} \right)^3 \quad (3.127)$$

En plus de fournir un potentiel aux longues distances plus précis, les potentiels DFT-D3 fournissent également une séparation plus nette entre les termes de courte et de longue portée.

### 3.2.7.3 La méthode de Tkatchenko-Scheffer (DFT-TS)

Cette méthode a été proposée en 2009 par Tkatchenko et Scheffler[152]. De nombreux travaux, essentiellement dédiés à l'étude de structures cristallines, ont été menés en utilisant cette correction. A titre d'exemple, citons notamment Kronik et Tkatchenko[153], qui ont étudié la structure du cristal d'Hemozoin, substance potentiellement responsable des fortes fièvres caractéristiques du paludisme. Les auteurs ont démontré la fiabilité de la méthode en calculant en particulier les paramètres de maille du cristal. D'autres cristaux ont été ainsi caractérisés, tels que le cristal de Brushite – qui n'est autre qu'un phosphate de calcium hydraté – au moyen de calculs de modes de vibration. Très récemment, Bučko et coll. [154] ont montré que cette méthode pouvait également servir à l'étude de systèmes ioniques, à condition d'apporter quelques modifications au calcul du volume effectif et des charges des atomes, comme nous allons le préciser ci-après.

Dans cette méthode, l'expression de l'énergie de dispersion est identique à celle établie pour la méthode DFT-D2. La différence majeure réside dans le calcul des coefficients de dispersion et de la fonction de damping qui sont, dans cette approche, dépendants de la densité de charge. Par ailleurs, la méthode DFT-TS est, en toute rigueur, également capable d'inclure les variations des contributions atomiques intervenant dans le calcul de l'énergie d'interaction, au moyen d'une prise en compte des effets d'environnement.

Les coefficients de dispersion, la polarisabilité et les rayons atomiques sont calculés à partir des expressions suivantes :

$$\alpha_i = \nu_i \alpha_i^{free} \quad (3.128)$$

$$C_{6ii} = \nu_i^2 C_{6ii}^{free} \quad (3.129)$$

$$R_{0i} = \left( \frac{\alpha_i}{\alpha_i^{free}} \right)^{\frac{1}{3}} R_{0i}^{free} \quad (3.130)$$

expressions dans lesquelles  $\alpha_i^{free}$  est tel que :

$$\alpha_i^{free} = \frac{V_i^{eff}}{V_i^{libre}} \quad (3.131)$$

Le terme  $\alpha$  représente le rapport entre le volume effectif occupé par un atome  $i$  dans une molécule ou un solide (noté  $V_i^{eff}$ ), et le volume de l'atome isolé (noté  $V_i^{libre}$ ) obtenu à l'aide d'une partition de Hirschfeld de la densité :

$$\nu_i = \frac{V_i^{eff}}{V_i^{libre}} = \frac{\int r^3 w_i(r) n(r) d^3 r}{\int r^3 n_i^{free}(r) d^3 r} \quad (3.132)$$

où  $n_i^{free}$  est la densité sphérique moyenne de l'atome isolé  $i$  et où  $w_i(r)$  représente le poids de Hirschfeld, défini à partir des densités atomiques libres :

$$w_i(r) = \frac{n_i^{free}(r)}{\sum_{j=1}^N n_j^{free}(r)} \quad (3.133)$$

Les paramètres  $\alpha_i^{free}$ ,  $C_{6ii}^{free}$  et  $R_{0i}^{free}$  sont tabulés pour tous les éléments des premières périodes de la classification périodique, à l'exception des lanthanides.

### 3.2.7.4 Self-consistent screening de la méthode de Tkatchenko-Scheffer (DFT-TS-SCS)

En l'absence de champ électrostatique, les interactions entre deux systèmes sont décrites en fonction de leurs moments permanents et des polarisabilités statiques. Pour un champ périodique de pulsation  $\omega$ , les interactions sont modifiées et les polarisabilités dépendent de cette pulsation. L'étude des forces de dispersion à longue distance nécessite alors la connaissance d'intégrales qui tiennent compte de l'ensemble des fréquences  $\omega$ . Pour des raisons mathématiques évidentes, les intégrations de fonctions discontinues restent complexes. Il convient alors de se ramener à la connaissance de la variation des polarisabilités dépendantes des fréquences imaginaires de chacun des deux systèmes pris isolément, puisque ces fonctions sont continues (certains détails figurent en Annexe de ce travail).

La variante DFT-TS-SCS de la méthode Tkatchenko-Scheffler (DFT-TS) a été proposée en 2012 par Tkatchenko et coll.[155] avec pour objectif de tenir compte de la perturbation due à un champ oscillant de fréquence  $\omega$ . Par la méthode DFT-TS-SCS, la polarisabilité dynamique s'obtient à partir de la résolution de l'équation auto-cohérente :

$$\alpha_i^{SCS}(i\omega) = \alpha_i^{TS}(i\omega) - \alpha_i^{TS}(i\omega) \sum_{i \neq j} \tau_{ij} \alpha_j^{SCS}(i\omega) \quad (3.134)$$

dans laquelle  $\alpha_i^{TS}(i\omega)$  représente la polarisabilité effective dynamique dépendante du champ, approximée par l'expression :

$$\alpha_i^{TS}(i\omega) = \frac{\alpha_i^{TS}(0)}{1 + (\omega/\omega_i)^2} \quad (3.135)$$

La fréquence/pulsation du champ est estimée au moyen d'une formule empirique du type :

$$\omega_i = \frac{4}{3} \frac{C_{6,ii}^{TS}}{(\alpha_i^{TS})^2} \quad (3.136)$$

dans laquelle les termes statiques de polarisabilité  $\alpha_i^{TS}$  et les coefficients de van der Waals d'ordre six  $C_{6,ii}^{TS}$  sont estimés selon les équations présentées ci-avant lors de la présentation de la méthode DFT-TS.

Finalement, les coefficients de dispersion d'ordre 6 utilisés dans l'expression de l'énergie de dispersion  $E_{disp}$ -D2 sont évalués au moyen de l'expression de Casimir-Polder :

$$C_6^{ij} = \frac{3}{\pi} \int_0^\infty \alpha_i^{SCS}(i\omega) \alpha_j^{SCS}(i\omega) d\omega \quad (3.137)$$

et les rayons de van der Waals des atomes sont recalculés au moyen de l'expression :

$$R_i^{SCS} = \left( \frac{\alpha_i^{SCS}}{\alpha_i^{TS}} \right)^{1/3} R_i^{TS} \quad (3.138)$$

qui fait intervenir les expressions des rayons atomiques, calculés par la méthode de Tkatchenko (DFT-TS).

Tout comme dans l'approche DFT-D2, le paramètre empirique  $s_R$  figurant dans l'expression de la fonction d'amortissement est généralement fixé à 0,97 (pour la fonctionnelle PBE) conformément aux recommandations de Bučko, Lebègue et al [156].

La méthode DFT-TS-SCS est disponible dans le code VASP grâce à Bučko, Lebègue et coll.[156]. Les auteurs ont étudié une large variété de solides (gaz nobles, cristaux ioniques et solides moléculaires), ainsi que des structures de type chaîne et des métaux. La méthode se révèle traduire avec une précision raisonnable les propriétés structurales et les énergies de cohésion, hormis pour les solides ioniques. La détermination des propriétés structurales, du volume de la maille, etc., sont autant de facteurs qu'il nous faut estimer avec le plus grand soin pour mener à bien les calculs de fréquences de vibration. Toutes les informations nécessaires à la comparaison des approches D2, D3, TS et TS-SCS seront données dans ce manuscrit dans le chapitre 6. Par ailleurs, les auteurs ont également montré que les effets dynamiques du champ périodique traités dans la méthode TS-SCS sont en général négligeables pour des systèmes en interaction faible, comme c'est le cas des atomes et des molécules neutres qui nous intéressent de prime abord dans ce travail.

## Chapter 4

# Discerning inter- and intramolecular vibrations of sulfur polyaromatic compounds

## 4.1 Introduction

Thiophenes are an important class of molecules in various domains of chemistry including petrochemistry [157, 158, 159], molecular electronics [160, 161, 162, 163] and optoelectronics[164]. In particular, they are of great interest in all that concerns oil upgrading [165, 166, 167]. Their presence as aromatic fragments in asphaltenes has been revealed by methods such as thermolysis analysis[168], mass spectroscopy [169, 170], and NMR and IR spectroscopies [171, 172]. Importantly, thiophenes are hypothesized to play a key role in the aggregation of asphaltenes [173, 174] which is of uppermost importance to the oil extraction process. Asphaltenes are known to cause the blocking of pipelines and refinery equipment.

The molecular structure of asphaltenes is difficult to determine since they are formed by thousands of molecules that tend to stick together in solution. Over the past six decades, the composition, molecular structure and colloidal state of asphaltenes in crude oil and bitumen have been considered a central enigma of petroleum chemistry [175]. The noncovalent interactions responsible for molecular recognition among components determine the microstructure of the aggregates. Two contrasting molecular architectures have been proposed to describe the dominant molecular motifs in asphaltenes: the island model in which asphaltene molecules contain one polyaromatic core with pendant aliphatic chains; and the archipelago model in which several aromatic motifs are bridged together via aliphatic chains [176, 12]. Interactions indicated by the island model are dominated by  $\pi$ -stacking among polyaromatic cores and steric repulsion among pendant chains. In the archipelago model diverse interactions arise from van der Waals forces, hydrogen bonding and interactions among heteroatom groups. Porphyrins also play major roles. The identification of interactions in asphaltenes, the understanding of how they influence molecular arrangements, and finally the experimental method by which they are identified, are major issues in production, transport, and refinery system design and operation [177]. For example, understanding the impact of inter- and intramolecular interactions occurring between thiophene-like molecules is a missing and necessary refinement to asphaltene aggregation models [19, 178, 176], that will lead to better process designs for cracking asphaltene aggregates and facilitate refining of thiophene-rich crude oils.

Several types of interactions coexist among aromatic molecules containing sulfur, mainly  $\pi$ - $\pi$ -stacking between aromatic rings, and  $\pi$ -S between an aromatic ring and a sulfur atom. The interaction between aromatic rings and sulfur atoms has been shown by Morgan *et al*[179] to play an important role in stabilizing the folded conformations of 8 different proteins. Tauer *et al* [180] studied the H<sub>2</sub>S-benzene dimer system in order to visualize molecular structures resulting from sulfur- $\pi$  interactions, and to establish an appropriate computational method capable of describing these interactions. The interactions driving the spatial arrangement of thiophene molecules have been studied theoretically by Tsuzuki *et al* [181] who showed that the major interaction among thiophene molecules was dispersive, whereas the electrostatic interactions were responsible for the orientation of the molecules.

In our laboratory previous work [49], they showed that interactions responsible for  $\pi$ - $\pi$  stacking for acene molecules could be visualized in infrared and Raman spectra. The frequencies related to these interactions were found between 15 and 100 cm<sup>-1</sup>. This range of frequencies is difficult to measure experimentally using conventional methods but with the development of terahertz spectroscopy, including synchrotron-based photacoustic (PA)

IR measurements, it is now possible to identify absorption bands specific to intermolecular interactions that occur at very low frequencies and to evaluate their importance.

In the current study we (1) identify the presence of thiophene-type molecules in complex systems, (2) identify  $\pi$ -stacking interactions when they are predominant in a complex system by means of vibrational spectroscopy experiments and high-precision quantum mechanical calculations, and (3) develop a new variation-perturbation method adapted to the resolution of the vibrational Schrödinger equation at very low wavenumbers for large-size systems (bigger than 15 atoms) and their dimers. These results are used to characterize and identify IR and Raman signatures of vibration modes found at very low wavenumbers that are characteristic of association, e.g. dimerization. More specifically, infrared and Raman spectra of benzothiophene **1a**, dibenzothiophene **1b**, 4,6-dimethylbibenzothiophene **1c** and 4-methyldibenzothiophene **1d** are reported. The observed bands are assigned using high level calculations based on anharmonic electrical and mechanical approximations. Fundamental bands previously assigned by Bree *et al* [182] are confirmed, and additional combination bands are identified. The approach and outcomes illustrate how inter- and intramolecular vibrations in single- and multi-component mixtures can be discriminated.

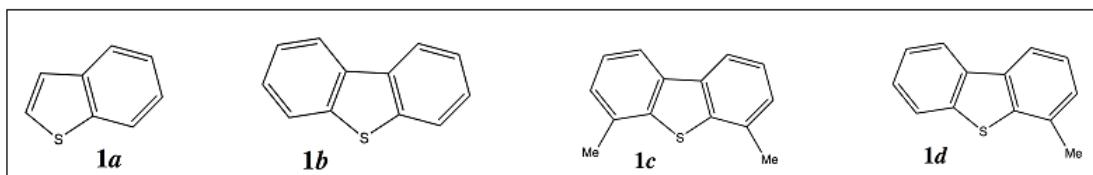


Figure 4.1 – Molecular structures studied: from left to right: benzothiophene **1a**, dibenzothiophene **1b**, 4,6-dimethylbibenzothiophene **1c** and 4-methyldibenzothiophene **1d**.

## 4.2 Experimental Spectra

Dibenzothiophene **1b**, 4,6-dimethylbibenzothiophene **1c** and 4-methyldibenzothiophene **1d** were obtained from commercial sources at purity levels of at least 98 per cent and analyzed as received. In some cases, samples were ground manually prior to acquisition of their spectra.

### 4.2.1 Raman spectra

Raman spectra were obtained using two spectrometers at CanmetENERGY (Devon, Alberta). In one set of experiments, a Renishaw InVia microscope-based system was coupled to a Spectra Physics 125 He-Ne laser, which provided excitation at 633 nm. Several milligrams of each solid were examined with the use of a  $50\times$ microscope objective. Spectra were acquired with a spherically focused incident beam; typical resolution was better than  $2\text{ cm}^{-1}$ . Total accumulation time was 5 min for each spectrum. The second group of experiments was performed with a Bruker IFS 88 FT spectrometer and FRA 106 Raman accessory. Excitation of the FT-Raman spectra was effected with a 1064-nm Nd:YAG laser at power levels between 90 and 260 mW. Ten 32-scan spectra with a resolution of  $8\text{ cm}^{-1}$  were averaged for each sample.

### 4.2.2 Far-infrared spectra

PA far-infrared spectra were acquired at a resolution of  $6\text{ cm}^{-1}$  using two different Bruker IFS 66v/S Fourier transform infrared (FT-IR) spectrometers. The first instrument, at CanmetENERGY, was operated in both rapid- and step-scan modes. The He-Ne laser modulation frequency was 1.6 kHz for the rapid-scan measurements. Large amplitude phase modulation [183] was employed in the step-scan experiments, with a Signal Recovery lock-in amplifier being utilized for demodulation. The second spectrometer was located at the Canadian Light Source (Saskatoon, Saskatchewan). Rapid-scan spectra were acquired under conditions similar to those just mentioned. A thermal (globar) source was employed with this instrument.

Standard MTEC 300 gas-microphone PA accessories were used with both spectrometers; the sample cells were fitted with polyethylene windows. Helium was employed as the carrier gas. Multilayer mylar beam splitters were installed in the interferometers. Spectra were recorded between about 50 and  $700\text{ cm}^{-1}$ , due to limits imposed by the optical materials. Carbon black powder or an MTEC carbon black reference yielded reference spectra that were used to correct sample spectra for the wavenumber-dependent response of each instrument.

### 4.2.3 Mid-infrared spectra

PA mid-infrared spectra were recorded at a resolution of  $6\text{ cm}^{-1}$  using the Bruker IFS 88 spectrometer mentioned above. The laser frequency was either 1.6 or 2.2 kHz. An MTEC 200 PA cell was utilized. Nitrogen was used to purge the spectrometer and as a carrier gas. A KBr window sealed the PA cell, and a Ge/KBr beam splitter was used in the spectrometer. Ten 32-scan spectra were averaged for each sample. Spectra were acquired from about 400 to  $4000\text{ cm}^{-1}$ . Carbon black powder was used to obtain reference spectra in these experiments.

## 4.3 Calculation method for thiophene-type molecules and their dimers

### 4.3.1 Structural details

The molecular structures of thiophene-like molecules and their respective dimers were calculated using Density Functional Theory (DFT) with the  $\omega$ B97X-D exchange-correlation functional and the 6-311++G\*\* basis set, implemented in the Gaussian 09 suite of programs. The  $\omega$ B97X-D functional developed by Chai *et al* [184] is based on optimized long-range corrected hybrid density functionals, which employ 100% Hartree-Fock (HF) exchange for long-range electron-electron interaction correction. This method, tested by Salzner *et al* [185] for  $\pi$ -conjugated oligomers, was also benchmarked successfully for interacting polycondensed aromatic molecules [49], thus proving it overcomes the failure of standard DFT calculation approaches to describe long-distance interactions. The 6-311++G\*\* basis set is of triple-zeta quality for the valence electrons. Among the basis sets tested by Wiberg [186], including the Dunning correlation basis sets such as aug-cc-pVTZ and aug-cc-pVQZ, the 6-311++G\*\* basis

set gave accurate geometries and frequencies at a relatively small computational cost. The calculations use planar “C<sub>2</sub>v symmetry” for dibenzothiophene, 4,6-dimethyldibenzothiophene, and C<sub>s</sub> for benzothiophene and 4-dimethyldibenzothiophene. The geometrical parameters of the molecules are benchmarked with experimental data when available.

For dimers, as an alternative to the supermolecular treatment, an approach based on symmetry-adapted perturbation theory (SAPT) [114] which utilizes the description of the interacting monomers in terms of Kohn-Sham (KS) orbitals was used in this work. The DFT-based SAPT approach [187], denoted as SAPT-DFT in this work, is exact for all major components of the interaction energy (asymptotically for exchange interactions) in the sense that these energies would be exact if the DFT description of the monomers were exact. It provides accurate results for small dimers such as (C<sub>6</sub>H<sub>6</sub>)<sub>2</sub> where results for the much more computationally expensive CCSD(T) method are well approximated [188]. The SAPT-DFT calculations were performed using the MOLPRO2012 package [189]. The PBE0 functional [134] was utilized as a monomer DFT functional in SAPT, using the aug-VTZ basis set.

## 4.4 Resolution of the vibrational Schrödinger equation (RVSE)

Without any consideration of molecular size, the theoretical process generally developed to solve the Vibrational Schrödinger Equation in the mechanical anharmonic hypothesis requires two restricted steps: the construction of the potential energy surface (PES), and the resolution of the vibrational equation (RVE) in order to obtain the energy levels and consequently the fundamental, overtone and harmonic wavenumbers comprising vibrational spectra. Calculations of the intensities in the electrical anharmonic hypothesis generally complete the analysis of the spectra.

Owing to the size of the systems studied in this work, we wondered about the precision we could get on PES that include hundreds of thousands of terms. The search for an optimal analytical expression of the PES is an ongoing topic of discussion, even in cases of low-dimensional systems. The skills and experience of some research groups such as that of V. Barone [190] show that the PES of large-size systems can be developed by expressing the potential function in a simple way, *i.e.*, a Taylor series expansion using curvilinear displacement coordinates. These series were generally truncated at fourth order. Quadratic, cubic, and quartic force constants were obtained by fitting the electronic energy data generally calculated by DFT methods for various nuclear configurations close to the respective optimized geometry.

Of course, vibrational spectroscopy modeling outcomes are directly dependent on the quality of the potential energy surface (PES) but also depend on the mathematical method used to solve the Schrödinger vibrational equation. In our last work,[71] where we focused on the search for a new variational approach to solve this equation, we compiled some methods - direct or indirect, variational or perturbative - that appeared among the most accurate from our point of view. It is generally assumed that variational or VCC (Vibrational Coupled Cluster) approaches take into account the main couplings between the 3N – 6 vibrational modes of a given system more efficiently than the perturbational approaches. However, the use of such methods is restricted to low-dimension systems (including less than 10 atoms)

or highly symmetrical systems. Conversely, vibrational studies of larger systems are now common since the VPT2 approach is implemented in most quantum chemistry codes. To improve the description of perturbations, other combined approaches that attempt to take advantage of both variational and perturbational approaches are utilized. These are known as variation-perturbation methods. Within these approaches, only a few vibrational states  $\phi_0^n(q)$  (from tens to thousands constructed from linear combinations of the  $3N - 6$  normal modes - see ref[71] for notations) strongly coupling the state(s) for which the vibrational characteristics are to be determined are selected by a procedure of interactively enriching the variational space. The energetic correction of all the non-selected states  $\varphi_0^n(q)$  is later estimated by perturbation theory. The sum  $\mathbf{B} = \phi_0^n(q) + \varphi_0^n(q)$  represents the global work basis. Use of these approaches reduces the dimension of the variational problem dramatically[191, 192] since the dimension of the active space to be diagonalized is, now, under control. At each iteration only a small predetermined number of vibrational states ( $\sim 100$ ) is needed to obtain the variational space containing the most important couplings. While this procedure converges the active space in a few cycles and leads to matrices (representations of the vibrational Hamiltonian on the basis of the normal modes) of only a few thousand terms for small systems, this approach has yet to converge for systems larger than 20-25 atoms.

Here we show that a compromise is always possible in order to reduce the active space dimension that is required to converge the variational problem. Considering that all  $3N - 6$  vibrational modes of a molecule are not strongly coupled, two “truncated” approaches were developed and tested.

Truncation approach n°1 includes verifying the influence of the truncation of the basis  $B$  for studying a mode  $x$  (belonging to the  $T = 3N - 6$  ensemble) taking into account, besides the mode being studied, only the  $X' = x + a'$  ( $a' \in \mathbb{N}$ ) lower-frequency and the  $Y' = [3N - 6] - y'$  higher-frequency normal modes (where  $y'$  stands for the last non-selected mode within the active space) (see Figure 4.2). Thus, one has a truncated problem dimension equal to  $T' = X' + Y' < [3N - 6]$ . The choice behind this truncation method relies on the fact that large amplitude modes, such as torsion and rotation, are generally coupled to the elongation modes that act on lighter atoms, such as hydrogen. We are interested in these particular modes. Within this truncation, only a few vibrational states  $\phi'_0(q)$ , strongly coupled to the state(s) for which the vibrational characteristics are to be determined, are selected by a procedure that enriches the variational space. The energetic correction of all the non-selected states  $\varphi'_0(q)$  is then estimated by perturbation theory. The sum  $\mathbf{B}' = \phi'_0(q) + \varphi'_0(q)$  represents the new reduced global work basis.

Truncation approach n°2 was used to verify a hypothesis normally assumed to be true - that every  $x$  state is *a priori* more strongly coupled with its closest energetic neighbors. To do so, we have prepared an alternate selection criterion. Concerning the selection of the lower-frequency modes, we state a new criterion based on the selection of  $X''$ : in order to study a vibrational state  $x$ , one must select  $2x''$  normal modes within the basis ( $x''$  being a positive integer) (see Figure 4.2). The selected  $2x''$  normal modes within the truncated basis are exactly those that are energetically as close as possible to  $x$ . This basis has  $X''=1+2x''$  lower-frequency and  $Y''=Y'$  higher-frequency states. The final dimension of this new basis is  $T''=X''+Y''$ . If  $x''$  equals zero, this is the same as studying the influence of high-energy stretching modes on  $x$ . This new procedure generates a new working basis denoted  $\mathbf{B}''$ .

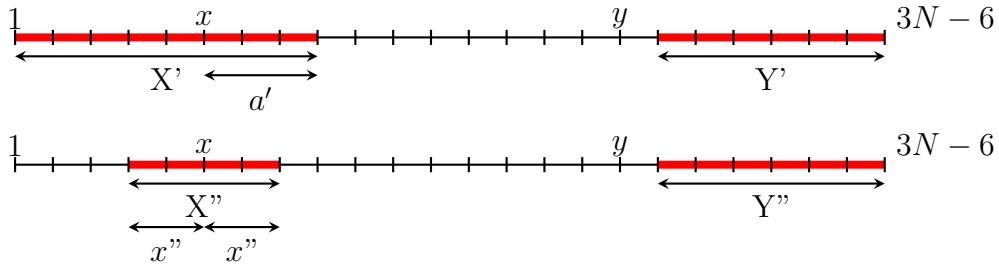


Figure 4.2 – Selection of normal modes which are stored in the expression of the PES : methodologies applied to realize the truncation of the vibrational basis a) truncation n°1  $T'=X'+Y'$  (at the top); b) truncation n°2  $T''=X''+Y''$  (at the bottom).

## 4.5 Results and Discussion

### 4.5.1 Monomers

Benzothiophene ( $N=15$ ) and dibenzothiophene ( $N=21$ ) were chosen to test basis truncation. For each of these systems, a fully converged calculation of the variation-perturbation type was performed. The ensembles of the results obtained using the two truncated bases for the first twelve modes for benzothiophene are reported in Figures 4.3 (truncation n°1) and 4.4 (truncation n°2). Of these twelve modes, modes 1, 2, 4, 5, 8, 10 and 12 occur in the  $C_s$  plane whereas modes 3, 6, 7 and 9 are out of the  $C_s$  plane.

For truncation n°1, the main results are reported here in Figures 4.3a - l. The information reported in these graphs represents the difference obtained on the calculation of a target vibration frequency in the reduced basis  $\mathbf{B}'$  from the result that would be obtained using the same variation-perturbation method with the full basis  $\mathbf{B}$ . For example, the analysis of the target mode 1 ( $X'=1$ , Figure 4.3a) shows that the presence of only the higher-frequency  $\omega_{CH}$  stretching modes suffices to describe the vibrational problem concerning this mode. This observation accords well with the hypothesis that only the large amplitude modes are strongly coupled with the elongation ones involving lighter atoms. The analysis of the other out-of-plane  $C_s$  deformation modes (2, 4, 5 and 8) leads to similar conclusions.

The results of the calculation on mode 2 are closer to those obtained for mode 1, which seems logical since these two vibrational modes are *de facto* similar. Concerning modes 4 and 5, regardless of the number of selected states within the basis created from the  $X'$  criterion, one can note the need to work with at least 13 modes on the  $Y'$  basis in order to be as close as possible to the reference calculation.

On the other hand, one observes that it is not necessary to take into account all the lower-frequency modes within the basis for the calculation to converge. This will be confirmed shortly by the analysis of the second truncation. Another striking observation also emerges from this analysis: if it is desirable to introduce some of the wagging, scissoring and stretching modes in the basis, the calculation will converge satisfactorily only if all of these modes are

introduced at the beginning. Concerning mode 8, it can be noted that no truncation seems capable of leading to an acceptable convergence.

We have added the results for modes 10 and 12 although they do not belong to the same low-frequency category considered in this study. These two modes lie at the interface between the low-frequency and the other modes. The main conclusion one can derive from them is that taking into account only the low frequency and  $\omega_{CH}$  modes in the basis no longer suffices for convergence of the calculations. These conclusions are generalized beyond mode 12.

The modes taking place in the plane of the molecule, *i.e.* modes 3, 6, 7, 9 and 11, again corroborate our hypothesis; particularly for low-frequency mode 3. In addition, the  $x'$  modes that are close to mode 3 seem to have no effect on the quality of the converged result. The descriptions of modes 6 and 7 show again that the results are not greatly dependent on the basis truncation unless all the  $\omega_{CH}$  modes were previously selected. Analyzing the results for modes 9 and 11 leads to the same conclusion, except that a strong truncation of  $Y'$  would cause a much more important deviation of the results from those observed for modes 3, 6 and 7.

The analysis of the results concerning the second truncation allows one to confirm even more the observations just presented. Only the results obtained for the modes 4, 5, 6 and 8 are reported here. Representations of these results were added as a second layer to those obtained beforehand (Figure 4.4a-d insets); the two situations are connected by a series of points they share in common, *i.e.* when  $a'=x''$ . The main information we can derive from these studies is that no low-frequency mode requires its closest neighbors to ensure a good convergence of the results. Every argument indicates that these modes are only coupled to those involving  $\omega_{CH}$  stretching. Only modes 4 and 5 seem to slightly depend on each other.

In summary, variational-perturbational studies using a truncated basis taking into account the closest neighbors of the low-frequency modes seems not to be an essential requirement for describing their vibrational properties as opposed to their coupling with the high-frequency stretching  $\omega_{CH}$  modes. Finally, variational-perturbational studies using a truncated basis were performed for systems **1a** (benzothiophene), **1b** (dibenzothiophene), **1c** (4,6-dimethylbenzothiophene) and **1d** (4-methylbenzothiophene). The same was done for their dimers and compared against experimental results obtained by IR and Raman spectroscopy.

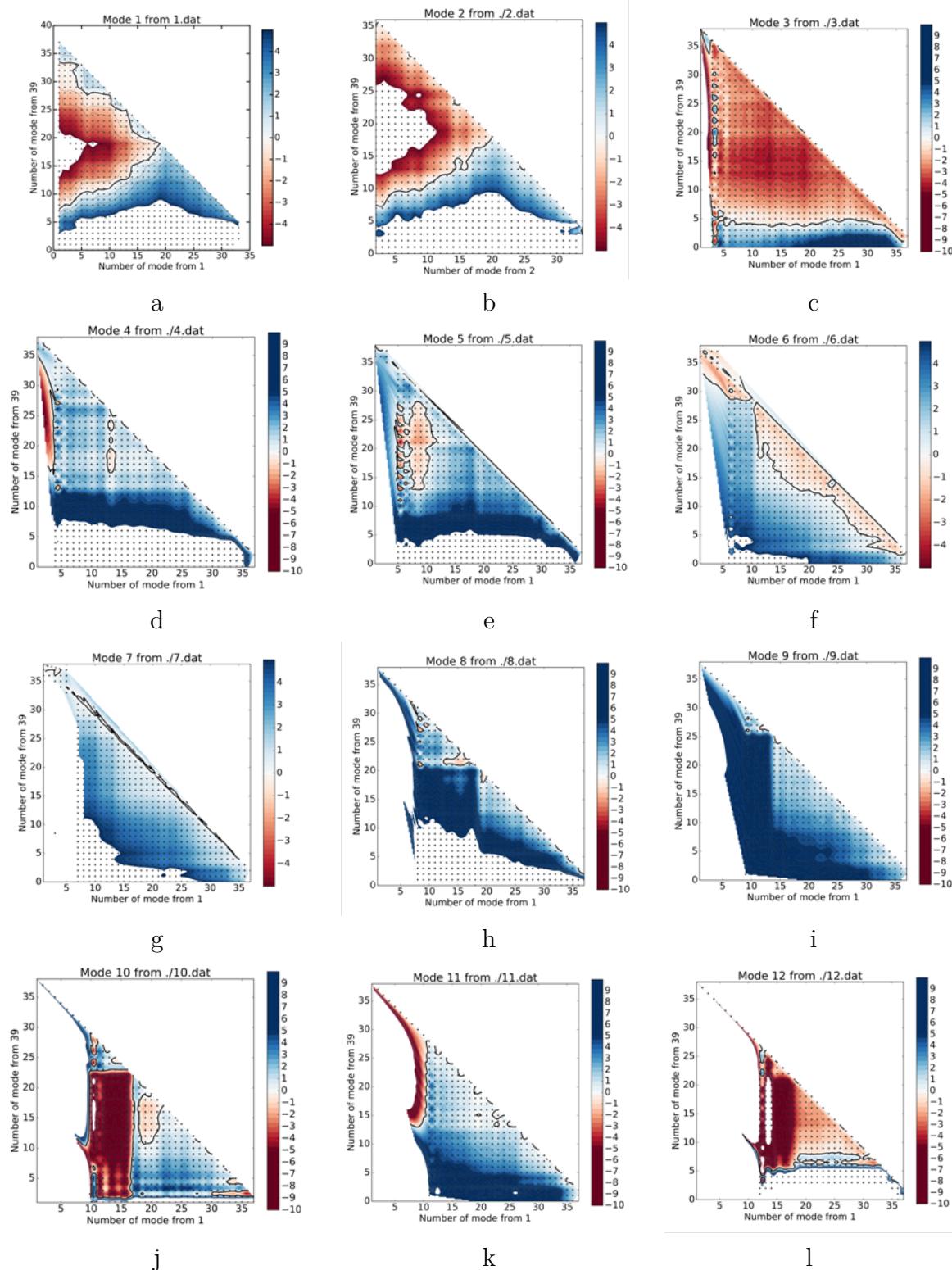


Figure 4.3 – Ensemble of results obtained using truncation n°1 for the first twelve modes of a benzothiophene molecule. Mode 1: Figure a; Mode 2: Figure b ... Mode 12: Figure l. In the z axis are reported the values of the difference obtained on calculating a frequency without any truncation and the same frequency calculated in base  $\mathbf{B}'$  ( $T'=X'+Y'$ ; X' axis x; Y' axis y; in accordance with the notation used in Figure 4.2).

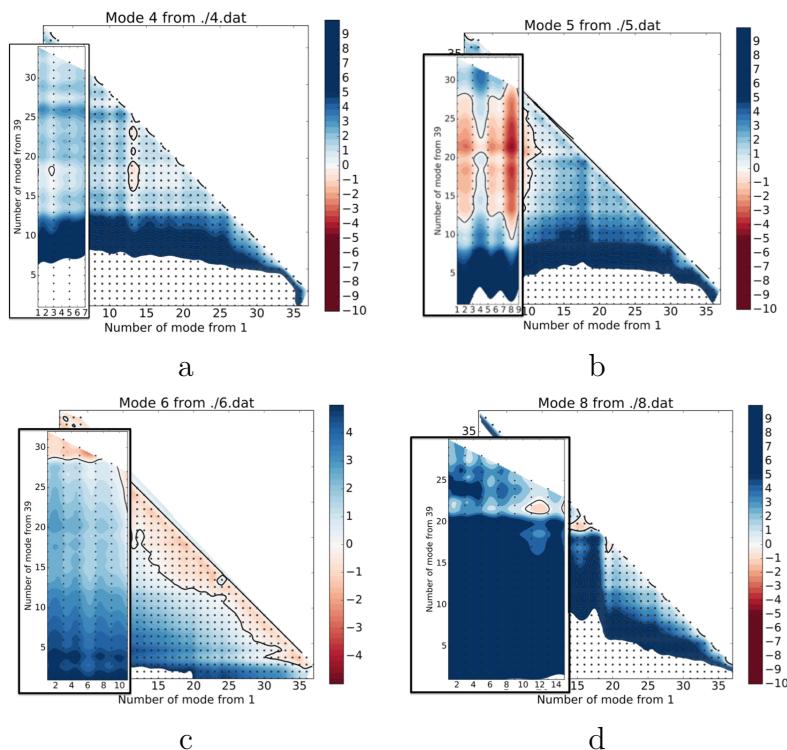


Figure 4.4 – Ensemble of results using the truncation n°2 for the first twelve vibration modes of a benzothiophene molecule. Mode 4: Figure a; Mode 5: Figure b; Mode 6: Figure c; Mode 8: Figure d. In the z axis are reported the values of the difference obtained on calculating a frequency without any truncation and the same frequency calculated in base  $\mathbf{B}''$  ( $T''=X''+Y''$ ; X'' axis  $x$ ; Y''=Y' axis  $y$ ; in accordance with the notation used in Figure 4.2).

PA infrared and Raman spectra of **1b** (dibenzothiophene), **1c** (4,6-dimethyldibenzothiophene) and **1d** (4-methyldibenzothiophene) were interpreted in light of the quantum chemistry calculations. Data for **1a** (benzothiophene) were taken from the literature. The low-wavenumber and “fingerprint” regions (comprising, together, the 20–2000  $\text{cm}^{-1}$  range) were studied in detail in these investigations. In addition to the predicted infrared-active bands, the PA spectra of these hydrocarbons display many features due to overtones and combinations. Moreover, numerous Raman-active (gerade) vibrations also give rise to bands in the PA spectra, which thus convey considerable information regarding the structures of these compounds.

*Low wavenumber region (100–700  $\text{cm}^{-1}$ ):* both Raman and infrared spectroscopy results for **1a** benzothiophene and **1b** dibenzothiophene are reported in Tables 4.1 and 4.2.

For the benzothiophene molecule, the reported Raman spectrum has no significant features in this region. The IR spectrum is dominated by three intense bands at 471, 560 and 690  $\text{cm}^{-1}$  which are characteristic of wagging modes. Two other combination modes are IR-active. They are not easily detected since they have very low intensity and also due to the fact that the  $\nu_2 + \nu_5$  mode has a signature very close to that of the active fundamental mode  $\nu_9$ . The analysis of the active modes of the dibenzothiophene molecule (Figure 4.5) shows that two of the three most intense wagging modes (*obs* 497 and 704  $\text{cm}^{-1}$ ), which are, incidentally,

characteristic of benzothiophene, can be clearly distinguished and do not allow one to easily separate the molecules from each other. However, one should note that a characteristic shift towards higher frequencies is found for these two modes of the dibenzothiophene system. Also note, in comparison to benzothiophene, an extinction (near  $560\text{ cm}^{-1}$ ) of the wagging signature because of symmetry induction and a new intense characteristic vibration located around  $615\text{ cm}^{-1}$ . Moreover, the IR spectrum of this molecule for the region between 100 and  $450\text{ cm}^{-1}$  is more complex and richer in spectral features. As a matter of fact, other stronger and medium intensity modes signify the presence of the second benzyl moiety within the dibenzothiophene molecule. This observation agrees very well with our calculations. The analysis of the eigenfunctions generated by the variational-perturbational process shows that these modes are mainly developed on the basis of vibrational states that are characteristic of torsion ( $\nu_3(a_1)$   $213\text{ cm}^{-1}$ ; *calc*  $217\text{ cm}^{-1}$ ), breathing ( $\nu_4(b_1)$   $228\text{ cm}^{-1}$ ; *calc*  $221\text{ cm}^{-1}$ ) and ( $\nu_6(a_1)$   $404\text{ cm}^{-1}$ ; *calc*  $416\text{ cm}^{-1}$ ) and wagging ( $\nu_8(b_1)$   $425\text{ cm}^{-1}$ ; *calc*  $433\text{ cm}^{-1}$ ) combination modes of the benzenic rings.

The Raman signature of the dibenzothiophene molecule is also more complex than that of benzothiophène. In the present context, this fact introduces only limited information since most of the active Raman modes are already the most intense and characteristic of the IR spectrum. All the other active modes have very low intensities which are again in agreement with our calculations. It is worth noting the confirmation provided by the Raman results concerning the shoulder observed in the IR spectrum around  $419\text{ cm}^{-1}$ . Similarly to the benzothiophene molecule two combination modes are foreseen by the calculations that can be, mainly for the  $\nu_2 + \nu_8$  mode, reasonably assigned experimentally.

In addition to some frequency shifts, the spectra of both 4-methyldibenzothiophene and 4,6-dimethyldibenzothiophene exhibit important changes in the intensities of the vibrations in the  $100 - 500\text{ cm}^{-1}$  region as compared with the spectrum of dibenzothiophene. The clearest change occurs in the  $400-460\text{ cm}^{-1}$  region; the presence of substituents induces a significant change in intensity, ranging from very intense modes down to nearly zero (notably for the  $\nu_{10}$  mode of 4-methyldibenzothiophene). These differences are even more pronounced in the Raman spectra. Whilst the presence of a single methyl group does not significantly change the Raman signature of the dibenzothiophene molecule, a second substitution in the 4,6 position does so very pronouncedly. The main change concerns the most characteristic mode of the di-substitution found at  $598\text{ cm}^{-1}$ , which represents an in-phase breathing mode of the three rings.

Table 4.1 – Low wavenumber Raman and infrared spectra of benzothiophene ( $C_s$  symmetry)

Raman	Observed <sup>b</sup> Infrared	Mode	Calculation $\omega^{a,c}$	$\nu^d$
192 vw	192 vw	$\nu_1 (A'')$	196 (0.22) [0.27]	194 (0.20)
	198-212 w	$\nu_2 (A'')$	204 (4.11) [0.12]	199 (2.97)
340 vw	340 vw	$\nu_3 (A')$	351 (2.82) [2.11]	345 (2.08)
387 vw		$\nu_1 + \nu_2$		391 (0.03)
	405-411 vw	$\nu_4 (A'')$	427 (3.45) [0.77]	414 (2.50)
	471 m	$\nu_5 (A'')$	482 (5.88) [0.53]	478 (6.02)
491 vw	491-497 vw	$\nu_6 (A')$	500 (0.30) [10.64]	494 (0.29)
526 vw		$\nu_7 (A')$	538 (0.25) [6.54]	533 (0.36)
	558-564 m	$\nu_8 (A'')$	578 (5.88) [0.70]	563 (5.24)
	669 vw	$\nu_9 (A')$	683 (2.32) [1.47]	675 (1.70)
		$\nu_2 + \nu_5$		679 (0.10)
	690 s	$\nu_{10} (A'')$	710 (34.70) [0.23]	692 (28.22)

<sup>a</sup> Infrared intensities (in parentheses) in km/mol and Raman intensities [in brackets] in  $\text{A}^4/\text{AMU}$ .    <sup>b</sup> Experimental data. Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder.

<sup>c</sup> Unscaled DFT frequencies calculated using  $\omega\text{B97X-D}/6-311++\text{G}^{**}$ .

<sup>d</sup> Anharmonic  $\omega\text{B97X-D}/6-311++\text{G}^{**}$  calculation using the A-VCI algorithm[71].

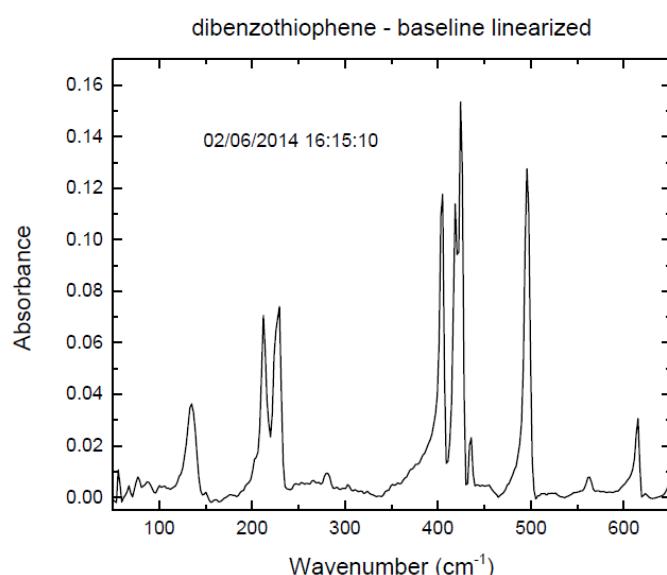


Figure 4.5 – Far-infrared PA spectra of dibenzothiophene

Table 4.2 – Low wavenumber Raman and infrared spectra of dibenzothiophene ( $C_{2v}$  symmetry)

Raman	Observed <sup>b</sup> Infrared	Mode	Calculation $\omega^{a,c}$	$\nu^d$
138 w	135 m	$\nu_1$ (b <sub>1</sub> )	102 (1.54) [0.28]	101 (0.66)
171 w		$\nu_2$ (a <sub>2</sub> )	135 (0.00) [0.07]	133 (0.01)
217 m	213 m	$\nu_3$ (a <sub>1</sub> )	217 (0.85) [1.25]	208 (1.23)
232 w	228 m	$\nu_4$ (b <sub>1</sub> )	221 (1.46) [1.03]	223 (1.42)
	269 vw			
282 w	282 vw	$\nu_5$ (a <sub>2</sub> )	279 (0.00) [2.90]	274 (0.00)
287 sh				
410 s	404 s	$\nu_6$ (a <sub>1</sub> )	416 (2.83) [13.05]	409 (2.80)
	419 sh	$\nu_7$ (b <sub>2</sub> )	428 (0.14) [2.72]	425 (0.15)
419 m	425 s	$\nu_8$ (b <sub>1</sub> )	433 (7.15) [0.01]	425 (5.06)
437 vw	436 w	$\nu_9$ (a <sub>2</sub> )	444 (0.00) [0.06]	438 (0.00)
	460 vw			
496 w	497 s	$\nu_{10}$ (1)	505 (0.28) [4.76]	499 (0.22)
		$\nu_{12}$ (b <sub>2</sub> )	511 (1.23) [4.67]	497 (1.48)
504 vw	512 w	$\nu_{11}$ (b <sub>1</sub> )	510 (2.74) [0.03]	504 (1.18)
	565 w	$\nu_2 + \nu_8$ (b <sub>2</sub> )		562 (0.12)
		$\nu_{13}$ (a <sub>2</sub> )	580 (0.00) [0.24]	568 (0.00)
	615 m	$\nu_{14}$ (b <sub>2</sub> )	635 (5.09) [0.04]	627 (4.62)
	667 w	$\nu_{17}$ (b <sub>1</sub> )	730 (0.82) [0.00]	666 (0.15)
		$\nu_3 + \nu_{10}$ (b <sub>2</sub> )		675 (0.01)
703 s	704 vs	$\nu_{15}$ (b <sub>2</sub> )	724 (6.53) [5.58]	713 (6.43)

<sup>a</sup> Infrared intensities (in parentheses) in km/mol and Raman intensities [in brackets] in  $\text{A}^4/\text{AMU}$ .    <sup>b</sup> Experimental data from this work and ref[182]. Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder.

<sup>c</sup> Unscaled DFT frequencies calculated using  $\omega\text{B97X-D}/6-311++\text{G}^{**}$ .

<sup>d</sup> Anharmonic  $\omega\text{B97X-D}/6-311++\text{G}^{**}$  calculation using the A-VCI algorithm[71].

## Dimers

Characterization of the spectral signatures of “isolated” thiophene molecules constitutes the first part of this study. The second aim of this work lays in the identification and characterization of the IR and Raman signatures of the vibration modes at very low wavenumbers which are characteristics of associated systems such as dimers. In order to foresee and identify the spectral signatures typical of these interactions, we have first identified all the stable forms of benzothiophene (and its derivatives) dimers, *i.e.*, **2a** d-benzothiophene, **2b** d-dibenzothiophene, **2c** d-4-methyldibenzothiophene and **2d** d-4,6-dimethylbenzothiophene. Among all the possible dimer configurations, eight types of benzothiophene dimers have been calculated; these are represented in Figure 4.6. In all tables and figures, the letter *d* means dimer. Our interest lies in  $\pi$ -stacking interactions, hence we have considered the molecules placed such that  $\pi - \pi$  stacking interactions are favored.

Although  $\pi - \pi$  stacking interactions have been proven to be predominant, electrostatic interactions also induce specific configurations [174]. In order to validate the calculation method used, additional calculations using SAPT-DFT have been performed and geometrical parameters as well as interaction energies calculated with the two methods are compared. The energies for the different configurations of the benzothiophene molecule are reported in Table 4.3.

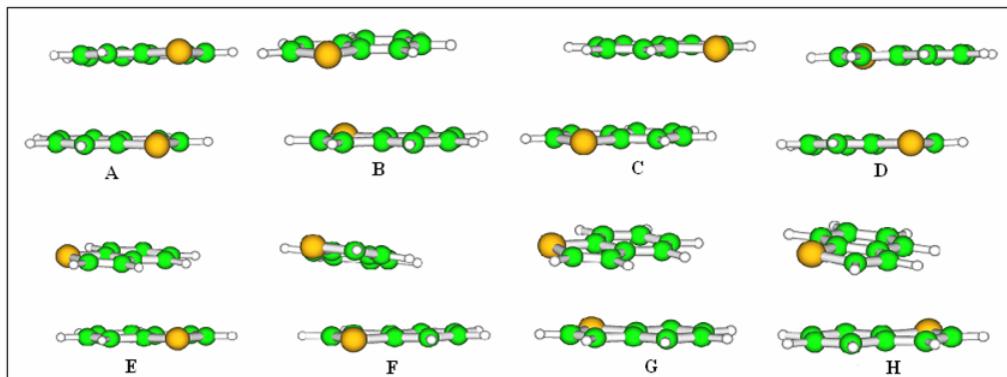


Figure 4.6 – Configurations of benzothiophene dimers optimized at the  $\omega$ B97X-D/6-311++G\*\* level of theory.

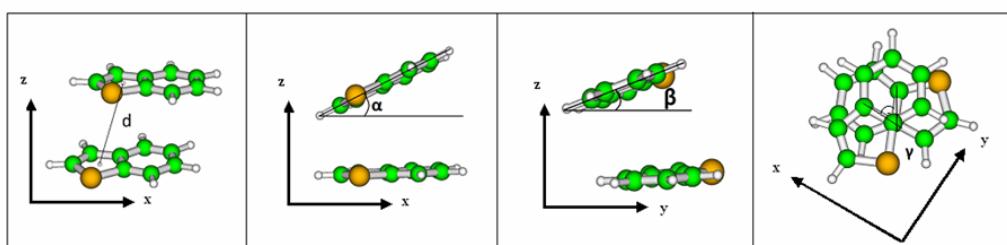


Figure 4.7 – Structural parameters of benzothiophene dimers

$d$  = Intermolecular distance between thiophene rings center of benzothiophene dimers.

$\alpha$  = Angle between the molecules inside  $x, z$  plane.

$\beta$  = Angle between the molecules inside  $y, z$  plane.

$\gamma$  = Angle of rotation in  $z$  axis.

Table 4.3 – (a) Interaction energies of benzothiophene dimers (kcal/mol) in 8 different configurations using  $\omega$ B97X-D/6-311++G\*\*, and compared with other methods. (b) Intermolecular distance between thiophene rings center of benzothiophene dimers (Å) (c) Angle between ring planes of benzothiophene dimers ( $^{\circ}$ ) according to Figure 4.7. (d) Angle between ring planes of benzothiophene dimers ( $^{\circ}$ ) according to Figure 4.7. Energies taking into account zero-point energy (ZPE) and Basis Set Superposition Error (BSSE) corrections.

(a)

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>
$\omega$ B97x-D/ 6-311++G**	-6.41	-7.05	-6.83	-7.11	-7.23	-7.01	-7.02	-7.00
DFT-SAPT(PBE0) aug-cc pVTZ	-5.33	-5.73	-5.56	-5.95	-5.83	-5.87	-5.72	-5.63
B971/ 6-31+G(d,p)-DCP <sup>a</sup>	-5.46	-5.80	-5.72	-5.68	-5.13	-5.40	-5.10	-4.76

(b)

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>
$\omega$ B97x-D/ 6-311++G**	3.94	3.81	5.28	3.72	4.49	3.56	3.74	4.37
DFT-SAPT(PBE0) aug-cc pVTZ	3.98	3.80	5.16	3.75	4.53	3.65	3.70	4.30
B971/ 6-31+G(d,p)-DCP <sup>a</sup>	3.99	3.84	4.65	3.83	5.03	3.66	3.75	3.70

(c)

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>
$\alpha$	2.94	1.73	1.06	0.44	2.23	4.38	2.05	6.37
$\beta$	3.59	4.23	3.14	0.51	5.07	13.28	9.03	5.64
$\gamma$					68.92	100.06	51.71	77.25

(d)

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>
$\alpha$	2.11	1.60	2.36	1.49	3.50	2.66	11.90	7.97
$\beta$	4.00	3.67	2.33	2.00	12.76	2.37	15.79	5.82
$\gamma$					108.52	105.27	107.62	51.44

<sup>a</sup> ref[178]

Eight conformations were highlighted for the benzothiophene dimer species. Benchmark

energy calculations were performed using the DFT-SAPT(PBE0)/aug-cc pvtz level of theory. The  $\omega$ B97X-D/ 6-311++G\*\* calculations, from which the vibrational calculations were performed taking into account both electrical and mechanical anharmonicities, lead to the same conclusions concerning the stability of these eight conformers (average deviation on the order of 1.25 kcal/mol). Regardless of the theory level that is employed, the two most stable conformations are **D** and **E**. In these conformations, the sulfur atoms are opposed to each other. This is also the case for the majority of the other systems studied in this work. The analysis of the energetic contributions linked to the intermolecular association phenomena is reported in Table 4.4.

Before commenting on the results obtained by these contributions, one should recall the basis of the SAPT method used here. The van der Waals interactions, responsible for the cohesion of the dimers reported in this work, are clearly distinct to other types of long-range interactions. In the intermediate region that includes the minimum, it is very difficult to unequivocally decompose the total interaction energy into different contributions. Several decomposition schemes based on the Rayleigh-Schrödinger perturbation theory have been proposed over the years. The principle of these approaches is the separation of the different contributions, such as electrostatic, induction and dispersion. On the other hand, the greatest drawback is the difficulty regarding convergence of the perturbative series (mainly the polarisation part, i.e., induction + dispersion), due to their truncation at low orders or their incompleteness. SAPT has been proposed to address this [114]. In this approach, each of the aforementioned components should also include one exchange correction in order to take the antisymmetric nature of the wave function of the created complex into account.

A simple way to briefly present this theory is to start from a classical partition of a molecular dimer Hamiltonian of two monomers, say A and B:

$$H = H^A + H^B + V \quad (4.1)$$

where  $V$  stands for the intermolecular perturbation.

The total Hamiltonian can be rewritten as:

$$H = H_0^A + W^A + H_0^B + W^B + V \quad (4.2)$$

For which the  $H_0^{A/B}$  terms are the Fock operators of both monomers A and B,  $W^{A/B}$  stands for the monomer's intramolecular perturbation terms and can be understood as the « normal » correlation of each fragment (i.e. monomers A and B Moller-Plesset operators), and  $V$  stands for the monomer's intermolecular perturbation term (i.e. the electrostatic attraction between A and B).

The first perturbation level of order  $N$  is obtained, in the first instance, from not considering the intramolecular correlation. The Rayleigh-Schrödinger perturbation theory provides a series of well-known corrections for the interaction energy [112]. However, it is usual to note that the perturbative series is seldom converged even though high-order terms ( $>2$ ) are employed. Sometimes these series can even diverge. In order to avoid such behaviour, an asymptotic attenuation term is used for the energy in the form:

$$E_{int} = \sum_{n=1}^N E_{pol}^{(n)} + O(1/R^{2(N+1)}) \quad (4.3)$$

Moreover, so that the anti-symmetric principle can be established within the SAPT approach, a classical anti-symmetrization operator is initially applied to the wave function of the systems allowing to establish the expression of the SRS (Symmetrized Rayleigh-Schrödinger - which is the origin of the SAPT name) energy corrected from different perturbation orders up to a given  $n$  large enough to contain all the long-range information.

The second perturbation level of this methodology is obtained considering the intramolecular correlation (once the intermolecular perturbation has been treated accordingly). In this way, each intermolecular perturbation order  $k$  one adds an intramolecular perturbative series respective to monomer A and B (of  $m$  or  $p$  order to develop the  $W^A$  or  $W^B$  operator, respectively). The summation over all the  $m$  and  $p$  orders leads to the following correction formula:

$$E_{pol}^{(k)} = \sum_{m,p} E_{pol}^{(kmp)} \quad (4.4)$$

$$E_{exch}^{(k)} = \sum_{m,p} E_{exch}^{(kmp)} \quad (4.5)$$

which leads to a globally corrected interaction energy:

$$E_{int} = \sum_{k=1}^N \left( \sum_{m,p} E_{pol}^{(kmp)} + \sum_{m,p} E_{exch}^{(kmp)} \right) + O(1/R^{2(N+1)}) \quad (4.6)$$

or even:

$$E_{int} = E_{int(HF,like)} + E_{corr-inter} + E_{corr-intra} \quad (4.7)$$

with the term

$$E_{int(HF,like)} = E_{pol}^{1,0,0} + E_{exch}^{1,0,0} + E_{ind}^{2,0,0} + E_{exch-ind}^{2,0,0} \quad (4.8)$$

representing the term calculated « without correlation » (from where the use of the *HF,like* acronym originates). Moreover, one also has the corrected term to the intermolecular correlation :

$$E_{corr-inter} = E_{disp}^{2,0,0} + E_{exch-disp}^{2,0,0} \quad (4.9)$$

One should note that the SAPT approach used for the study of non-polar complex systems (for which van der Waals forces are predominant) brings up little information and it is unnecessary to develop the perturbation beyond the second order. The so-obtained results (truncated at the order 2) for the four dimers **B**, **D**, **E** and **F** of the benzothiophene system are reported in Table 4.4. The very small numerical differences observed show that the nature of the interaction energies for different conformers does not make it possible to prove the existence of one as opposed to the others. Accordingly one could propose the concomitant presence of all these conformations as representative of inhomogeneous mixtures of asphaltenes, justifying the analysis made herein.

Table 4.4 – Interaction energies components in kcal/mol of benzothiophene dimers using DFT-SAPT(PBE0)/aug-cc pVTZ

	<b>B</b>	<b>D</b>	<b>E</b>	<b>F</b>
$E_{pol}^{1,0,0}$	-3.52	-3.93	-3.48	-4.48
$E_{exch}^{1,0,0}$	10.99	11.04	10.76	10.41
$E_{ind}^{2,0,0}$	-4.65	-4.83	-4.46	-4.69
$E_{exch-ind}^{2,0,0}$	4.38	4.54	4.19	4.38
$E_{disp}^{2,0,0}$	-14.01	-13.78	-13.97	-12.37
$E_{exch-disp}^{2,0,0}$	2.01	2.00	1.97	1.80
$E_{int}(HF,like)$	7.20	6.82	7.01	5.62
$E_{corr-inter}$	-12.00	-11.78	-12.00	-10.57
$E_{corr-intra}$	-0.93	-0.99	-0.84	-0.92
$E_{int}$	-5.73	-5.95	-5.83	-5.87

In this work the Non-Covalent Interactions (NCI) formalism developed by Johnson and coworkers[193] was also used to analyze the energetic differences between all stable dimer forms. This formalism is based on the reduced electronic density gradient ( $s$ ), a fundamental quantity in DFT theory which describes the deviation of the electronic density of a system ( $\rho$ ) from the homogeneous case. Plotting  $s$  versus  $\rho$  reveals the basics of intramolecular interactions: for small  $\rho$  and large  $s$ , one can identify the exponential decay of the density far from the nuclei. For small  $s$  and large  $\rho$ , the points correspond to the covalent bond between the atoms. Situations where  $\rho$  and  $s$  are small are diagnostic for non-covalent interactions. To distinguish between attractive and repulsive non-covalent interactions, one needs to consider the Laplacian of the density ( $\nabla^2\rho$ ).

The components of the decomposition along the three coordinates are the three eigenvalues  $\lambda_i$  of  $\rho$ 's Hessian matrix in a way that  $\nabla^2\rho = \lambda_1 + \lambda_2 + \lambda_3$  is true for  $\lambda_3 \geq \lambda_2 \geq \lambda_1$ . Around the nuclei, there is a maximum for  $\rho$  and all  $\lambda_i$  are negative. For the interatomic regions between bonded atoms, the situation  $\lambda_3 > 0$  and  $\lambda_1$  and  $\lambda_2 < 0$  exists. In this formalism, for covalent interactions, the negative contributions are dominant and the Laplacian is then negative. On the other hand, for non-covalent interactions, the Laplacian in the interatomic region is positive, irrespective of the bonding or non-bonding character of the interaction. Bonding interactions are, in this way, identified by a negative value of  $\lambda_2$  whereas a positive value indicates non-bonding contacts. In other words, the sign of this eigenvalue is used to distinguish bonded  $\lambda_2 < 0$  and non-bonded interactions  $\lambda_2 > 0$ .  $\rho$  itself indicates the strength of the interaction, as can be seen from Figures 4.8a-h.

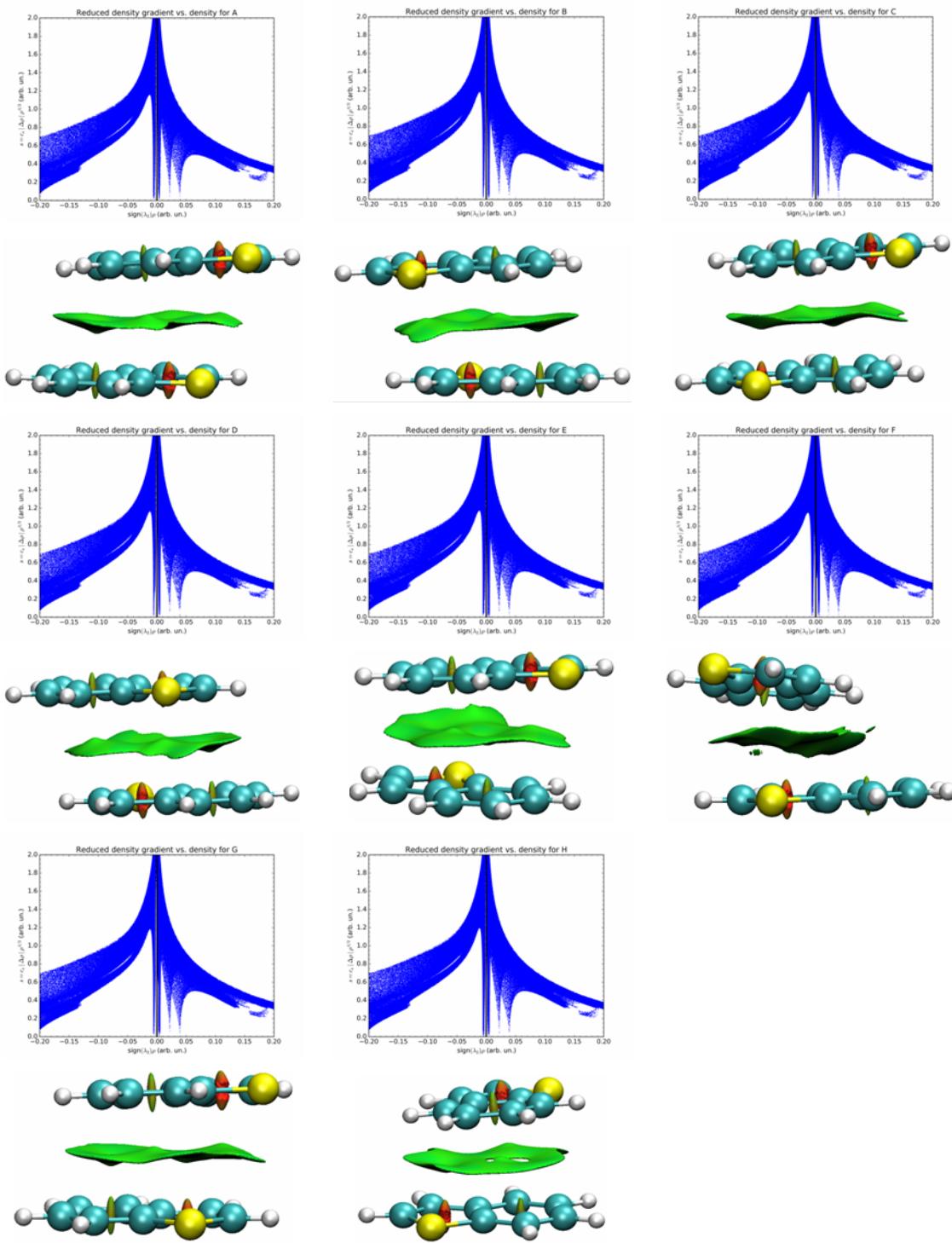


Figure 4.8 – Plots of the electronic reduced gradient density's against against the density  $\rho$  times the sign of  $\lambda_2$

In this Figure, the low density, high gradient regions correspond to the exponential decay of the electronic density away from the nuclei of the molecules. The regions of gradient values around 0.5 arbitrary units and density values around  $\pm 0.2$  au correspond to the covalent interactions: in the positive region this is related to the repulsive character of the covalent

intramolecular bonds and the negative region is correlated with attractive character.

Moreover, the plots of these surfaces allow one to visualize the regions of space where the interactions are stronger. The color mapping indicates their repulsive character (towards red) or attractive (toward green). The intra-ring nuclear repulsion can be clearly distinguished from the interaction surfaces between the molecules of the dimer system in this way.

From these plots, one can deduce that the conformations of the dimers have, to a limited extent, some influence on the non-covalent interactions. Although very little can be said from the reduced density gradient versus the density plots, the non-covalent surfaces show singular features for the studied conformers: conformer **F** has attractive interactions between the hydrogen atoms of one molecule and the carbon atoms of the other while conformer **H** has regions where no interaction takes place or where they mutually cancel (indicated by the holes on the surface). Moreover, analysis of the positioning of the intra-ring nuclear repulsion interactions shows that systems **C**, **D** and **E** have the surfaces of each benzyl or thiophene ring in opposition to each other. This increased interaction between the planes might indicate that the translational (sliding) vibration modes of one plane over the other should be slightly shifted to higher wavenumbers.

#### *Very Low wavenumber region ( 30–100 cm<sup>-1</sup>):*

The first conclusion to be drawn from our calculations for all conformations for all species is that the very low wavenumber features in the spectra for all of the dimers are very similar to each other.

This observation is also consistent with our spectra recorded between 5 and 26 cm<sup>-1</sup> for the dibenzothiophene **1b**, 4,6-dimethylbibenzothiophene **1c** and 4-methyldibenzothiophene **1d** systems (see Figure 4.9). The assignation of the six libration modes of the most stable conformers for the four species (**2a-2d**) is reported in Table 4.5. The parameters of the variation-perturbation calculations used for each species are also reported in this table. These results show that the modes with the lowest wavenumbers, *i.e.*, the sliding mode along the *xy* plane (Figure 4.7) and the rotation around the main axis perpendicular to the plane of the isolated molecules (*y* direction) are mainly coupled; these modes are also coupled with several CH stretching modes. On the other hand, the other libration modes and, particularly, the sixth corresponding to the  $\alpha$ -type rocking mode are coupled, at the same time, to the stretching and wagging  $\omega_{CH}$  modes. This justifies why these modes must imperatively be present in basis **B'** (or **B''**) in which the variation-perturbation work is performed so that convergence can be achieved. The calculated frequencies for each one of the dimers are reported in Table 4.6. Moreover, the results concerning the dibenzothiophene molecule are reported in Table 4.7 and compared to the only experimental data available in the literature and also to our PA measurements. All the results that were gathered from the literature do not lead to the same frequency due to the different methodologies in use but rather to a dispersion (around the frequency found by us). This is unsurprising since the measurement conditions differ among these experiments. The calculations performed for the dibenzothiophene dimer are in good agreement with the ensemble of experimental observations. They confirm that the very low-intensity bands observed experimentally are indeed the signatures of the intermolecular modes. The results in Table 4.6 also show that the rocking modes, found between 60 and 100 cm<sup>-1</sup>, are not strongly dependent on the specie under study, and have more intense

signatures than the libration modes. The rotation and sliding modes are much less intense and their locations are more dependent on the type of molecule. They are also more difficult to characterize in all types of spectroscopy: however, we have recently identified weak infrared absorption bands that may correspond to the sliding mode in spectra acquired using coherent synchrotron radiation (Figure 4.9). Finally, the presence of one or several substituents at different positions on the molecule can also change the intensity of the signatures of the libration modes significantly, particularly for those implying, through vibration, the movement of the substituents towards each other, which gives rise to new intermolecular interactions.

Previously, our group showed the extent to which these interactions play a significant role in the aggregation of molecules constituting the asphaltenes and the importance of characterizing them [194]. Among these interactions, a non-negligible part consists in intermolecular interactions. Characterizing them through spectroscopy is then crucial for understanding the aggregation phenomena of asphaltene mixtures and, thus, estimating the energies required to separate them during the refining process and avoid the asphaltene sedimentation phenomenon widely observed in the pipelines. Our work using molecular dynamics sheds some light on the nature of the intermolecular interactions that can take place in the mixture for different heteroatoms, different lateral chains, etc. Among our future aims is the identification of characteristic features in vibrational spectra which permit identification of each interacting species for different heteroatoms and lateral chains that may be present in these aggregated systems.

In summary, the present work shows that: 1) solving the vibrational Schrödinger equation for isolated molecules allows one to identify, for molecules containing the sulfur heteroatom, the vibration modes for different molecular moieties as well as their singular spectroscopic properties; 2) utilizing both IR and Raman spectroscopy is very useful towards the full identification of these modes; 3) variation-perturbation calculations in the 100-700 cm<sup>-1</sup> region make it possible to complete the assignment of the modes of each monomer (which was previously initiated in the literature); 4) in the same region, vibrational modes of the dimers correspond to those of the isolated molecules making it possible to determine based solely on these modes, whether the molecules are interacting or not; 5) these intermolecular interaction modes are expected only at very low wavenumbers, and 6) in the region below 100 cm<sup>-1</sup>, each molecule containing different configurations has characteristic signatures in the vibrational spectrum: these signatures arise from the sliding or rotation modes of one molecular plane with respect to another and can be in- or out-of-phase. Even though no general rule could be established concerning the behavior of these modes for different species or configurations, it is clear that their possible existence can serve as a marker for the presence or absence of intermolecular aggregation.

Additionally use of the NCI formalism on a wide range of molecules containing different substituents and different heteroatoms should enable us to characterize the nature of these interactions in the future, and to quantify the energies involved in these interactions. This will make it possible to predict the behaviors for different molecular architectures and heteroatom arrangements.

Finally, the description of the intermolecular modes chosen in this work is a first step of a project aimed at fully characterizing the aggregation interactions. In future work, we plan to a)

perform solid state calculations, which are closer to actual experimental conditions since these interaction modes are linked to the deformations of the crystalline network of the solid (related results were already published by our group concerning the vibrational modes near  $100\text{ cm}^{-1}$  for a family of acenes); and b) study families of molecules with different chemical substituents and different heteroatoms (N, O and S).

Table 4.5 – Parameters used in the variation-perturbation process (values of X' and Y as defined in Figure 4.2 parameters and dimension of the variational problem  $\mathbf{B}''$ ) for each intermolecular mode of the four most stable dimers **2a-2d**

Mode	Benzothiophene			Dibenzothiophene			4-methyl dibenzothiophene			4,6-dimethyl dibenzothiophene		
	type	X" ; Y" ; $\mathbf{B}''$	type	X" ; Y" ; $\mathbf{B}''$	type	X" ; Y" ; $\mathbf{B}''$	type	X" ; Y" ; $\mathbf{B}''$	type	X" ; Y" ; $\mathbf{B}''$	type	X" ; Y" ; $\mathbf{B}''$
1	x,y	1 ; 7 ; 2500	$\gamma$	1 ; 10 ; 3200	$\gamma$	1 ; 14 ; 8100	$\gamma$	1 ; 17 ; 10300	$\gamma$	1 ; 17 ; 10300	$\gamma$	1 ; 17 ; 10300
2	$\gamma$	3 ; 7 ; 2500	x,y	3 ; 10 ; 4100	x,y	3 ; 14 ; 9000	x,y	3 ; 17 ; 15100	x,y	3 ; 17 ; 15100	x,y	3 ; 17 ; 15100
3	x,y	5 ; 7 ; 3000	x,y	5 ; 10 ; 4500	x,y	5 ; 14 ; 10500	x,y	5 ; 17 ; 16600	x,y	5 ; 17 ; 16600	x,y	5 ; 17 ; 16600
4	$\alpha$	7 ; 12 ; 3600	$\alpha$	7 ; 15 ; 8800	$\alpha$	7 ; 20 ; 12700	$\alpha$	7 ; 23 ; 20500	$\alpha$	7 ; 23 ; 20500	$\alpha$	7 ; 23 ; 20500
5	$\beta$	3 ; 12 ; 3500	$\beta$	3 ; 15 ; 6300	$\beta$	3 ; 20 ; 9900	$\beta$	3 ; 23 ; 17400	$\beta$	3 ; 23 ; 17400	$\beta$	3 ; 23 ; 17400
6	$\alpha$	9 ; 12 ; 4200	$\alpha + \omega$	11 ; 15 ; 15000	$\alpha + \omega$	11 ; 20 ; 17200	$\alpha + \omega$	11 ; 23 ; 24900	$\alpha + \omega$	11 ; 23 ; 24900	$\alpha + \omega$	11 ; 23 ; 24900

Table 4.6 – calculated intermolecular vibrations modes of the four most stable dimers **2a-2d**

Mode	Benzothiophene		Dibenzothiophene		4-methyl dibenzothiophene		4,6-dimethyl dibenzothiophene	
	$\omega^a$ (I) <sup>b</sup>	$\nu^a$	$\omega^a$ (I) <sup>b</sup>	$\nu^a$	$\omega^a$ (I) <sup>b</sup>	$\nu^a$	$\omega^a$ (I) <sup>b</sup>	$\nu^a$
1	17 (0.02)		18 (0.05)		22 (0.01)		17 (0.00)	
2	18 (0.08)		28 (0.01)	24	36 (0.17)	32	22 (0.01)	
3	45 (0.06)	39	36 (0.09)	35	37 (0.01)	31	41 (0.01)	38
4	67 (0.32)	58	70 (0.32)	66	65 (0.07)	55	74 (0.52)	66
5	74 (0.02)	67	89 (0.01)	82	81 (0.03)	67	83 (0.07)	69
6	91 (0.06)	88	93 (0.56)	92	96 (0.06)	85	84 (0.01)	78

<sup>a</sup> :  $\text{cm}^{-1}$     <sup>b</sup> :  $\text{km/mol}$

Table 4.7 – Experimental and calculated intermolecular vibration modes ( $\text{cm}^{-1}$ ) of dibenzothiophene dimer (**B**)

	$\omega\text{B97x-D/6-311++G**}$	$\text{IR}^a$	$\text{Raman}^a$	$\text{IR}^b$
Mode	$\omega$ (I) <sup>c</sup> [ $I^{\text{Raman}}$ ] <sup>d</sup>	$\nu$	$\nu$	$\nu$
1	18 (0.05) [0.62]			
2	28 (0.01) [7.96]	24		
3	36 (0.09) [9.15]	35	53 w	48 - 49
4	70 (0.32) [0.51]	66	72 w	63
5	89 (0.01) [0.51]	82	89 vw	79 - 82
6	93 (0.56) [3.33]	92	104 - 101 w 109 w	103 - 106 126

<sup>a</sup> : liquid/crystal ref[182].    <sup>b</sup> : this work (crystal)    <sup>c</sup> : km/mol    <sup>d</sup> :  $\text{A}^4/\text{AMU}$

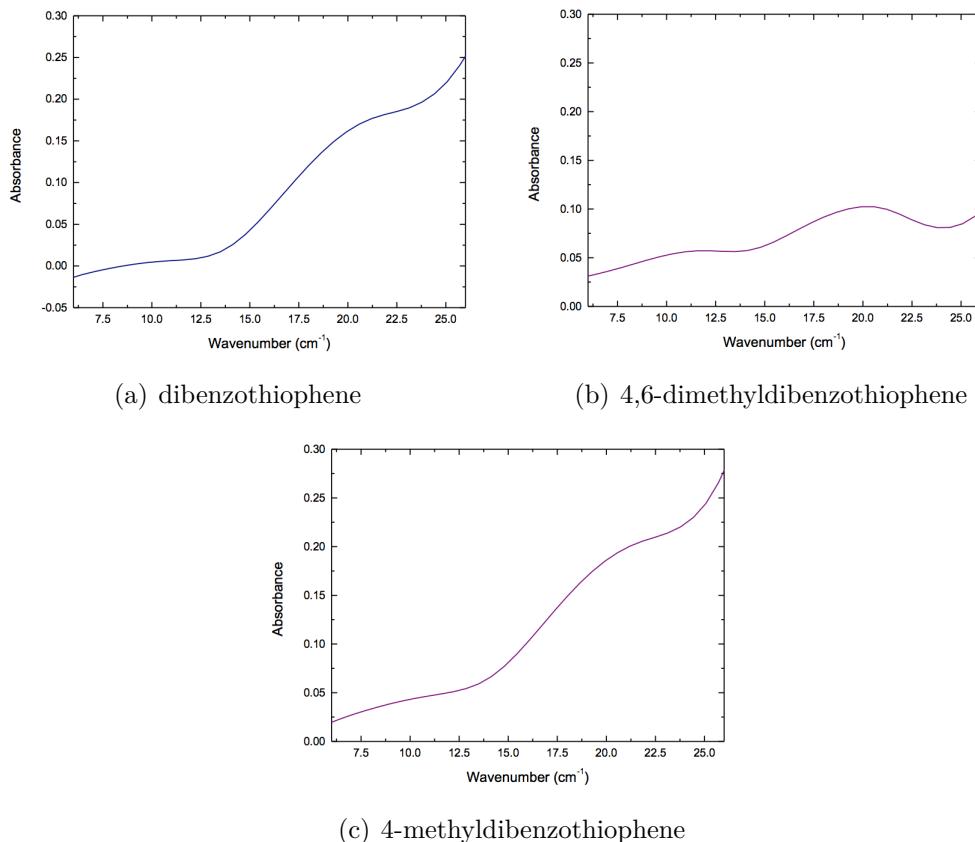


Figure 4.9 – Far-infrared spectra of dibenzothiophene **1b**, 4,6-dimethylbenzothiophene **1c** and 4-methylbenzothiophene **1d**. Data were acquired using coherent synchrotron radiation and a Bruker IFS 125HR FT-IR spectrometer.

# **Chapter 5**

## **Heteroatoms**

## Introduction

The chemical complexity of asphaltenes, which represent the heaviest fraction of oil, make their analysis and modelling very challenging, since the detailed molecular composition remains up to now unknown. However, some studies have reported that they consist of a heterogeneous mixture of polycondensed molecules, containing heteroatoms  $X$  such as sulfur, oxygen, and nitrogen. On the other hand, vibrational spectroscopy has shown to be a tool for the identification and characterization of molecules in ill-defined mixtures, in combination with predictive modelling.

Considering this, in this chapter, the far- and mid-infrared spectra of a series of heteroaromatic (N, O and S) hydrocarbons and their dimers have been calculated using Density Functional Theory (DFT) within the  $\omega$ B97X-D/6-311++G\*\* level of theory. The perturbational-variational method presented in the previous chapter coupled with potential truncation was incorporated in order to provide anharmonic corrections of monomers. Identification and quantification of peaks from *inter* and *intra*-molecular vibrations in experimental spectra have been performed. Also, experimental IR bands in the far-IR region were identified, allowing the differentiation of the *inter*-molecular and *intra*-molecular vibrations. In addition, the interaction energies were studied using DFT-SAPT<sup>1</sup>

These methods reproduce energies in the same order of magnitude and identify  $\pi - \pi$  stacking as the dominant electronic interaction. Understanding the interaction between these primary units helped us to enrich the knowledge of the *inter* and *intramolecular* interactions in asphaltenes and why they tend to aggregate and then flocculate from the oil condensed phase. This outcome illustrates that the spectral signatures of heteroaromatic compounds can be used to probe the molecular and sub-molecular composition, and the *inter*-molecular interactions present in asphaltenes by spectral decomposition.

## 5.1 Structural details

The molecular structures of thiophene, benzofuran, carbazole and fluorene-like molecules and their respective dimers were calculated using Density Functional Theory (DFT) with the  $\omega$ B97X-D exchange-correlation functional and the 6-311++G\*\* basis set, implemented in the Gaussian 09 suite of programs. The  $\omega$ B97X-D functional developed by Chai *et al*[184] is based on optimized long-range corrected hybrid density functionals, which employ 100% Hartree-Fock (HF) exchange for long-range electron-electron interaction correction. This method, tested by Salzner *et al*[185] for  $\pi$ -conjugated oligomers, was also benchmarked successfully for interacting polycondensed aromatic molecules,[49] thus proving it overcomes the failure of standard DFT calculation approaches to describe long-distance interactions. The 6-311++G\*\* basis set is of triple-zeta quality for the valence electrons. Among the basis sets tested by Wiberg,[186] including the Dunning correlation basis sets such as aug-cc-pVTZ and aug-cc-pVQZ, the 6-311++G\*\* basis set gave accurate geometries and frequencies at a relatively small computational cost. The geometrical parameters of the molecules are

---

<sup>1</sup>The symmetry-adapted perturbation theory based on density functional theory) at the PBE0/aug-cc-pVTZ level of theory.

benchmarked with experimental data when available.

For dimers, as an alternative to the supermolecular treatment, an approach based on symmetry-adapted perturbation theory (SAPT)[114] which utilizes the description of the interacting monomers in terms of Kohn-Sham (KS) orbitals was used in this work. The DFT-based SAPT approach,[187] denoted herein as SAPT-DFT, is exact for all major components of the interaction energy (asymptotically for exchange interactions) in the sense that these energies would be exact if the DFT description of the monomers were exact. The SAPT-DFT calculations were performed using the MOLPRO2012 package.[189] The PBE0 functional[134] was used as a monomer DFT functional in SAPT, using the aug-cc-pVTZ basis set.

## 5.2 Results and Discussion

### 5.2.1 Global analysis

The methodology behind this part of the work is close to that of the ‘design of experiments’ [195]. Indeed, studying all the moieties, all the heteroatoms, all the substituents as well as all the allowed *inter*-molecular associations of the hypothetical constituents of the asphaltenes (for which, *cf.* chapter 1, both the composition and formulation are unknown) is utopic and unrealistic. The general principles behind the approaches of the type “design of experiments” rely on the organisation and limitation of the experiments/calculations when a link between a response function  $y$  and variables  $x_i$  (designated a ‘factor’) is sought after. Within the context of this work, the searched response function is knowing whether or not, for any given compound constituted of a given moiety, presenting or not an heteroatom, interacting or not among them, owns an active vibrational signature. These three criteria represent the three principle factors taken into account in our calculations. Unluckily, our problem is more complex since, besides being active, this vibrational mode should also be identifiable within a spectral zone where the density of modes is weak enough to avoid hiding the signature that is looked for. The notion of spectral zone is the last factor that has been considered in this study.

Actually, we also could have multiplied the chemical moieties *ad infinitum* and accumulate more and more vibrational information for these systems, but it would be very complicated and hard to manage and explore all these data. We have chosen to concentrate on a strict domain of definition for which the main pieces of information are controlled so that useful information can be provided to spectroscopists: “is it possible for a given asphaltene to use IR and Raman to characterise the presence of associated systems?” and, “is it possible to assign these signatures to the presence of the three main heteroatoms (oxygen, sulphur and nitrogen)?”

Infrared and Raman spectra of indene **1a**, fluorene **1b**, benzofluorene **1c**, cyclopentaphenanthrene **1d**, 1-methylfluorene **1e**, 1,8-dimethylfluorene **1f**, indole **2a**, carbazole **2b**, benzocarbazole **2c**, 4,5-iminophenanthrene **2d**, 1-methylcarbazole **1e**, 1,8-dimethylcarbazole **1f**, benzofuran **3a**, dibenzofuran **3b**, benzonaphofuran **3c**, tribenzofuran **3d**, 4-methyldibenzofuran **3e**, 1,8-dimethyldibenzofuran **3f**, benzothiophene **4a**, dibenzothiophene **4b**, benzonaphthothiophene **4c**, tribenzothiophene **4d**, 4-methyldibenzothiophene **4e** and 4,6-dimethyldibenzothiophene **4f** were studied in this work. In complement, the observed bands of the monomers are assigned using high level calculations based on anharmonic electrical and mechanical approximations. All computational outcomes are reported in the annexe C.

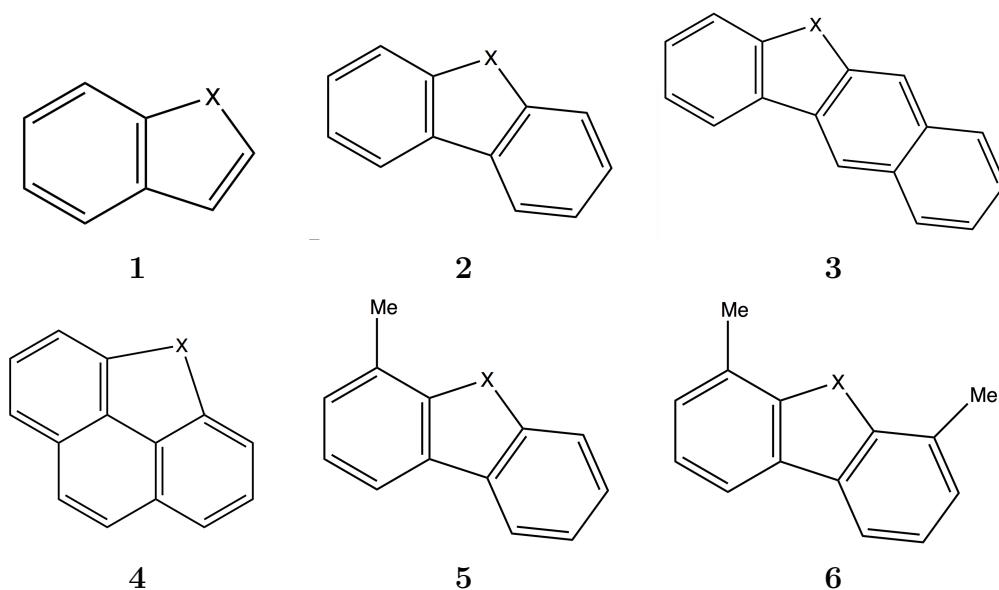


Figure 5.1 – Molecular structures studied: **a** Benzo, **b** Dibenzo, **c** Benzonaphtho, **d** Tribenzo, **e** methyldibenzo, **f** dimethyldibenzo. X= **1** C, **2** N, **3** O and **4** S.

The two first factors we studied were: 1- the nature of the heteroatom (factor  $x_1$ ), for which the bounds were arbitrarily set to  $x_{1(min)} = C$  and  $x_{1(max)} = S$ ; 2- the nature of the chemical moiety, which is discrete and not easily quantifiable ( $x_{2(min)} = \text{benzo}$  and  $x_{2(max)} = \text{tribenzo}$ ). The spectral zone is then a third factor to be considered and which definition is also discrete. Four spectral zones were defined: the very low wavenumbers one ( $x_{3(min)} = 1$ ), which displays the signatures of the *inter*-molecular modes (represented on the figures by the “+” symbol), butterfly-like (“○”) and torsion (“△ and ◇”) modes, which are particularly active in this region; another one ( $x_3 = 2$ ), covering the signatures of the cycle breathing and scissoring (“◀”), twist (“▽”) and wagging modes that are impacted by the presence of the hetero atom (pure: ◇, X-H: ○, wagging/carbon backbone wagging: □);  $x_3 = 3$  covering the C-X (○) stretching, the CCC and CXC angular deformations (★) modes of the aromatic cycles and the C-H wagging modes ones; and, finally, the  $x_3 = 4$  zone covering carbon-carbon and C-H stretching modes found beyond  $1300\text{ cm}^{-1}$ . A table summarizing the modes that are followed in this study as well as their respective symbols is presented in Table 5.1. The modes of interest and which are represented in our Figures by the use of pictograms are the ones for which the hetero atom contribution is strong for the ensemble of the 24 studied molecules, in average. The last factor considered in this work concerns the presence ( $x_{4(min)} = \text{yes}$ ) or not ( $x_{4(max)} = \text{no}$ ) of *inter*-molecular interactions by analysing the dimers and monomers, respectively. However, one should note that this chapter only reports the results of the work done in the molecular state for which the only aim is defining if it is relevant to consider one or several spectra signatures (*intra*- and *inter*-molecular) that could allow one to characterise the heteroatoms present in the asphaltene mixtures and their associated vibrational modes. In chapter 6, the description of the *inter*-molecular modes will be discussed in more details by the development of solid state calculations taking into account this time of the global environment of the studied systems.

The implementation of an “ordered” strategy able to take into account the characteristics discriminating spectroscopically all the systems is justified when one take a look at Figure 5.2.

Table 5.1 – Symbols and assigment of normal modes studied in vibrational spectra calculations

Symbol	Assigment Mode	Symbol	Assigment Mode
+	Intermolecular		
○	Butterfly	▽	Twisting
△	Torsion	◇	X-H Wagging pure
◊	Torsion	○	X-H Wagging
◀	Cycle Breathing / scissoring	□	Carbon Backbone Wagging
◇	C-X Streching	*	CCC and CXC Angular deformation

In this figure, the ensemble of the calculated modes of the 24-monomer systems is reported without distinction neither of the spectral zone nor the intensity. We can easily realise the extent of the data available to treat this problem. Only the modes for which the heteroatom is directly implied in the development of a given vibration are reported in Figure 5.2 as colored points.

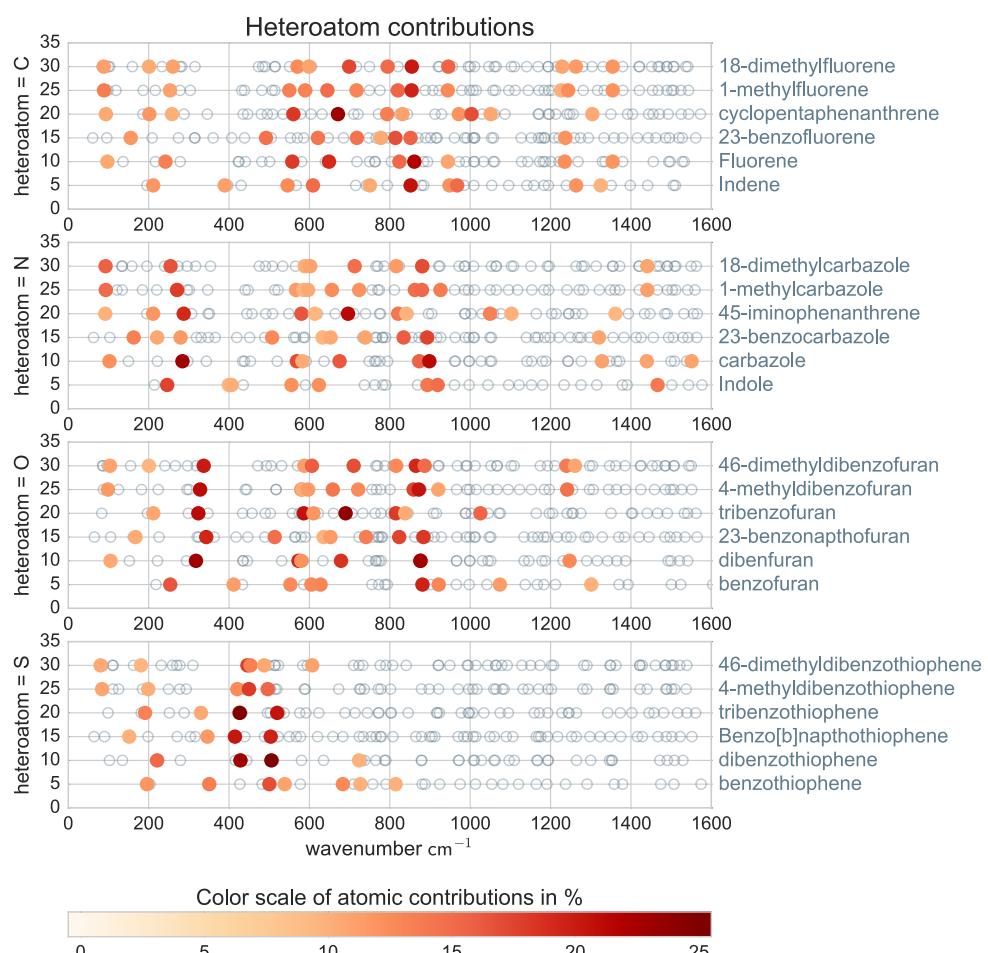


Figure 5.2 – Vibration modes heteroatoms contributions &gt; 5% for studied molecules.

Several pieces of informations can be extracted from this figure. First of all, the number of modes having an influence of the heteroatom  $X = C, N$  and  $O$  is of the order of 50% of all the modes, regardless of the molecule. This percentage is reduced by half for the molecules of the thiophene family. The contributing modes are, in this case, concentrated in the low wavenumber spectral zones ( $x_3 = 1$  and 2 – what is due to a simple mass effect). This first analysis also shows that within the three first spectral zones localised below  $1300\text{ cm}^{-1}$  there are much more similarities between the signatures of  $x_1 = N$  and  $O$  heteroatoms than there are between  $S$  and  $C$ . Lastly, and except for sulfur, it is clear that below  $1000\text{ cm}^{-1}$  the signature of each moiety is singular. This observation is particularly pronounced in the lowest spectral zone  $x_3 = 1$ , where the torsion and deformation modes of the backbone can be found (including, this time, the thiophenic molecules). The moiety seems to be, at this stage, a major discriminatory criterion of the spectral signature of these molecules.

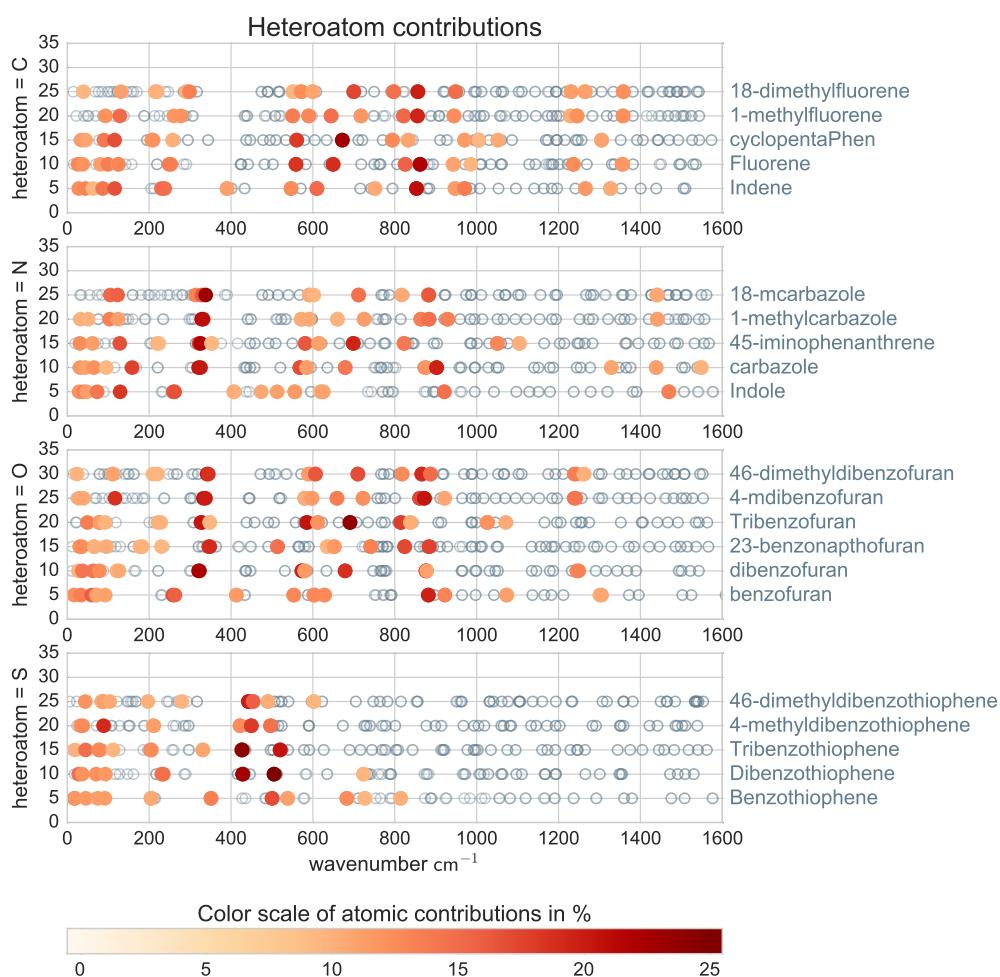


Figure 5.3 – Vibration modes heteroatoms contributions  $> 5\%$  for dimers configurations

In figure 5.3, the same information reported in figure 5.2 is presented for the ensemble of the most stable dimer of each specie. Figure 5.2 and 5.3 look alike apart from the emergence of *inter*-molecular modes at very low wavenumbers for which the contribution of the hetero atoms are strong. Another remarkable difference concerns the nitrogen heteroatom since it seems that it is the only one with a density and rate of contributing modes evolving. This is particularly

true for the zones  $x_3 = 1$  around  $150 \text{ cm}^{-1}$  and  $x_3 = 2$  around  $400\text{-}500 \text{ cm}^{-1}$ . No other significant difference can be highlighted from figure 5.3. This corroborates the idea stating that the signatures of the *intra-molecular* modes specific to each molecule constituting the dimers are, except these two precedent observations, the very same reported for the monomers. Another difference that can be noted is the spectral behaviour of the 24 molecules within the study domain depicted in Figure 5.4. In this figure the spectral signatures taking into account the intensities and filtrated by each moiety are reported. For each one of them, one can observe again at the lowest frequencies a singular behaviour of the signatures associated to the molecules of the carbazol family. The modes associated seem to own more intense signature for which the position seems to depend on the moiety more than for any other heteroatom.

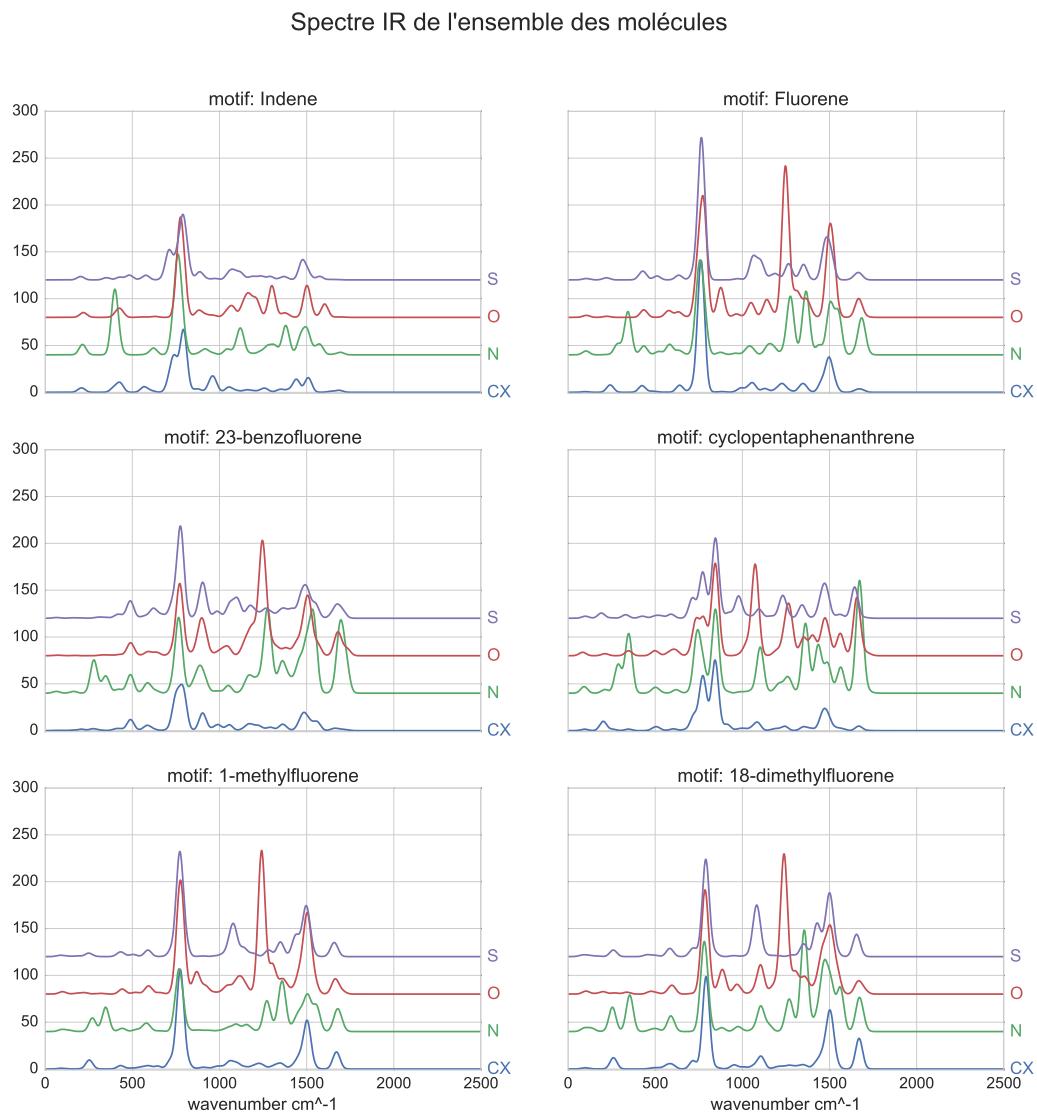


Figure 5.4 – Calculated vibrational spectra for all molecules, motifs and heteroatoms comparisons.

Another information concerns now the thiophene family of molecule. If in figure 5.2 and 5.3 one could verify a non-negligible translation of contributing modes towards the low-wavenumber zone, figure 5.4 shows that, in the  $700\text{-}1600 \text{ cm}^{-1}$  (sparsely populated, though) the activity of

the remaining modes have not (or very little) lowered in comparison to the equivalent signatures observed for the other hetero atoms (for those, there are no sparsely populated region). The reason behind it is very simple: the majority of the remaining modes imply vibrations of the carbonic backbone that are not very dependant on the hetero atom and are common to all moieties. In this way, it is clear that this spectral zone is of no (or a little) interest to discriminate the spectral behaviours of the different systems. The  $x_3 = 3$  zone is also not very well suited to the discrimination of the vibrators, since it concentrates the most active modes (common to all systems) which the amplitude of the signals can hide all the other signatures. Finally, zones  $x_3 = 1$  and 2 seem to be more adapted to this discrimination. A finer analysis, mode per mode of these two spectral zones is still necessary in order to identify the singularities of each moiety and each hetero atom, what is reported hereafter.

### 5.2.2 Analyse

At this point, we should remember that such analysis has a sense only if one tries to product a global information of the spectral behaviour of the molecules (or moieties) potentially present in asphaltenes and, particularly, the signatures of *inter*-molecular signatures on the basis of the samples of chosen model molecules. Of course, if we had the exact formulation of the molecules present in the asphaltenic medium available, it would suffice to model the spectrum of each molecule, for every single moiety and every single heteroatom in order to find, for each molecule, one or several singular characteristics able to discriminate it among the others. Employing mathematical methods dedicated to the resolution of the vibrational Schrödinger equation such the ones already presented in chapter 4 (developed within the anharmonic hypothesis) would be enough to provide to spectroscopists useful information (fundamental bands, combination bands, overtones, medium and near IR, Raman, ...) The full description of the 24 molecules studied in this work can be found in the Appendix of the manuscript.

The afore reported analysis allows one to bring to light the singular behaviours linked to the presence of heteroatoms and the associated molecules. Despite the fact that this analysis is qualitative and does not allow to conclude concerning the role of each factor. A more detailed analysis of the different modes implicated in each molecule was then necessary to address this problem. In view of the very specific role that the nitrogen hetero atom seem to play, a first detailed study of the carbazole molecule is presented, for which experimental data is available. This will be followed by a second one performed on the dibenzothiophene molecule for which exhaustive experimental data is also available. We should finish by resuming the ensemble of data for each one of the 24 molecules studied in this work, focusing on the differences observed in the  $x_3 = 1$  and 2 spectral zones (which were previously identified as the one allowing one to discriminate plus efficiently the role of each moiety and heteroatom on the association process).

## 5.3 Carbazole

The molecules of the carbazole family are the ones for which the heteroatom contribution to the vibrational modes are the strongest. Another particularity of this family is that the most intense contributions can be found in the lower zones of the spectra. The carbazole monomer presented in figure 5.5 depicts these observations.

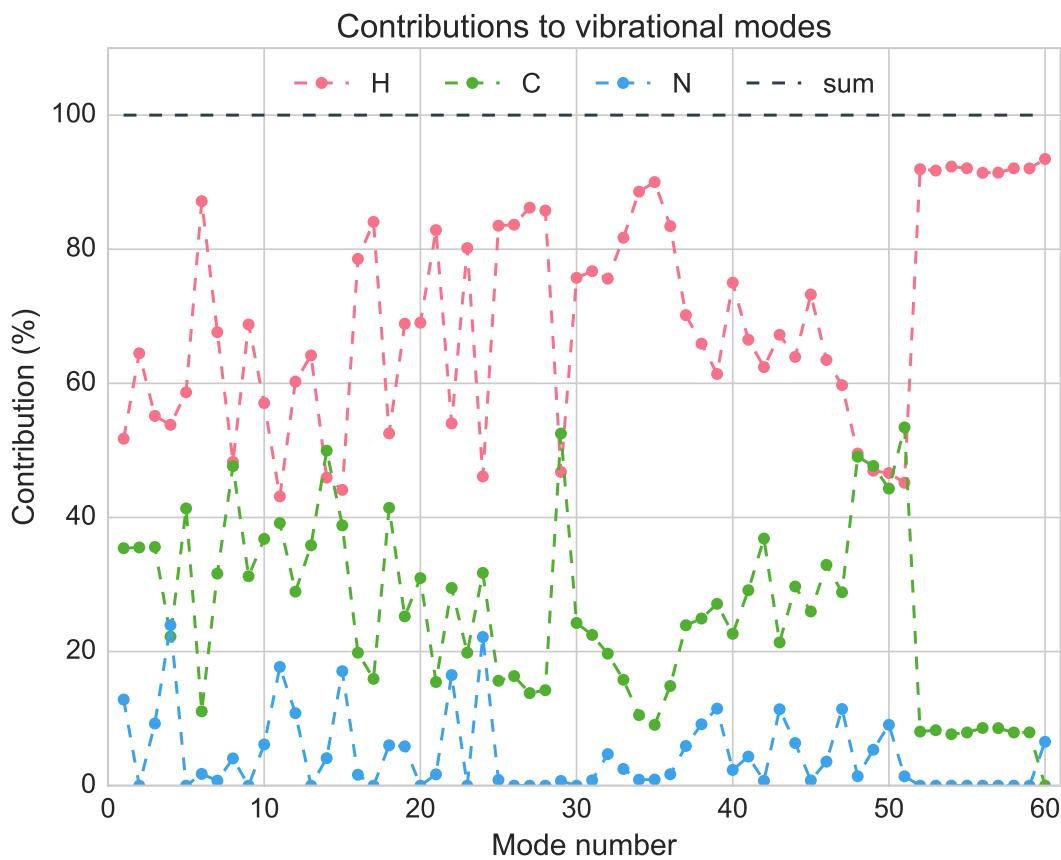


Figure 5.5 – Atomic percentage contributions to vibrational modes in Carbazole molecule.

Similarly, it becomes clear from this picture that the variations of heteroatom contributions spread over the three spectral zones  $x_3 = 1 - 3$  and are, exception made to mode 4 ( $\omega_{N-H}$ ), perfectly correlated to the contributions linked to the case where carbon is the heteroatom. These three spectral zones, distributed over around  $900\text{ cm}^{-1}$ , characterise the ensemble of the signatures of the 25 first modes of carbazole. Among these 25 modes, those for which the contribution of the heteroatom is the most important are reported in figure 5.6 by the aforementioned pyctograms. One can note that, except the C-X stretching mode ( $\square$ ), all the modes depicted by pictograms can be found for carbazole. This information is not enough to allow a rigorous analysis of carbazole's spectrum. To do so, it is also necessary to take into account at the same time the intensity of each mode and having in hand simultaneously the equivalent information for the dimer molecules if one wishes to characterise the signatures of the *inter*-molecular modes. The ensemble of these pieces of information is reported in Figure 5.7.

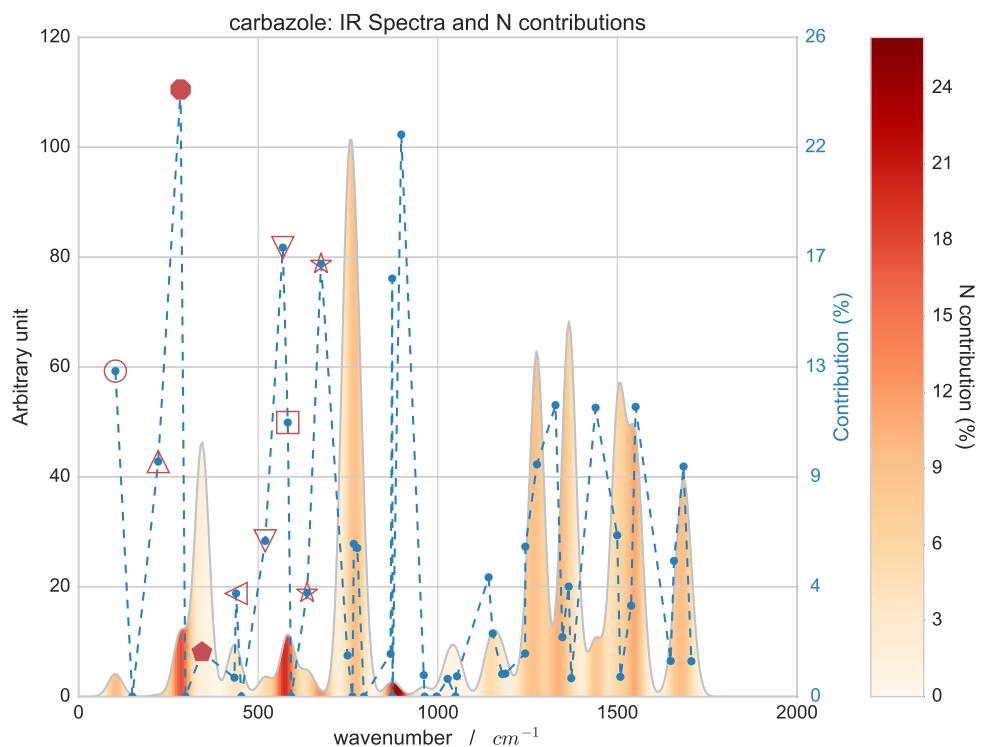


Figure 5.6 – Calculated vibrational spectra of Carbazole molecule, showing heteroatom contribution and selected modes.

The information present in this picture reveals several distinct signs of monomeric and associated carbazole. 1 – a first category of modes seems to distinguish themselves from the others, i.e., those which have a strong contribution of the hetero atom and a strong intensity at the same time. For instance, carbazole's first modes to own the two characteristics are the twist ( $\nabla$ ) and wagging ( $\square$ ) ones, found between  $550$  and  $600\text{ cm}^{-1}$ . Even if the response of these modes seems to be the most accessible of a spectroscopic point of view, we also could realise that these characteristics are also identically present for the simulations of carbazole dimers. Because of these strong similarities, these modes are thus not useful to our analysis. 2- the presence of *inter-molecular* modes is, on the other hand, more useful. This has been brought to light by our calculations based on the appearing of low intensity signatures at very low wavenumbers which positioning, at least apparently, agrees well with the experimental data reported by K. Michaelin and J. Shaw (figure 5.8). We shall present in this chapter only the results of these specific modes. The quantical treatment of these *inter-molecular* modes will be the subject of development in chapter 6 of this work. However, we should note that these modes are particularly coupled with the others of the deformation type, such as the “butterfly” mode ( $\circlearrowright$ ). Consequently, these couplings induce the presence of distinctive differences between the signatures of the monomers and of the dimers within the region nearby the one associated to the *inter-molecular* modes, as it is depicted in Figure 5.8. This is then a zone of interest to us that shall need a very precise description when we will model the vibrational signatures of the systems at the solid state (chapter 6).

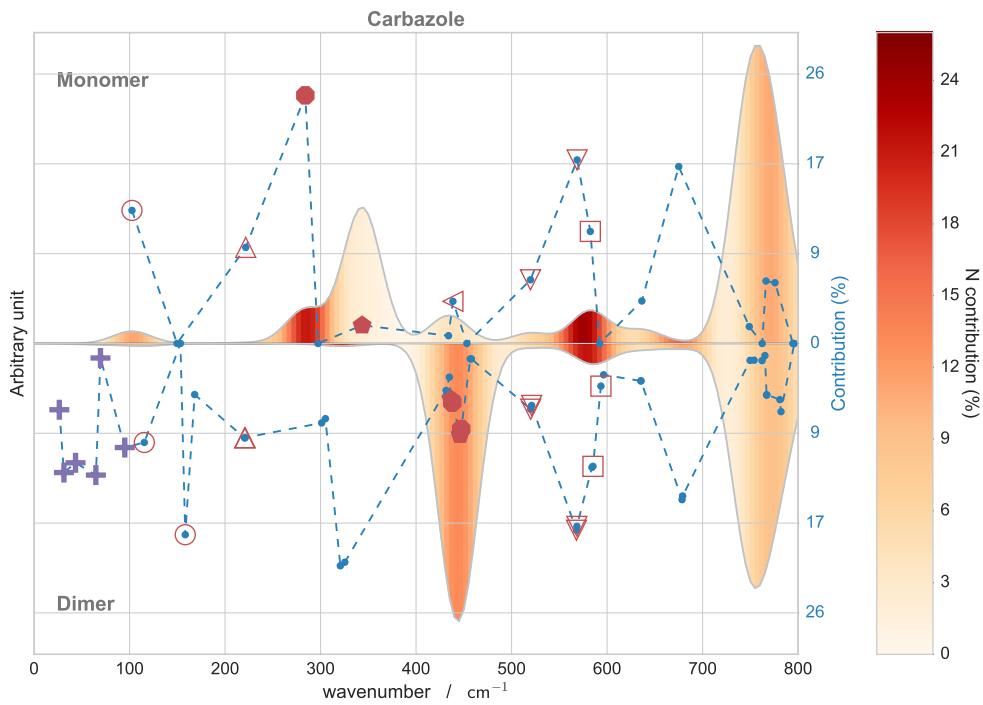


Figure 5.7 – Calculated vibrational spectra of Carbazole monomer and dimer, between 0- 800  $\text{cm}^{-1}$ , showing heteroatom contribution and selected modes.

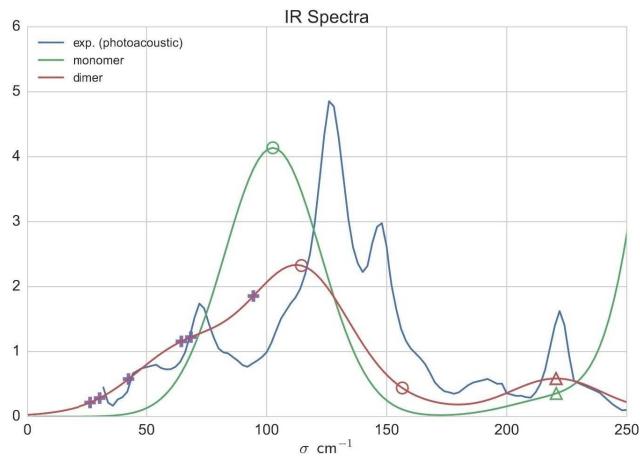


Figure 5.8 – Comparison between calculated vibrational spectra of monomer and dimer with IR experimental at 0- 250  $\text{cm}^{-1}$  of Carbazole

Another category of modes that are specific to the carbazole family will allow us to highlight the presence of moieties associated as shown by the experiments conducted by K. Michaelian. So that the characterisation of these specific modes can be done, we have reported in figure 9 the

evolution of the density of vibrational states against the vibrational frequencies for carbazole's monomer and dimer. It appears distinctly in figure 5.9 two zones with strong densities within the zone  $x_3 = 2$  (comprised between  $250\text{-}500\text{ cm}^{-1}$ ) that should indicate strictly the presence of two spectral zones in the experimental spectrum. Yet, one can find in this spectrum only intense zone centered at  $450\text{ cm}^{-1}$ . Nevertheless, a very weak activity is noted around  $300\text{ cm}^{-1}$ , precisely at the same place where there is a strong density of vibrational states.

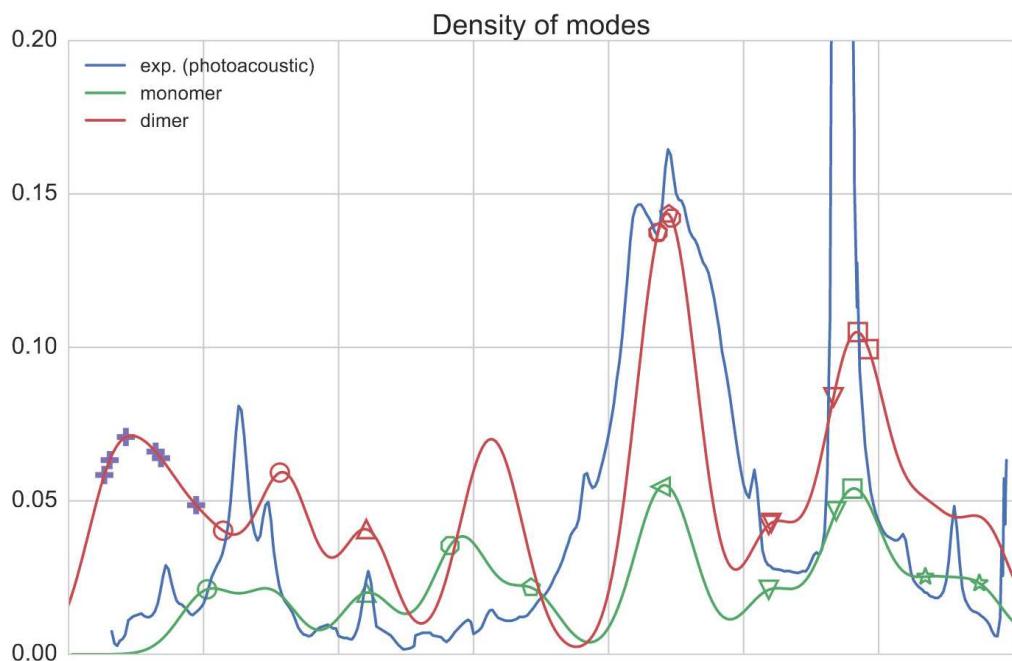


Figure 5.9 – Comparison of density of modes at low wavenumber region of Carbazole.

Our calculations also show that all the dimer's modes (however more numerous than those of the monomers in this zone) present around  $300\text{ cm}^{-1}$  are practically all inactive. On the other side, if the density of modes for the monomer in this region is weaker than the one of the dimer, the  $280\text{-}300\text{ cm}^{-1}$  zone remains potentially the place where detectable signatures could be found for this specie (as it is depicted in figure 5.7). The analysis of a second regions centred at  $450\text{ cm}^{-1}$  reveals this time that two species (monomer and dimer) have both of them a stronger density of modes, but oppositely to the dimer, no monomer's vibration mode localised within this region have a strong signature. Another very important information concerns the N-H wagging modes ( $\square$  and  $\circlearrowleft$ ) at the origin of monomer's spectral signature in the  $280\text{-}350\text{ cm}^{-1}$  zone. As a result of the presence of *inter-molecular* interactions, the position of these modes is strongly displaced towards higher wavenumbers for the dimer. Surprisingly, they are displaced to around  $450\text{ cm}^{-1}$ , what induces an augmentation of the dimer's modes density in this region. Experimentally, this region is where a strong, wide spectral band having several bumps can be found and this seems to agree well with what was calculated for the associated specie (and reported in figure 5.10, on which the modes' intensities are also shown alongside their densities). Moreover, the  $280\text{-}350\text{ cm}^{-1}$  zone is here the experimental spectral signatures

are weak and this corroborates a little more the fact that carbazole is not present as a monomer (at least not predominantly) when the spectrum was recorded.

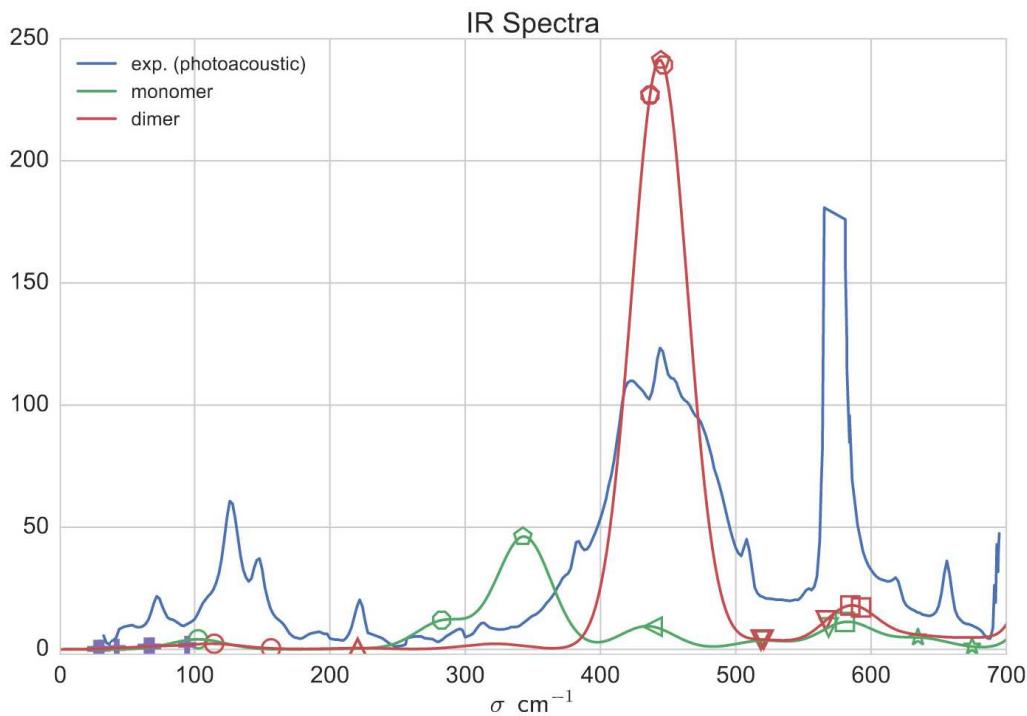


Figure 5.10 – Comparison of calculated vibrational spectra and experimental data taking into account theoretical intensities for Carbazole.

Given the fact that we do not have experimental data for the other molecules of the carbazole family, we can not verify the validity of our hypothesis for all the species. But we can report the shifts of the N-H wagging modes predicted by our calculations and show that they are not specific to the carbazole system only for the family having nitrogen as the heteroatom (figures 5.11). Yet, the spectra reported in figures 5.11 clearly show that the shifts of the wagging modes signatures are not all equivalent and apparently they strongly depend on the system in question. In figure 5.12, the pieces of information concerning these shifts were regrouped aiming to find an explanation concerning the influence of the substituent or the moiety. Unluckily, no direct link connecting the intensity of these shifts and any particular symmetry of these moieties, or the position of the substituent in the molecule, the number of them, etc, for instance, could be found. On the other hand, these results clearly show that these shifts concern only the species of the carbazole family. A deeper study will then be presented in the following paragraphs.

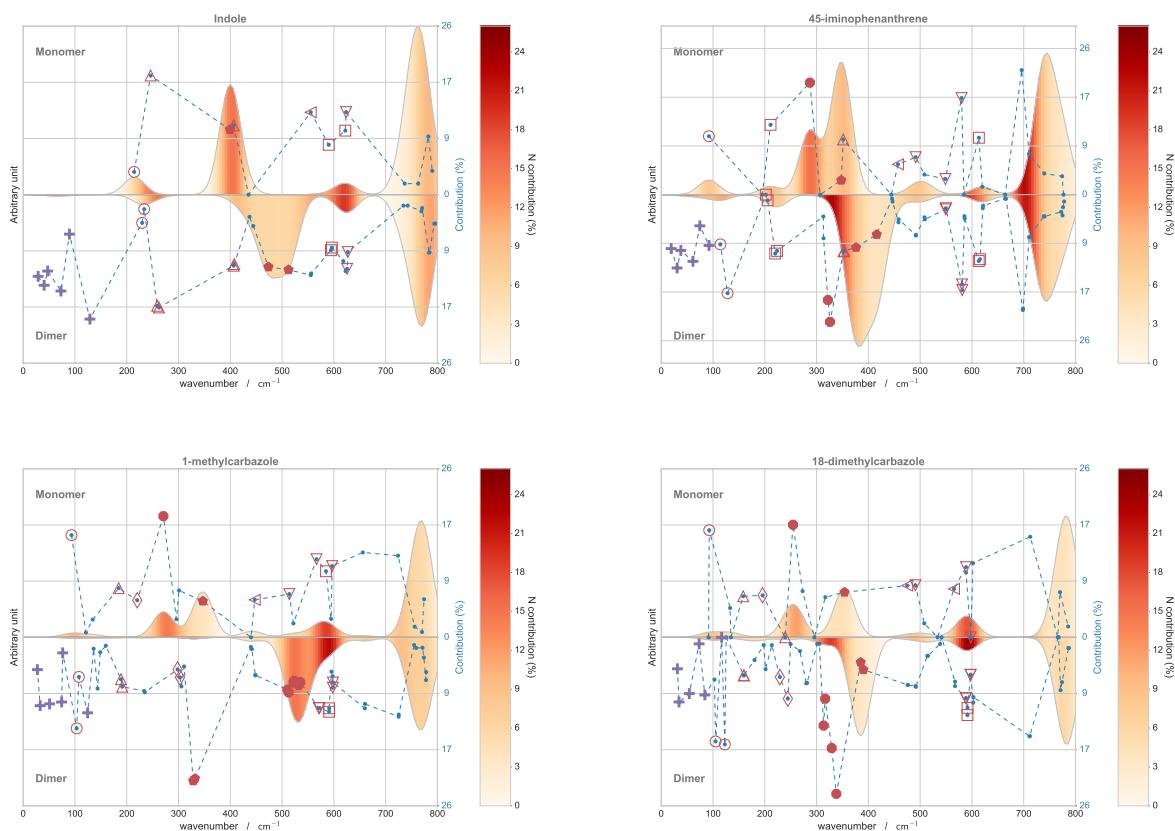


Figure 5.11 – Comparison between momoner and dimer calculated vibrational spectra for **a** Indole, **b** 4,5-iminophenanthrene, **c** 1-methylcarbazole and **d** 1,8-dimethylcarbazole at 0- 800 cm<sup>-1</sup> interval.

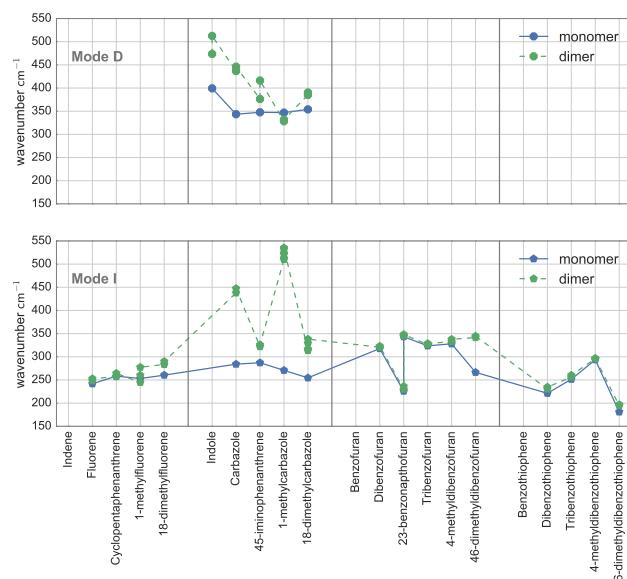


Figure 5.12 – Variation of D and I modes between monomer- dimer of studied molecules

In summary, two type of modes characteristic to the associated systems bearing nitrogen as the heteroatom were highlighted in our work: they are *inter*-molecular modes localised at very low wavenumbers and N-H wagging modes for which the frequency, *in fine*, strongly depends on the presence or not of *inter*-molecular interactions.

## 5.4 Dibenzothiophene

For dibenzothiophene monomer and dimer, their results are very similar for what concerns the frequency, intensity and sulphur contribution to all the vibrations. The only exceptions are the particularities listed for the nitrogen as heteroatom case (N-H wagging).

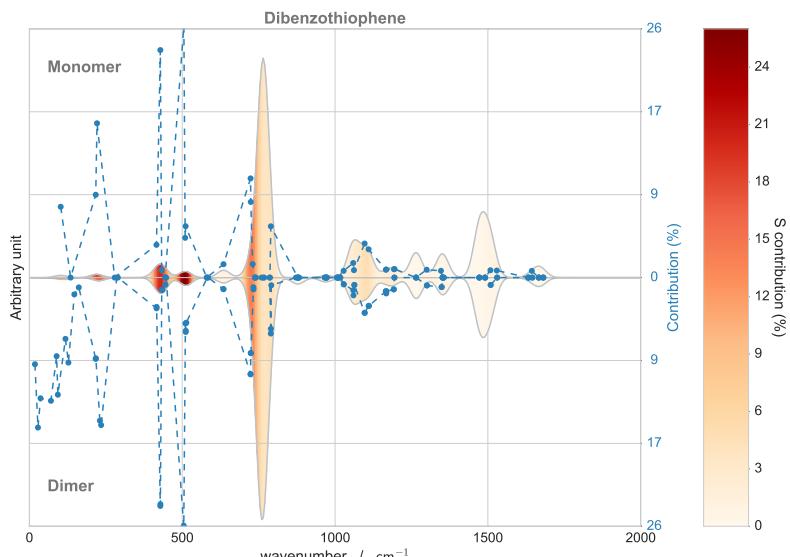


Figure 5.13 – Monomer and dimer comparison of calculated vibrational spectra of Dibenzothiophene

Analysing the  $0\text{--}800\text{ cm}^{-1}$  spectral zone for which we had particularly noted modifications (on frequencies and intensities) due to the wagging of nitrogen shows now no particularity when we have sulphur as hetero atom, neither for frequency's positioning (figure 5.14) nor mode densities (figure 5.15). In any case and for all the spectral zones, there is an excellent theory-experiment agreement. The strong similarities found for dibenzothiophene's monomer and dimer show that it is probably not possible to discriminate spectroscopically the presence of associated systems, exception made for the spectral zone  $x_3 = 1$  relative to the *inter*-molecular modes and the ones coupled to them, such as the butterfly ones).

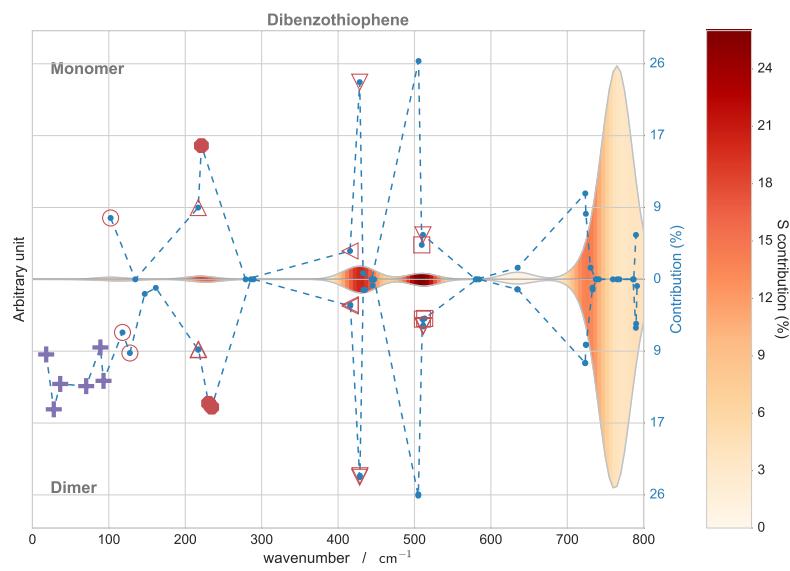


Figure 5.14 – Calculated vibrational spectra of Carbazole monomer and dimer, between 0- 800  $\text{cm}^{-1}$ , showing heteroatom contribution and selected modes.

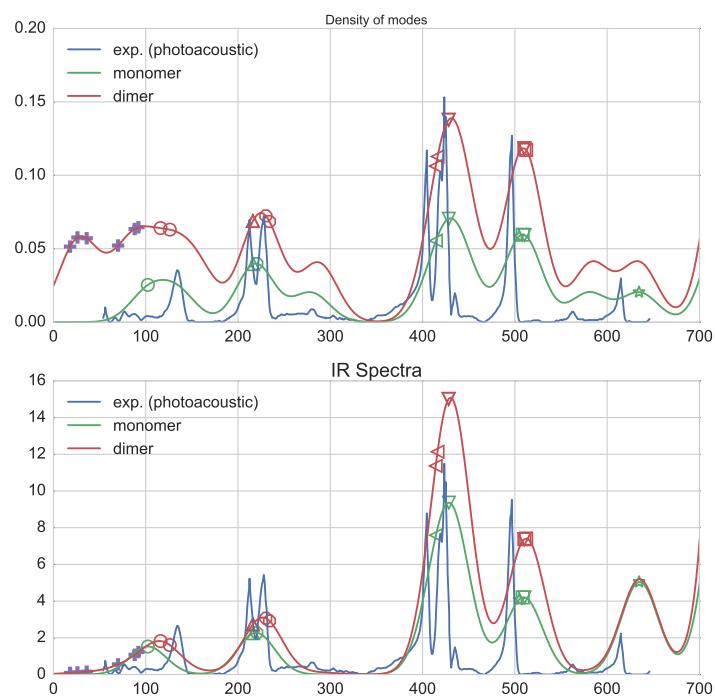


Figure 5.15 – Comparison between monomer, dimer calculated and experimental spectra, Density of modes at the top and normal vibrational modes and IR intensities at the bottom.

These ensembles of conclusions are also valid for all the other molecules of the thiophene family (figure 5.16). They also apply to the ensemble of molecules studied in this work having oxygen as heteroatom or carbon (figure 5.17 and 5.18). Finally, apart the *inter-molecular* modes and the ones that are coupled with them, only the N-H wagging mode for the molecules of the nitrogen family have enough singularities to allow the discrimination of the monomeric and associated states.

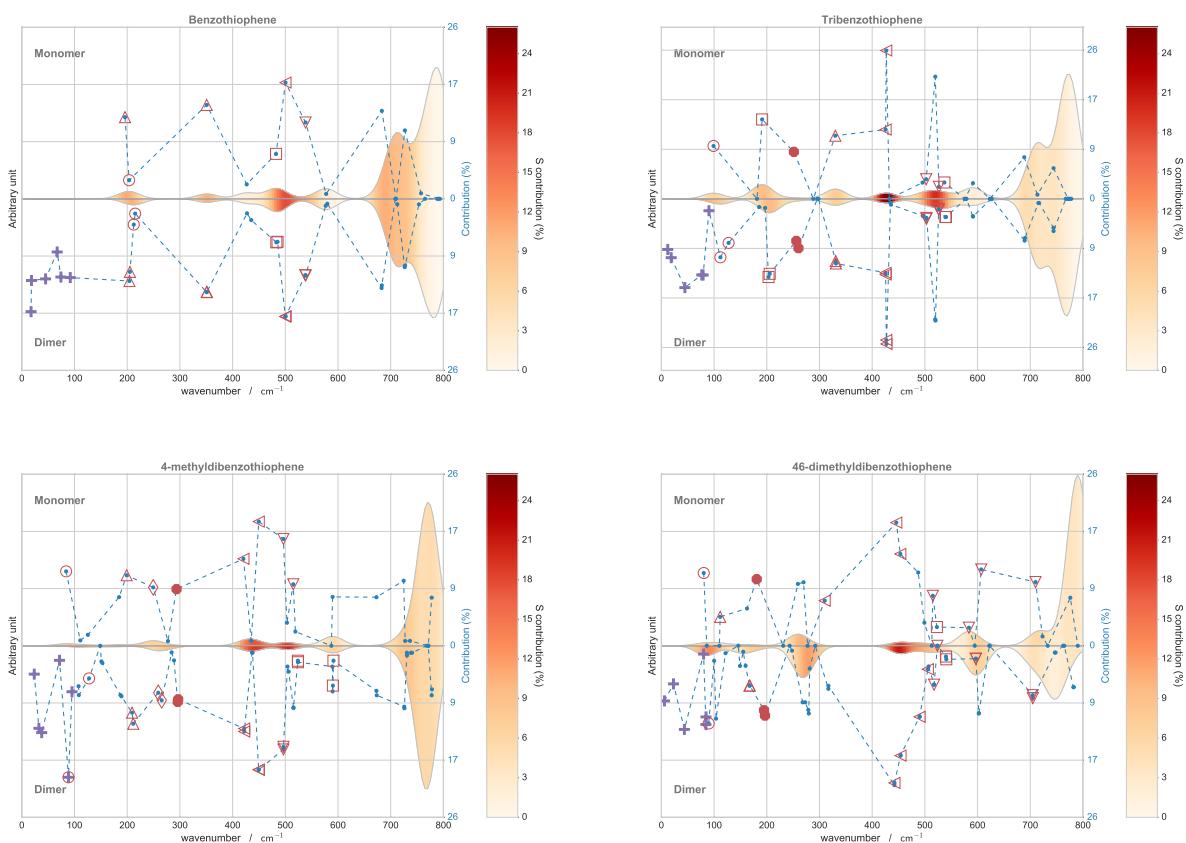


Figure 5.16 – Comparison between monomer and dimer calculated vibrational spectra for **a** Benzothiophene, **b** Tribenzothiophene, **c** 4-methyldibenzothiophene and **d** 4,6-dimethyldibenzothiophene at 0- 800  $\text{cm}^{-1}$  interval.

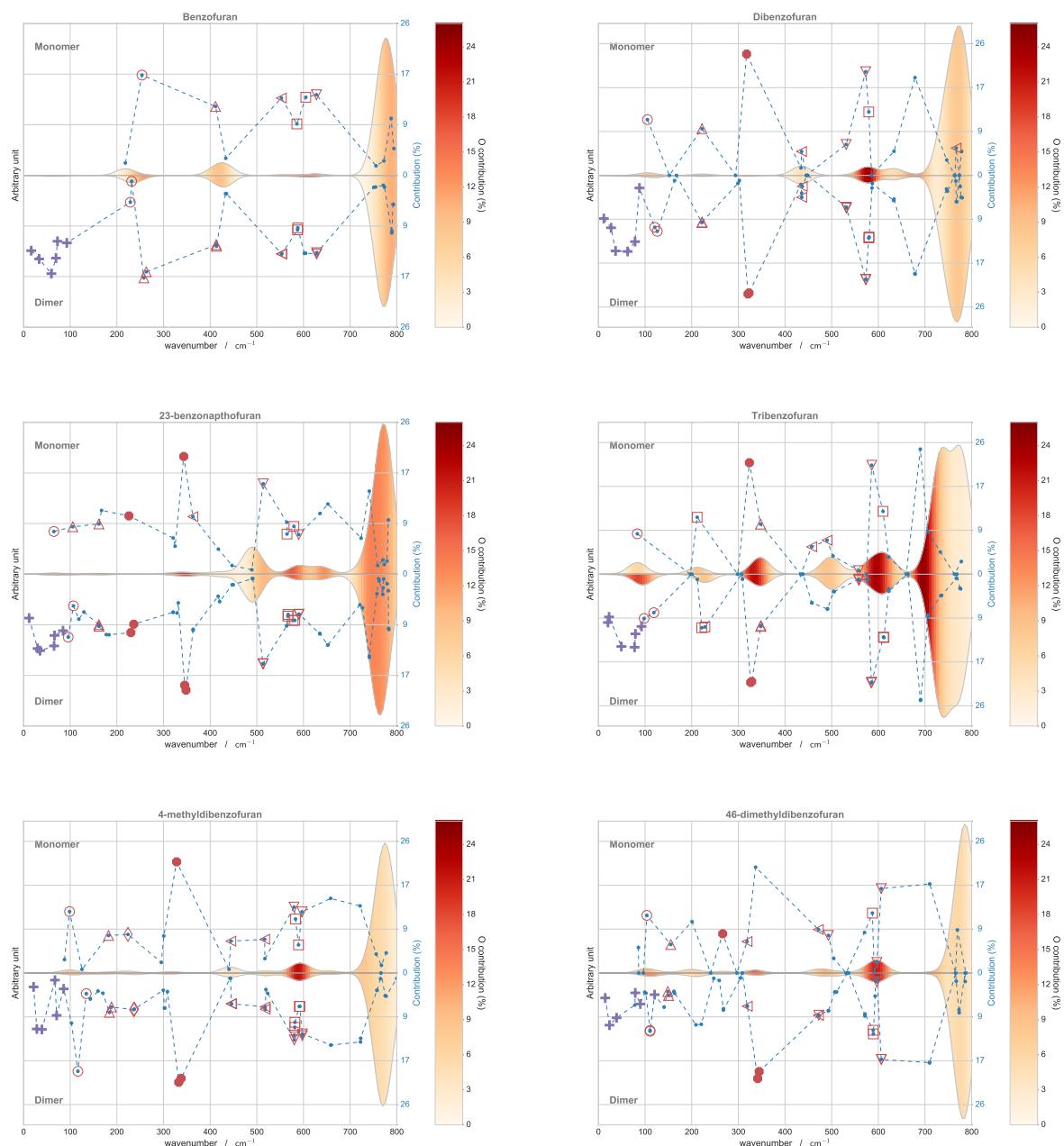


Figure 5.17 – Comparison between momoner and dimer calculated vibrational spectra for **a** Benzofuran, **b** Dibenzofuran, **c** Benzonaphthofuran, **d** Tribenzofuran, **e** 4-methylbenzofuran and **f** 4,6-dimethylbenzofuran at 0- 800 cm<sup>-1</sup> interval.

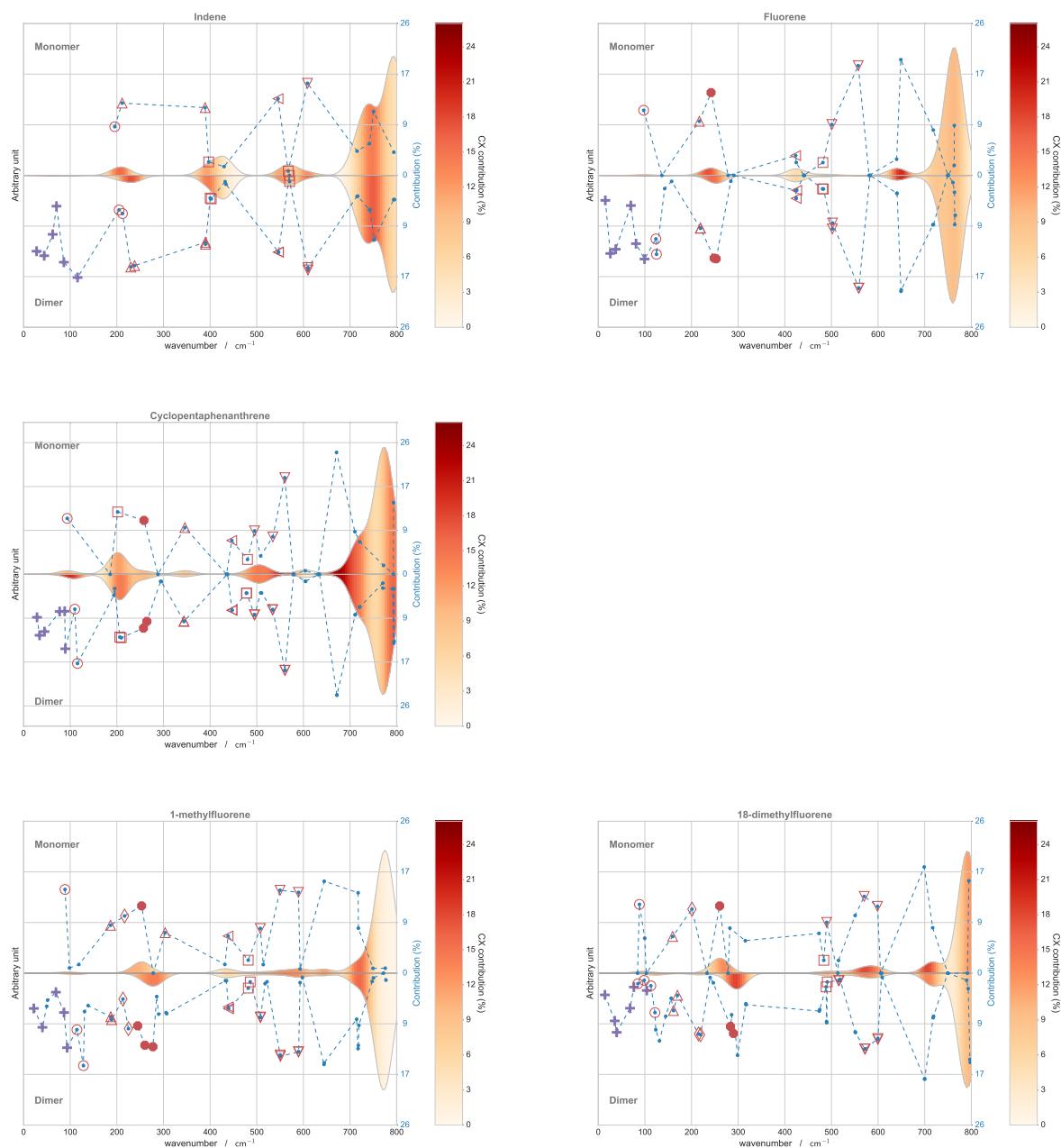


Figure 5.18 – Comparison between momoner and dimer calculated vibrational spectra for **a** Indene, **b** Fluorene, **c** Cyclopentaphenanthrene, **d** 1-methylfluorene and **e** 1,8-dimethylfluorene at 0- 800 cm<sup>-1</sup> interval.

# **Chapter 6**

## **Network modes identification**

## Introduction

The aim of the present chapter is to perform high-precision quantum mechanical calculations to compute spectra for systems in solid state in order to characterize the signature of the intermolecular vibration for a large variety of aromatic compounds.

First, acenes, a family of linear polynuclear aromatic compounds possessing only  $\pi - \pi$  interactions [196, 197, 198, 199, 200], for which high precision experimental data are available up to pentacene [201], comprise a first benchmark example of this chapter that illustrates both the potential and the computational challenges found in this line of inquiry.

Second, additional calculations were performed for other compounds of the acene family, as well as other systems of the aromatic polycycle molecules family. The aim of these additional calculations is verifying if the position and activity of the network modes identified for the acene molecules in the solid state, and particularly the more active mode, depend on the size, length, the moiety or on the presence of hetero elements of the polycycle system in study.

### 6.1 Computational details

Spectral assignments, determined on the basis of both wavenumber and relative intensity values, has already been made primarily using computed spectra for individual molecules, and dimers [49]. Figure 6.1, derived from this work, gives a direct comparison of experimental and calculated spectra with our conclusion as for assignment. Unfortunately, in this study both calculations and experiments do not give information about very low-wavenumbers and no real assignment of the inter molecular modes has been achieved. However, this experimental information exist. Far-infrared spectra provide information about intermolecular and inter-ring interactions and about  $\pi$ -stacking in particular. For example, by using coherent synchrotron radiation it is now becoming possible to make experimental photoacoustic IR measurements in the range 7 to 30  $\text{cm}^{-1}$  as demonstrated recently by Billinghamurst and Michaelian[202]. Concerning acene, Hineno et al [203] measured spectra of naphthalene and anthracene in the region 190-50  $\text{cm}^{-1}$  at liquid helium temperature. In the same way, bands in the experimental vibrational spectra of tetracene and pentacene that cannot be interpreted by Hineno et al based on computed wavenumbers for a single molecule are expected to arise from lattice modes. the failure of these calculations is normal. The accurate description of these network modes requires solid state calculations. In this chapter we propose to use this approach in order to describe these network modes. These calculations were finally 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)[204]. The generalized gradient approximation (GGA) in the PBE parameterization was considered as the exchange correlational functional [133]. 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/ $\text{\AA}$ , 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 [83, 205, 152, 155, 87]. To obtain accurate band gaps, we have used the GW approximation [206]. 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 500 eV.

Second, periodic calculations for tetracene were also performed in order to calculate IR intensities with CRYSTAL14 program[207], 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  $\Gamma$  Point were calculated for the optimized geometry. The empirical dispersion correction by Grimme [83] was performed to take into account weak interactions.

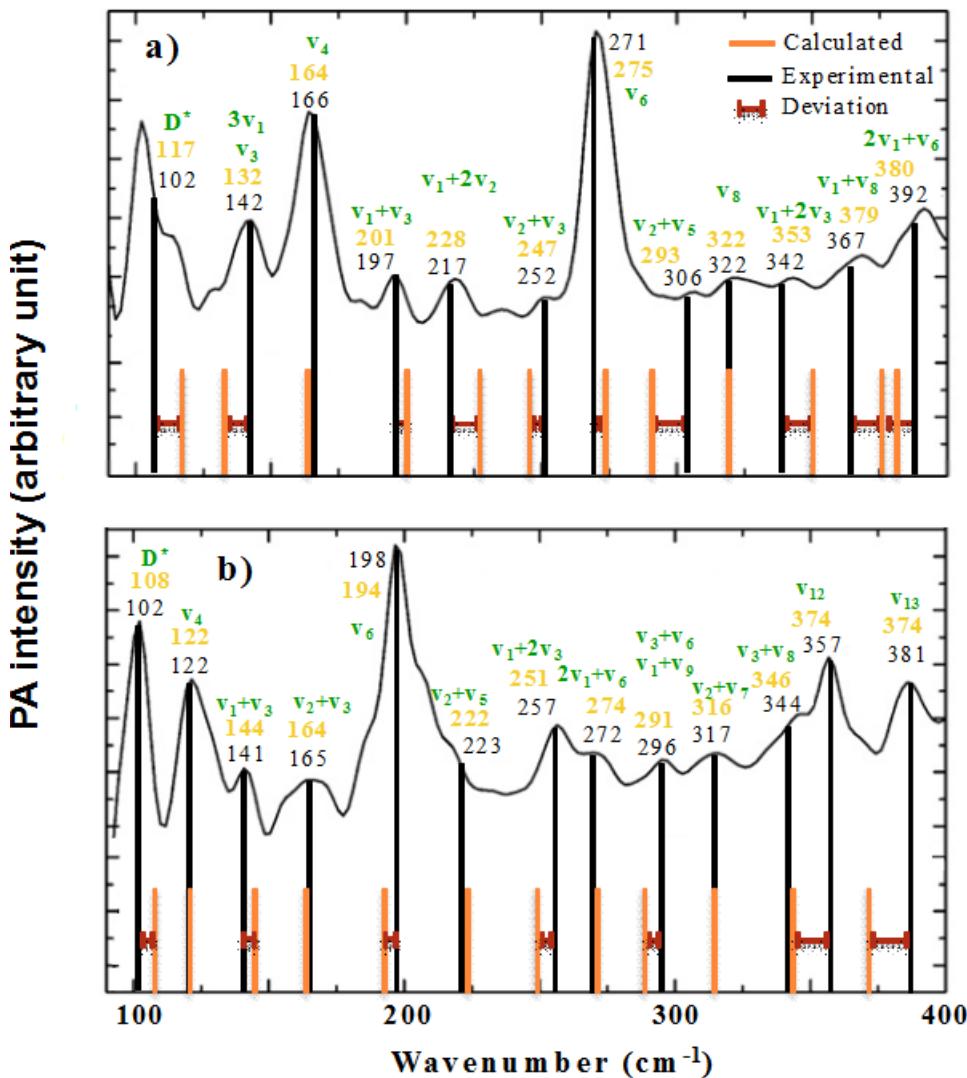


Figure 6.1 – Final assignment of the experimental infrared transitions for a) tetracene and b) pentacene. Calculated wavenumbers are indicated on the spectra with orange lines, and corresponding values above. Experimental values are indicated for comparison, and represented with dark lines

Only the 0-150 cm<sup>-1</sup> zone is herein commented in this chapter since it is where the signatures we intend to identify are located. Notwithstanding, for the ensemble of systems, the

Table 6.1 – Lattice parameters of Tetracene molecule calculated in VASP code

	<b>D2</b>	<b>D3</b>	<b>TS</b>	<b>TS-SCS</b>	<b>PBE*</b>	<b>Exp ref[196]</b>
<b>a</b>	7.54 (7.41)	7.92	7.71	7.70	9.25	7.90
<b>b</b>	5.96 (5.83)	6.00	5.99	6.08	7.21	6.03
<b>c</b>	13.28 (13.39)	13.38	13.38	13.47	16.62	13.53
$\alpha$	102.10 (102.21)	101.21	101.36	100.97	101.36	100.30
$\beta$	114.33 (113.23)	113.26	113.61	114.39	112.14	113.20
$\gamma$	85.20 (85.11)	85.89	85.74	85.68	86.67	86.30
<b>Volume (Å<sup>3</sup>)</b>	531.41 (519.01)	572.96	554.74	563.74	1007.16	582.85

\* without dispersion correction

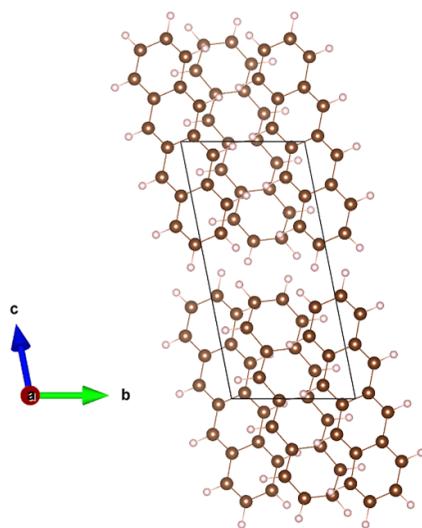
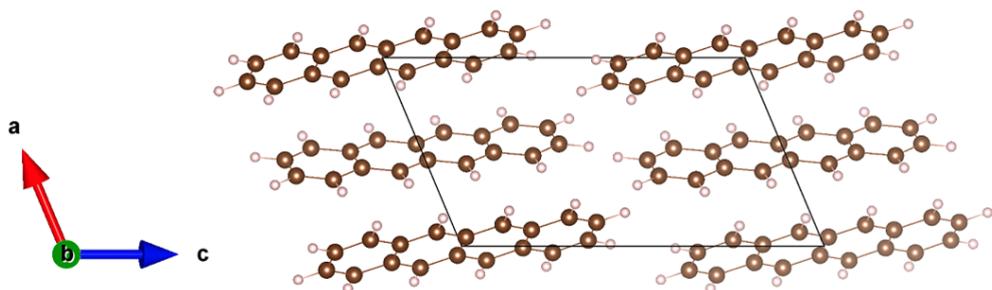
<sup>0</sup> Parenthesis values were calculated with CRYSTAL program

calculated data were also extended to englobe the 150-400 cm<sup>-1</sup>, allowing in this way both the validation of our calculations in a better experimentally identified zone and the identification of the specific modes of the ‘non interacting’ modes. As such, the solid state calculations herein performed are generally in a very good agreement with the available experimental data. This also contributes to the validation of our results using the same calculation conditions for the regions below 150 cm<sup>-1</sup>.

## 6.2 A first case: Tetracene

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 the tetracene compound are presented in Table 6.1. 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 [208], 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. 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 6.2 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  $\nu_2$  and  $\nu_8$ ) are modified with respect to the isolated system due to intermolecular interactions. Besides, our previous study on tetracene dimers revealed that

Figure 6.2 – The *bc* plane of Tetracene CrystalFigure 6.3 – The *ac* plane of Tetracene Crystal

intermolecular interactions induced the appearance of some characteristic spectral signatures at very low wavenumbers. For instance, the mode experimentally found at  $106\text{ cm}^{-1}$ , 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  $102\text{ cm}^{-1}$  for the pentacene system in our previous work. 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  $69.7\text{ cm}^{-1}$  ( $\nu_1$ ),  $128.5\text{ cm}^{-1}$  ( $\nu_3$ ) and  $320.8\text{-}323.8\text{ cm}^{-1}$  ( $\nu_8$ ). Unfortunately, there is no experimental data available below  $100\text{ cm}^{-1}$  for the tetracene system. But it seems that the  $\nu_3$  and  $\nu_8$  modes can now be identified as the modes of medium intensity reported from experimental data at  $142$  and  $322\text{ cm}^{-1}$ , respectively. Not 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 between  $200\text{-}400\text{ cm}^{-1}$ , whatever their intensities are, they do

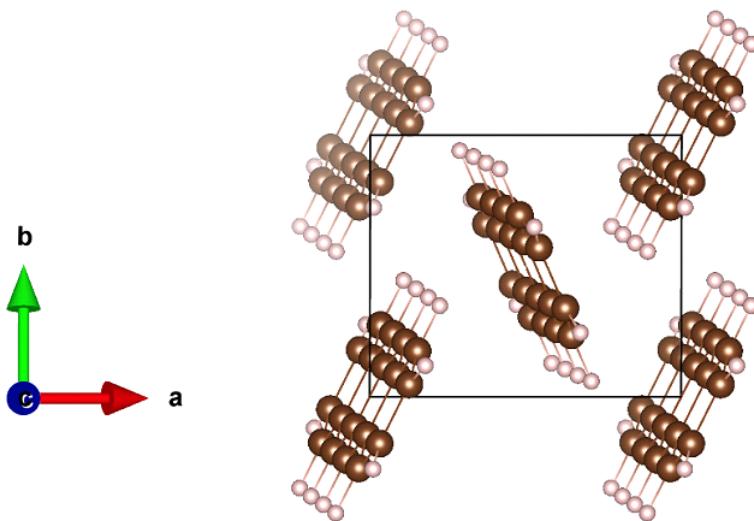


Figure 6.4 – The *ab* plane of Tetracene Crystal

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  $200\text{ cm}^{-1}$  during the molecular study of tetracene and its dimer.

The tetracene molecule is a favorable case to detect the network signature. Indeed, as indicated in Table 6.2, the isolated tetracene molecule strictly owns a single signature in the zone of study ( $93\text{ cm}^{-1}$ ) that could potentially hide the network's one. Nevertheless, as the intensity of this mode for the monomer is expected to be zero accordingly to our calculations, one should not have problems assigning it. This situation is very favorable and does not represent the most common situation. Occasional overlap of vibrators could be a limiting factor for our capability to identify the nature of these observed vibrators in some cases. Still, this fact does not question the presence of these active network modes that we intend to identify, neither the fact that these modes can always be associated to one or two intra-molecular vibrations.

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 spectroscopic results of acenes molecules in the crystalline state.

Table 6.2 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE tetracene system).

	<b>Monomer</b>	<b>Dimer</b>	<b>Experimental (P.A.) this work</b>	<b>VASP/CRYSTAL</b>
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)
$\nu_1$	57 (0.8)	15 (<0.01) 31 (0) 52 (0)		69.7 (8.2)
$\nu_2$	93 (0)	59 (0.6) 80 (1.0) 87 (0.1) 94 (0) 112 (0) 124 (0)	106 (m) + sh	102.4 (2.1) 105.7 (9.5)
$3\nu_1$	132 (<0.01)			
$\nu_3$	153 (0.01)		142 (m)	128.5 (7.1) 148 (0.1)
$\nu_4$	164 (1.3)	162 (0) 168 (1.7) 169 (0)	166 (s) + sh	164.6 (15.5)
$\nu_5$	199 (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)	
$\nu_6$	275 (1.03)	282 (1.6) 286 (0)	271 (m)	267.4 (3.0) 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)
$\nu_8$	322 (<0.01)	325 (0.06) 335 (0.03) 335 (0)	342 (w)	323.8 (0.5)
$\nu_1 + 2\nu_3$	353 (<0.01)			
$\nu_1 + \nu_8$	379 (<0.01)			
$\nu_{10}$	385 (0.04)	390 (0.02)	392 (m)	
$2\nu_1 + \nu_6$	380 (<0.01)			

Italic: Intermolecular modes

### 6.3 Impact of the chain-length parameter on the vibrational signature of the crystalline network.

We are aware that it is not possible to jump to conclusions on the basis of the results of a single system. In the same way, additional calculations were performed within this work for other compounds of the acene family, as well as other systems of the aromatic polycycle molecules family. The aim of these additional calculations is verifying if the position and activity of the network modes identified for the tetracene molecule, and particularly the mode clearly localise at around  $106\text{ cm}^{-1}$ , depend on the size, length or on the moiety of the polycycle system in study.

The first parameter we studied is the chain-length within the acene family. A series of calculations on the benzene, naphthalene, anthracene and pentacene calculations was performed. The crystalline structures from which we were able to do our calculation are presented hereafter. More particularly, the benzene molecule has three different phases: Phase I correspond to an orthorhombic Pbca space group that contain four molecules per unit cell. In opposite, the two other phases called II and III contain two molecules in the unit cell and crystallizes in a monoclinic system (space group P21/c) at high pressure [209, 210]. In that case molecule configuration observed by differents study is T-Shape [210]. Under ambient conditions of temperature and under low-pressure, crystals of the naphthalene belong to the monoclinic system (space group P21/a) and contain two molecules per unit cell [211]. Under the same conditions, the crystals of both anthracene and tetracene crystallize in a monoclinic phase P21/a [200] and in a triclinic phase P $\bar{1}$  [212] respectively. Crystals of the pentacene system crystallize into a triclinic phase (space group P $\bar{1}$ ); it is reported to have two polymorphs structures in its bulk phase, polymorph C [213] and polymorph H [197]. The ensemble of the numerical results relative to the cell parameters and cell volumes can be found in Tables 6.3-7. A representation of these crystalline cells is given in Figures 6.5-8.

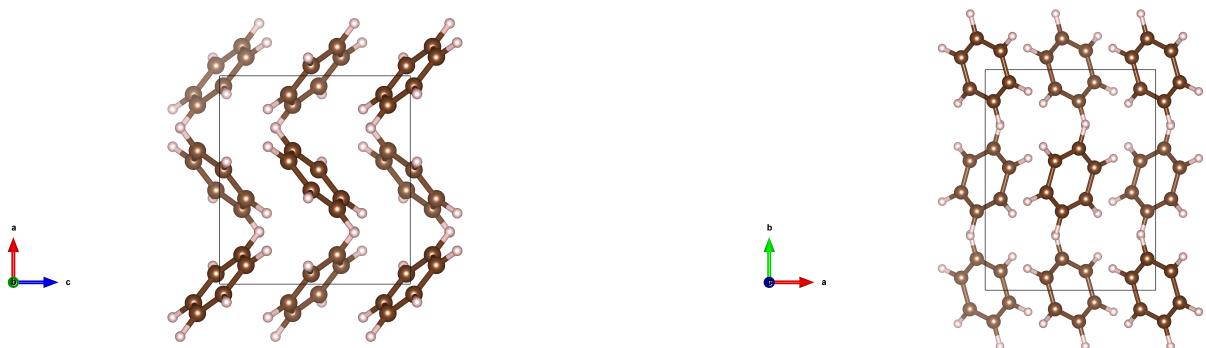


Figure 6.5 – The *ac* and *ab* planes of Benzene crystal

Table 6.3 – Lattice parameters of Benzene molecule calculated in VASP code

	D2	D3	TS	TS-SCS	PBE*	Exp ref[209]
<b>a</b>	7.24 (6.90)	7.45	7.32	7.33	8.29	7.36
<b>b</b>	9.09 (9.13)	9.54	9.41	9.42	10.35	9.37
<b>c</b>	6.63 (6.55)	6.84	6.71	6.72	7.73	6.70
<b>Volume (Å<sup>3</sup>)</b>	434.57 (412.45)	485.94	462.14	463.15	662.85	465.60

\* without dispersion correction

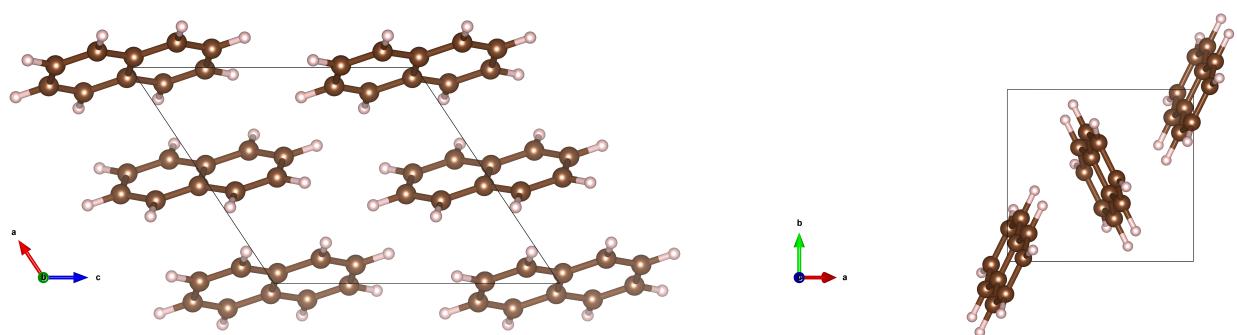
<sup>0</sup> Parenthesis values were calculated with CRYSTAL programFigure 6.6 – The *ac* and *ab* planes of Naphthalene crystal

Table 6.4 – Lattice parameters of Naphthalene molecule calculated in VASP code

	D2	D3	TS	TS-SCS	PBE*	Exp ref[211]
<b>a</b>	7.71 (7.78)	8.00	7.85	7.89	9.64	8.03
<b>b</b>	5.84 (5.74)	6.07	6.01	6.03	6.51	5.89
<b>c</b>	8.50 (8.57)	8.68	8.63	8.65	9.14	8.57
$\beta$	125.18 (125.25)	123.40	123.68	123.69	121.46	123.59
<b>Volume (Å<sup>3</sup>)</b>	313.24 (312.75)	351.97	338.45	342.52	488.96	337.65

\* without dispersion correction

<sup>0</sup> Parenthesis values were calculated with CRYSTAL program

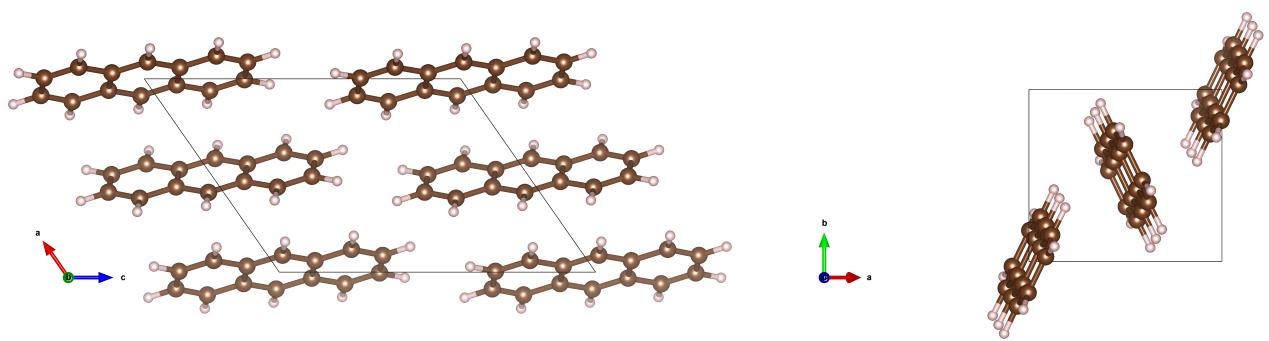
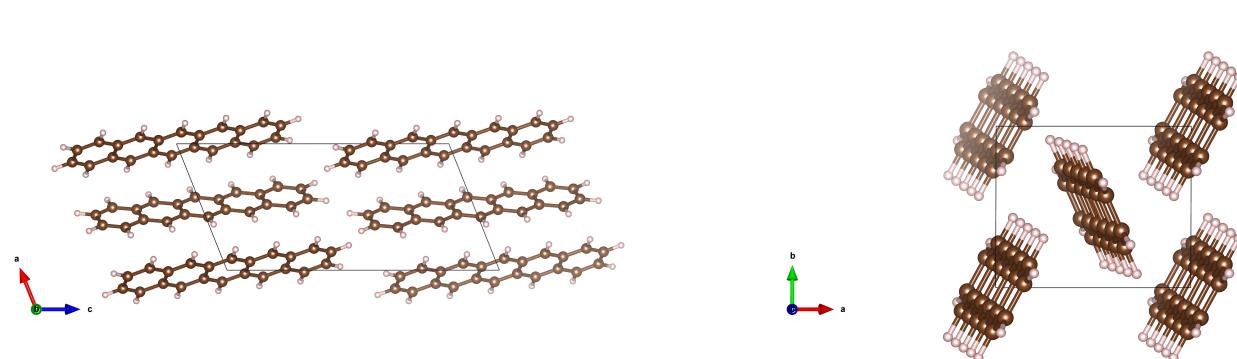
Figure 6.7 – The *ac* and *ab* planes of Anthracene crystal

Table 6.5 – Lattice parameters of Anthracene molecule calculated in VASP code

	D2	D3	TS	TS-SCS	PBE*	Exp ref[200]
<b>a</b>	8.04 (8.40)	8.39	8.23	8.26	9.97	8.55
<b>b</b>	5.99 (5.67)	6.12	6.03	6.07	6.54	6.02
<b>c</b>	11.14 (11.30)	11.25	11.17	11.21	11.60	11.17
$\beta$	126.09 (128.26)	124.95	125.46	125.42	121.61	124.60
<b>Volume (Å<sup>3</sup>)</b>	433.85 (422.71)	473.17	452.06	457.75	644.25	473.17

\* without dispersion correction      <sup>0</sup> Parenthesis values were calculated with CRYSTAL program

Figure 6.8 – The *ac*, *ab* planes of Pentacene polymorph C.

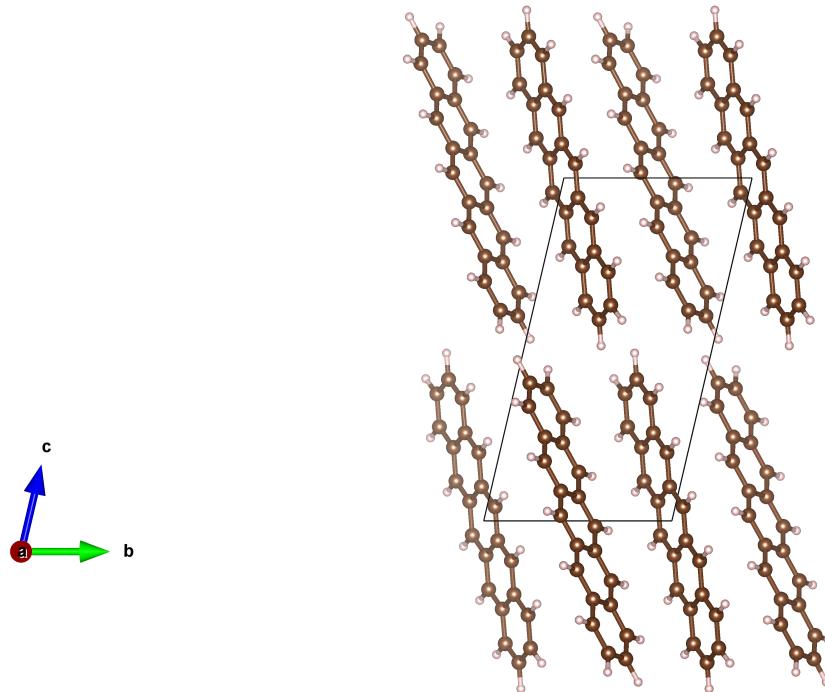


Figure 6.9 – The  $bc$  plane of Pentacene polymorph H at the bottom

Table 6.6 – Lattice parameters of Pentacene molecule (Polymorphe H) calculated in VASP code

	<b>D2</b>	<b>D3</b>	<b>TS</b>	<b>TS-SCS</b>	<b>PBE*</b>	<b>Exp ref[197]</b>
<b>a</b>	6.16 (6.18)	6.27	6.13	6.37	6.93	6.27
<b>b</b>	7.25 (7.32)	7.77	7.54	7.68	8.79	7.78
<b>c</b>	13.96 (13.98)	14.48	14.33	14.46	15.36	14.53
$\alpha$	78.99 (78.92)	76.72	77.89	77.62	72.10	76.48
$\beta$	88.69 (90.73)	88.19	87.76	87.66	86.85	87.68
$\gamma$	83.83(83.00)	84.43	84.49	84.42	85.07	84.68
<b>Volume (Å<sup>3</sup>)</b>	608.37 (615.63)	683.04	644.52	687.62	886.39	685.15

\* without dispersion correction      <sup>0</sup> Parenthesis values were calculated with CRYSTAL program

Table 6.7 – Lattice parameters of pentacene molecule (polymorphe C) calculated in VASP code

	<b>D2</b>	<b>D3</b>	<b>TS</b>	<b>TS-SCS</b>	<b>PBE*</b>	<b>Exp ref[213]</b>
<b>a</b>	7.40 (7.02)	7.81	7.67	7.61	9.03	7.90
<b>b</b>	5.98 (5.70)	6.09	6.02	6.09	6.54	6.06
<b>c</b>	15.70 (15.68)	15.97	15.82	15.92	16.23	16.01
$\alpha$	102.67 (89.98)	101.97	102.08	101.40	101.91	101.90
$\beta$	113.48 (97.19)	112.31	112.73	112.57	110.91	112.60
$\gamma$	84.98 (90.00)	85.33	85.53	85.63	86.42	85.80
<b>Volume (Å<sup>3</sup>)</b>	622.05 (622.95)	686.75	659.16	668.18	878.28	692.38

\* without dispersion correction

<sup>0</sup> Parenthesis values were calculated with CRYSTAL program

The results associated to the acene family, i.e., benzene, naphthalene, anthracene and pentacene, are reported in Tables 6.3- 7 hereafter. The calculation conditions in this piece of modelling are identical to those used just before for the tetracene molecule. The main conclusions arising from the study of these systems are in agreement with those obtained for the tetracene system. Indeed, for these systems, the structural parameters and unit cell volumes are better described by the PBE-D3 level of theory if a comparison is made with the other methods used in this work (except for the two anthracene and benzene systems where the TS and the TSC data are in relative better agreement in comparison to the experimental data [156]). The agreement between our theoretical and experimental data meets the requirements to the frequency calculation of the crystalline network. Moreover, the inclusion of the dispersive terms (like D3 or the modified D2:TS) must be introduce onto the frequency calculation because of the well-known dispersive character of the intermolecular modes to describe. The vibrational results calculated for these systems are reported in Table 6.8- 11 in which the frequency values can also be found for each acene family-member monomer and dimer. According to the conclusions drawn for the tetracene system case, one can find for these systems at least an active signature associated to a network vibrational mode around 100 cm<sup>-1</sup> (88 and 106 cm<sup>-1</sup> for naphthalene, 98 cm<sup>-1</sup> for anthracene, and 102 cm<sup>-1</sup> for pentacene). One should note that the benzene molecule (also in the acene family) in its T-Shape conformation needs a specific analysis. As a matter of fact, for this conformation, the intermolecular interactions of the crystal network do not correspond any longer to the simple  $\pi$ - $\pi$  interactions (as it is the case of the other systems herein studied) but to interactions of the type  $\pi$ -H, as demonstrated by Katrusiak et al [210]. Yet, although the interactions are of different nature, our calculations show that the main active vibrational signature of the crystalline network remains localised in the 100-130 cm<sup>-1</sup> region. Moreover, in the following paragraphs, it will be shown that the spectral position remains unchanged for different systems only for this (these) specific mode(s). For all the other modes, depending on the symmetry of the system and on the strength of the long-range interactions between neighbour molecules, the spectral signatures can vary in position and in intensity. This observation only constitutes the major information to be extracted from our calculations given the aim of this work.

Table 6.8 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state  
(PBE Benzene system).

	<b>Monomer</b>	<b>Dimer</b>	<b>Experimental</b>	<b>VASP/CRYSTAL</b>
Assignment	$\nu(\text{cm}^{-1})$	$\nu(\text{cm}^{-1})$	$\nu(\text{cm}^{-1})$	$\omega(\text{cm}^{-1})$
	Int(km/mol)	Int(km/mol)		Int(km/mol)
	<i>7 (0.04)</i>			
	<i>12 (&lt;0.01)</i>			
	<i>48 (&lt;0.01)</i>			
	<i>56 (0.07)</i>			
	<i>63 (0.08)</i>		63 (0.69)	
	<i>75 (0.03)</i>		71 (0.05)	
			101 (0.69)	
			119 (4.24)	
			126 (1.11)	
			128 (4.83)	
$\nu_1$	426 (0)	424 (0.003)	410 (0.05)	
		425 (0.003)	419 (0.79)	

Italic: Intermolecular modes

Table 6.9 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Naphthalene system).

Monomer		Dimer	Experimental ref [214] $\nu(\text{cm}^{-1})$	VASP/CRYSTAL $\omega(\text{cm}^{-1})$ Int(km/mol)
Assignment	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$ Int(km/mol)		
		33 (0.04) 36 (0) 55 (0) 75 (0.21) 88 (0.08) 101 (0)		
$\nu_1$	177 (3.2)	181 (3.8) 188 (0) 201 (1.2)	176	174 (2.60)
$\nu_2$	191 (0)	201 (0)	195	190 (0.04) 208 (0.03) 210 (1.26)
$\nu_3$	379 (1.5)	372 (2.2) 372 (0)	359	362 (0.35) 364 (1.54)
$\nu_4$	403 (0)	396 (0.02)		386 (0) 391 (0)

Italic : Intermolecular modes

Table 6.10 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Anthracene system).

	<b>Monomer</b>	<b>Dimer</b>	<b>Experimental</b> ref [215]/ ref [216]	<b>VASP/CRYSTAL</b>
Assignment	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$	$\omega(\text{cm}^{-1})$ Int(km/mol)
		22 (0.004)		
		34 (0)		
		57 (0)	63 (m)	
		75 (0.05)	72 (w)	
		87 (0.8)		98(1.37)
$\nu_1$	93 (1.5)	111 (0.8) 112 (1.2)	106/104(m) 110 (s)	101 (4.03) 119 (0.05)
$\nu_2$	124 (0)	148 (0.6) 149 (0.02)	137/126(s) 166 (s)	153 (1.77) 155 (0.07)
$\nu_4$	243 (1.4)	244 (2.1) 244 (0)	234 235 (s)	239 (1.25) 239 (0.33)
$\nu_5$	276 (1.1)	250 (0) 253 (0.2)	244	236 (0) 237 (0)
		285 (0.02)	287	278 (0)
		286 (0.08)		281 (0)
$\nu_6$	394 (0.05)	396 (0.2) 396 (0.04)	380/361(vw) 380 (w)	375 (0.23) 379 (0.36)
$\nu_7$	407 (0)	407 (0) 407 (0.04)	397	389 (0) 391 (0)

Italic: Intermolecular modes

Table 6.11 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Pentacene system).

Monomer	Dimer	Experimental (P.A.) ref [49]	VASP/CRYSTAL
Assignment	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\omega(\text{cm}^{-1})$ Int(relative)
	19 ( $<0.01$ ) 27 (0.0) 38 (0.5)	42 (0.4) 57 (0.0) 72 (0.9) 74 (0.07) 84 (0) 95 (0.0)	58 (0.46)
$\nu_3$	97(0.003)	101 (0.002) 108 (0.2)	102 (m)  102 (5.17)
$\nu_4$	122 (1.0)	122 (1.5)  125 (0.0) 130 (0.08)	122 (m)  117 (0.99) 127 (0.21)  134 (6.1)
$\nu_1 + \nu_3$	144 (0.04)		141 (m)  140 (0.68) 148 (1.58)
$\nu_2 + \nu_3$	164( $<0.01$ )	168 (0.2) 168 (0.05)	165 (vw)  188 (w) 198 (m)  180 (2.54)
$\nu_6$	194 (1.4)	204 (2.2)	204 (2.36)
$\nu_3 + \nu_4$	224 (0.0)	210(0.002)	223 (vw)
$\nu_2 + \nu_5$	222 (0.002)		
$\nu_1 + 2\nu_3$	251 (0.0)	247 (0.05)	254 (0.29)
$\nu_9$	265 (0.03)	255 (0.02)	257 (w)
$2\nu_1 + \nu_6$	274 ( $<0.01$ )		272 (vw)  272 (0.45)
$\nu_3 + \nu_6$	291 ( $<0.01$ )	302 (0.005)	296 (w)
$\nu_1 + \nu_9$	292 ( $<0.01$ )	303 (0.07)	
$\nu_2 + \nu_7$	316 ( $<0.01$ )		317 (w)
$\nu_3 + \nu_8$	346 (0.03)		344 (vw)
$\nu_{12}$	374 (0.1)		352 (1.12) 355 (0.27)
$\nu_{13}$	374 (0.03)		382 (0.01) 386 (0.34)

Italic: Intermolecular modes

## 6.4 Impact of both the chain-length and "widths" parameters on the vibrational signature of the crystalline network.

The second parameter is the one characterising at the same time the chain lengths and "widths". Calculations for the Phenanthrene, Pyrene, Chrysene and Perylene molecules were then performed. Furthermore, an additional calculation was performed for the fluorene molecule in order to verify that the signature of the interaction modes does not depend of the nature of the aromatic cycles composing the studied systems. The crystalline structures used for our modelling correspond to those already found in literature for low-pressure room-temperature conditions, *i.e.*, monoclinic P21 [211], monoclinic P21/a [211], monoclinic I2/c [217], monoclinic P21/c [218] and orthorombic Pnam [219], for phenanthrene, pyrene, chrysene, perylene and fluorene, respectively. The numerical results so-obtained are reported in Tables 6.12-21, and the corresponding crystalline cells are shown in Figures 6.10-15.

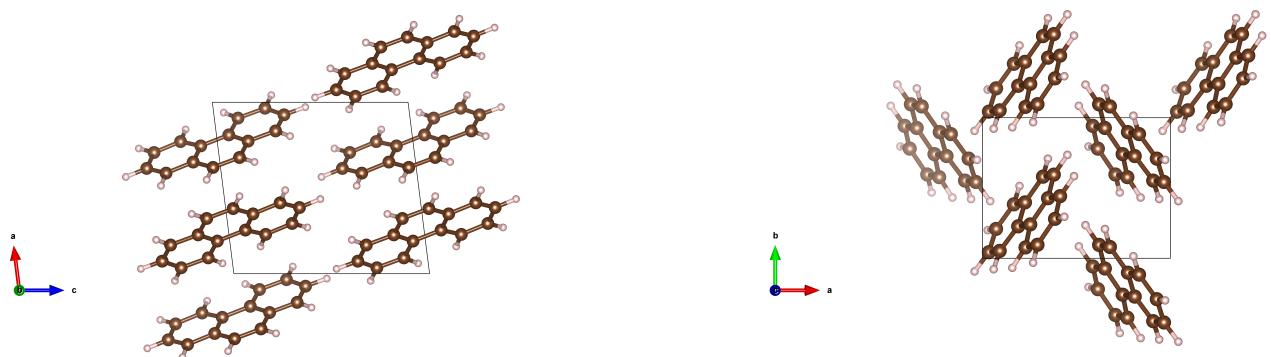


Figure 6.10 – The *ac* and *ab* planes of Phenanthrene crystal

Table 6.12 – Lattice parameters of Phenanthrene molecule calculated in VASP code

	D2	D3	TS	TS-SCS	PBE*	Exp ref[211]
<b>a</b>	7.91 (8.61)	8.31	8.12	8.14	9.59	8.47
<b>b</b>	6.04 (5.46)	6.13	6.06	6.09	6.86	6.17
<b>c</b>	9.30 (9.24)	9.44	9.35	9.37	10.08	9.47
$\beta$	95.89 (95.32)	97.18	96.73	96.73	100.43	98.01
<b>Volume (Å<sup>3</sup>)</b>	441.56 (432.83)	477.42	456.69	460.92	651.81	489.72

\* without dispersion correction

<sup>0</sup> Parenthesis values were calculated with CRYSTAL program

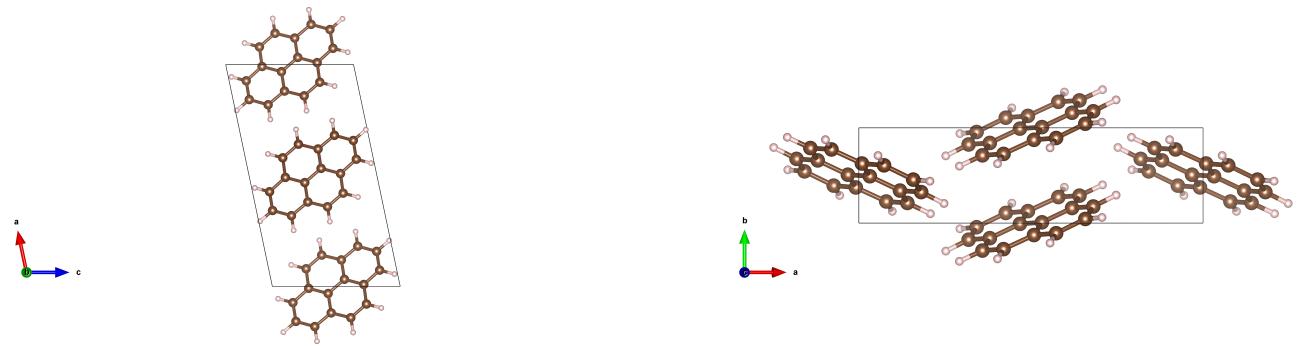
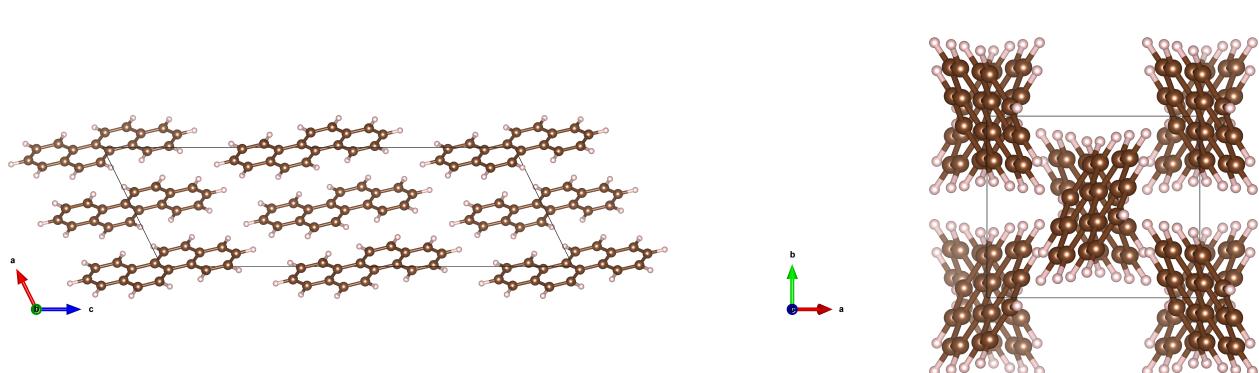
Figure 6.11 – The *ac* and *ab* planes of Pyrene crystal

Table 6.13 – Lattice parameters of Pyrene molecule calculated in VASP code

	D2	D3	TS	TS-SCS	PBE*	Exp ref[211]
<b>a</b>	14.64 (13.22)	14.93	15.02	15.10	16.17	15.35
<b>b</b>	3.98 (4.41)	4.04	3.93	4.00	4.64	3.85
<b>c</b>	8.03 (8.01)	8.42	8.43	8.42	9.10	8.65
$\beta$	101.56 (98.28)	101.88	102.02	101.83	103.82	103.03
<b>Volume (Å³)</b>	470.08 (461.94)	497.74	487.22	497.70	690.67	498.00

\* without dispersion correction      <sup>0</sup> Parenthesis values were calculated with CRYSTAL program

Figure 6.12 – The *ac* and *ab* planes of Chrysene crystal

6.4. Impact of both the chain-length and "widths" parameters on the vibrational signature of  
145 the crystalline network.

Table 6.14 – Lattice parameters of Chrysene molecule calculated in VASP code

	<b>D2</b>	<b>D3</b>	<b>TS</b>	<b>TS-SCS</b>	<b>PBE*</b>	<b>Exp ref[217]</b>
<b>a</b>	7.85 (8.67)	8.18	8.13	8.04	9.61	8.39
<b>b</b>	6.08 (5.40)	6.23	6.13	6.13	6.54	6.20
<b>c</b>	25.10 (24.88)	25.39	25.19	25.15	26.10	25.20
$\beta$	117.05 (114.80)	116.39	116.56	116.87	115.76	116.20
<b>Volume (Å<sup>3</sup>)</b>	1067.63 (1057.28)	1157.58	1123.04	1105.90	1476.90	1175.00

\* without dispersion correction

<sup>0</sup> Parenthesis values were calculated with CRYSTAL program

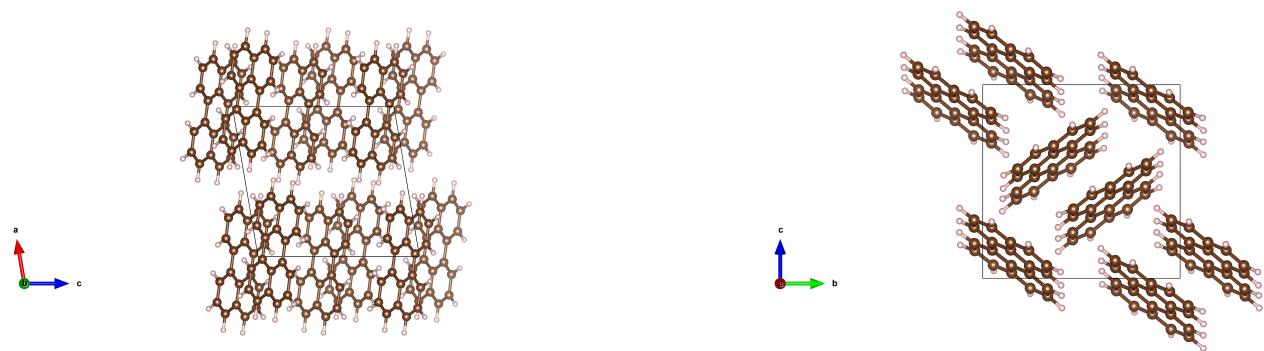


Figure 6.13 – The *ac* and *bc* planes of Perylene crystal

Table 6.15 – Lattice parameters of Perylene molecule calculated in VASP code

	<b>D2</b>	<b>D3</b>	<b>TS</b>	<b>TS-SCS</b>	<b>PBE*</b>	<b>Exp ref[218]</b>
<b>a</b>	9.99 (10.37)	10.29	10.21	10.27	10.80	10.27
<b>b</b>	10.67 (10.27)	11.00	10.91	10.98	11.41	10.84
<b>c</b>	10.51 (10.91)	10.92	10.64	10.83	12.85	11.28
$\beta$	99.54 (103.62)	99.70	99.25	99.55	102.75	100.53
<b>Volume (Å<sup>3</sup>)</b>	1105.54 (1128.72)	1218.72	1169.03	1203.94	1544.16	1234.29

\* without dispersion correction

<sup>0</sup> Parenthesis values were calculated with CRYSTAL program

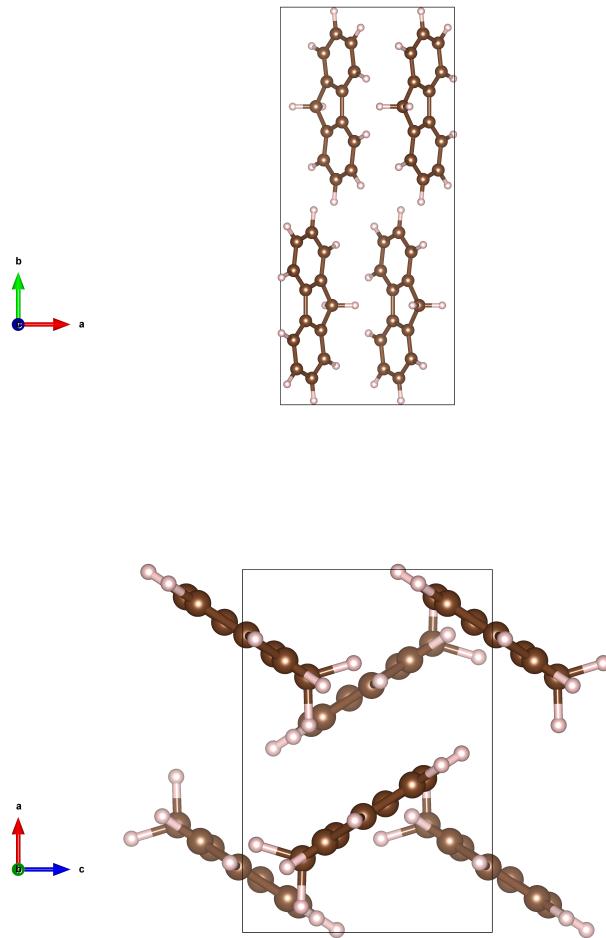
Figure 6.14 – The *ab* and *ac* planes of Fluorene crystal

Table 6.16 – Lattice parameters of Fluorene molecule calculated in VASP code

	<b>D2</b>	<b>D3</b>	<b>TS</b>	<b>TS-SCS</b>	<b>PBE*</b>	<b>Exp ref[219]</b>
<b>a</b>	7.96 (8.25)	8.30	8.08	8.16	9.77	8.48
<b>b</b>	18.40 (18.61)	18.90	18.77	18.82	19.54	18.92
<b>c</b>	5.54 (5.26)	5.72	5.64	5.68	6.16	5.72
<b>Volume (Å³)</b>	810.91 (807.12)	897.69	855.73	871.59	1176.23	916.60

\* without dispersion correction      <sup>0</sup> Parenthesis values were calculated with CRYSTAL program

6.4. Impact of both the chain-length and "widths" parameters on the vibrational signature of  
147 the crystalline network.

Table 6.17 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state  
(PBE Phenanthrene system).

Assignment	Monomer		Dimer	Experimental ref [220]/ ref [221]	VASP/CRYSTAL
	$\nu(\text{cm}^{-1})$	Int(km/mol)	$\nu(\text{cm}^{-1})$	$\nu(\text{cm}^{-1})$	$\omega(\text{cm}^{-1})$ Int(km/mol)
			11 (0.003) 14 (0.005) 42 (0.02) 69 (0.2) 74 (0.06) 81 (0.02)		
				89	57 (0.4) 63 (1.05) 73 (0.01) 88 (0.91) 93 (1.88)
$\nu_1$	99 (0)		102 (1.04) 113 (0.61)	97	103 (0.34) 118 (1.88)
$\nu_2$	103 (1.06)		123 (0.44) 126 (0.02)	124/127 143 (sh)	119 (0.05) 127 (0.53)
$\nu_3$	235 (5.01)		239 (9.23) 244 (0.35)	229(w)/ 233 (s)	231 (0.44) 241 (0.86)
$\nu_4$	251 (0)		250 (0.31) 255 (0.5)	248(s)/	241 (0.12) 243 (4.21)
$\nu_5$	255 (0.43)		255 (0.21) 259 (0.13)	250 (m) 289(w) 398(vw)/ 395 (vw)	250 (2.51) 251 (0.04) 394 (0.34) 395 (0.11)
$\nu_6$	412 (0)		411 (0.03) 415 (0.28)	409(w)/ 408 (ms)	403 (0.1) 408 (0.63)

Italic: Intermolecular modes

Table 6.18 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Pyrene system).

	<b>Monomer</b>	<b>Dimer</b>	<b>Experimental</b> ref [183]	<b>VASP/CRYSTAL</b>
Assignment	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$	$\omega(\text{cm}^{-1})$ Int(km/mol)
		<i>17 (&lt;0.01)</i>		
		<i>30 (0)</i>		
		<i>40 (0)</i>		
		<i>69 (0.01)</i>		<i>50 (0.05)</i>
		<i>94 (0.01)</i>	83 (m)	<i>94 (5.98)</i>
		<i>95 (0)</i>		
$\nu_1$	100 (0.66)	115 (1.60)	98 (m)	<i>97 (2.31)</i> 105 (0.33)
			121 (w)	
			138 (vw)	135 (2.71)
$\nu_2$	157 (0)	162 (0.03)	157 (w)	<i>157 (0.77)</i> 162 (4.27)
			181 (vw)	
			197 (vw)	
$\nu_3$	220(12.56)	225(24.34)	221 (m)	<i>211 (4.64)</i> 231 (7.89)
$\nu_4$	260 (0)	<i>264 (0)</i>	260 (w)	<i>261 (0)</i>
$\nu_5$	269 (0)	275 (0.18)	284 (w) 305 (vw) 323 (w)	<i>264 (0)</i>
$\nu_6$	370 (1.54)	370 (2.05)	353 (m)	<i>349 (0.53)</i> 350 (1.71)
$\nu_7$	412 (0)	416(0.004)	396 (vw)	<i>391 (0.82)</i> 394 (0.04)

Italic: Intermolecular modes

6.4. Impact of both the chain-length and "widths" parameters on the vibrational signature of  
149 the crystalline network.

Table 6.19 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state  
(PBE Chrysene system).

	<b>Monomer</b>	<b>Dimer</b>	<b>Experimental</b> ref [222]	<b>VASP/CRYSTAL</b>
Assignment	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$	$\omega(\text{cm}^{-1})$ Int(km/mol)
$\nu_1$	50 (0.21)	53 (0.2)	15 (0.002) 26 (0.02) 34 (0)	41 (0.01) 69 (0.11)
$\nu_2$	79 (0.34)	76 (0.02) <i>86 (0)</i> 88 (0.18) 95 (0.95)		112 (2.88) 114 (3.19)
$\nu_3$	138 (0)	100 (0) 115 (0) 145 (0) 157 (0.01)		100 (0.09) 134 (0.63) 158 (0.26)
$\nu_4$	181 (0)	190 (0) 194 (0.13)		
$\nu_5$	192 (0.52)	194 (0.73) 195 (0)		189 (0.82)
$\nu_6$	243 (7.50)	248 (14.70) 254 (0)	232	247 (1.18) 251 (10.72)
$\nu_7$	300 (0)	302 (0) 303 (0.06)		
$\nu_8$	301 (0.24)	308 (0) 309 (0.44)		292 (2.56) 295 (0.09)
$\nu_9$	396 (0)	396 (0.005) 396 (0)		

Italic: Intermolecular modes

Table 6.20 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Perylene system).

Monomer		Dimer	Experimental ref [183]	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(km/mol)
$\nu_1$	17 (0.00)	14 (0.001)		
		33(0.00)		
		31 (0.00)		39 (0.12)
		35 (0.00)		45 (0.06)
		73 (0.10)	83 (m)	48 (0.17)
		84 (0.15)		87 (3.61)
		89 (0.00)		
				95 (0.29)
		110 (2.00)	98 (m)	103 (1.03)
		126 (0.00)		104 (9.86)
$\nu_2$	100 (1.00)		117 (m)	111 (7.38)
				117 (2.72)
			124 (vw)	127 (0.52)
		138 (0.00)	142 (w)	148 (0.68)
$\nu_3$	128 (0.00)	144 (0.08)	157 (w)	152 (1.73)
				185 (1.10)
		190 (15.52)	185 (m)	187 (16.72)
$\nu_4$	182 (8.02)	192 (0.00)		189 (0.55)
$\nu_5$	212 (0.00)	218 (0.00)	205 (vw)	219 (0.00)
		223 (0.02)	222 (vw)	233 (0.06)
$\nu_6$	245 (0.00)	249 (0.09)		244 (0.03)
		253 (0.00)	254 (w)	249 (0.20)
				250 (0.01)
$\nu_7$	264 (0.19)	264 (0.00)		299 (0.02)
		264 (0.27)	283 (vw)	300 (0.62)
$\nu_8$	307 (0.00)	312 (0.00)	311 (vw)	
		315 (0.72)	323 (vw)	348 (0.04)
$\nu_9$	366 (0.00)	367 (0.00)		350 (0.06)
		367 (0.04)	354 (w)	361 (0.13)
				363 (0.14)
$\nu_{10}$	375 (0.00)	375 (0.00)	395 (vw)	
		375 (0.04)		410 (0.31)

Italic: Intermolecular modes

6.4. Impact of both the chain-length and "widths" parameters on the vibrational signature of  
151 the crystalline network.

Table 6.21 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state  
(PBE Fluorene system).

	<b>Monomer</b>	<b>Dimer</b>	<b>Experimental</b> ref [223]	<b>VASP/CRYSTAL</b>
Assignment	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$	$\omega(\text{cm}^{-1})$ Int(km/mol)
		20 (0.01) 23 (0.06) 35 (0.02) 74 (0.06) 85 (0.1) 89 (0.02)		
$\nu_1$	100 (0.84)	119 (1.16) 124 (0.02)	108 (m) 122 (w)	100 (5.17) 101 (2.18) 110 (1.13) 118 (0.83)
$\nu_2$	140 (0)	148 (0.04) 162 (0.04)	140 (w)	123 (0.93)
$\nu_3$	220 (0.22)	221 (0.18) 222 (0.17)	217 (m)	160 (0.02) 222 (0.04) 226 (2.87)
$\nu_4$	254 (9.39)	265 (6.20) 277 (16.91)	256 (m) 266 (m)	248(14.23) 262 (42.37) 266 (0.01)
$\nu_5$	286 (0)	292 (0.04) 297 (0.11)	302 (w)	
$\nu_6$	429 (0.45)	430 (0.54) 430 (0.23)	411 (s)	405 (9.39) 405 (4.71)
$\nu_7$	439 (8.29)	440 (19.00) 442 (1.09)		407(19.44)

Italic: Intermolecular modes

The vibrational data obtained within the same calculation conditions for the polyaromatic phenanthrene, pyrene, chrysene and perylene molecule lead to the same conclusions: each system owns at least one active signature found around  $100\text{ cm}^{-1}$  and inherent to the network ( $93\text{ cm}^{-1}$  for phenanthrene, 94 and  $97\text{ cm}^{-1}$  for pyrene, 112 and  $114\text{ cm}^{-1}$  for chrysene, 87, 103 and  $104\text{ cm}^{-1}$  for perylene). The fluorene molecule, besides having a different aromatic constitution (other than benzenic rings) lead to the similar results, for which the modes were estimated to be found at 88 and  $100\text{ cm}^{-1}$ .

The available vibrational data for these systems are not easily exploitable for the network modes characterization. For these modes, indeed, either the data is incomplete or the intramolecular signatures (highlighted by the calculations of the monomers) overlap the intermolecular ones in the  $100\text{-}150\text{ cm}^{-1}$  zone. For example, the modes associated to pyrene (monomer calc  $100\text{ cm}^{-1}$  int  $0.66\text{ km.mol}^{-1}$  – solid state calc  $97\text{ cm}^{-1}$  int  $2.31\text{ km.mol}^{-1}$ ) and fluorene (monomer calc  $100\text{ cm}^{-1}$  int  $0.84\text{ km.mol}^{-1}$  – solid state calc  $100\text{ cm}^{-1}$  int  $5.71\text{ km.mol}^{-1}$ ) are respectively measured at 98 and  $108\text{ cm}^{-1}$ , illustrating this problem.

## 6.5 Impact of heteroatoms on the vibrational signature of the crystalline network.

Third and last parameter studied was heteroatoms influence inside intermolecular modes. Calculations for Carbazole, Dibenzofuran and Dibenzothiophene were performed, those molecules are nitrogen, oxygen and sulphur analogues of Fluorene. The crystallines structures used for our calculations correspond to low- pressure and room- temperature conditions found in literature. All unit cell contain four molecules and crystallize in an orthorhombic Pbnm [224], orthorhombic Pnam [225] and monoclinic P21/c [226], for carbazole, dibenzofuran and dibenzothiophene respectively. Results for lattices parameter are reported in tables 6.22- 24, and the corresponding crystallines cells are shown in Figures 6.15- 17. In the same way calculations with D3 dispersion correction exhibit the better values for network systems, except for dibenzothiophene where TS-SCS is comparable to D3, because it present unlike only 0.02%.

Intermolecular mode around  $100\text{ cm}^{-1}$  appear not to depend to the side and ring cycle structure. Heteroatoms influence was studied by fluorene analogues, vibrational calculation for monomer, dimer and solid state are shown in tables 6.25- 27. These systems present actives network modes at  $72$ ,  $114$  and  $118\text{ cm}^{-1}$  for carbazole;  $106$ ,  $108\text{ cm}^{-1}$  for dibenzofuran and  $91$ ,  $108$  and  $113\text{ cm}^{-1}$  for dibenzothiophene, accordingly our previous results for acenes family and other systems heteroatoms presence have not impact in frequence values for these actives bands which are produced by intermolecular interactions.

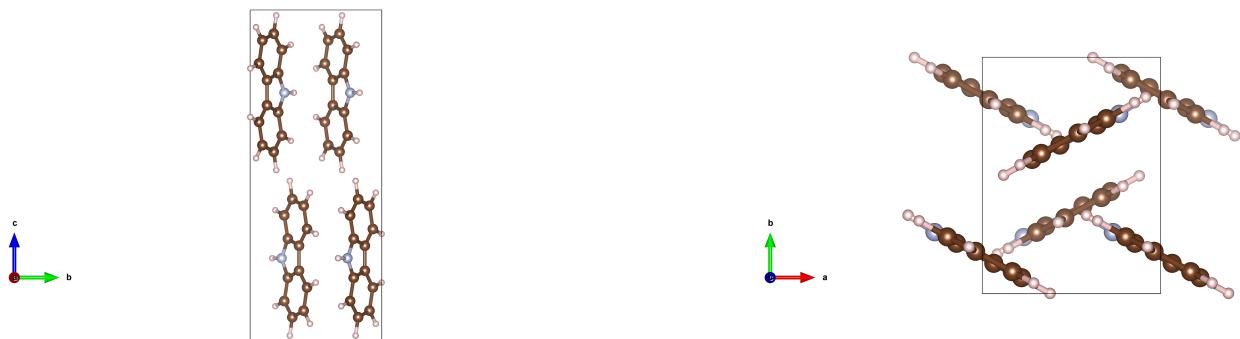
Figure 6.15 – The *bc* and *ab* planes of Carbazole crystal

Table 6.22 – Lattice parameters of Carbazole molecule calculated in VASP code

	D2	D3	TS	TS-SCS	PBE*	Exp ref[224]
<b>a</b>	5.59 (5.51)	5.77	5.70	5.70	6.07	5.71
<b>b</b>	7.27 (7.21)	7.64	7.39	7.44	8.67	7.78
<b>c</b>	18.75 (18.86)	19.20	19.08	19.07	19.67	19.11
<b>Volume (Å³)</b>	762.34 (749.46)	846.60	804.18	808.71	1035.61	848.88

\* without dispersion correction      <sup>0</sup> Parenthesis values were calculated with CRYSTAL program

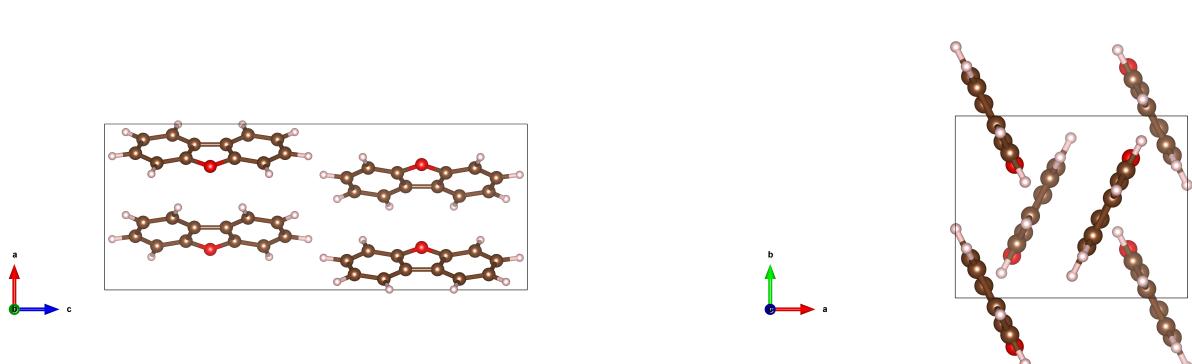
Figure 6.16 – The *ac* and *bc* planes of Dibenzofuran crystal

Table 6.23 – Lattice parameters of Dibenzofuran molecule calculated in VASP code

	<b>D2</b>	<b>D3</b>	<b>TS</b>	<b>TS-SCS</b>	<b>PBE*</b>	<b>Exp ref[225]</b>
<b>a</b>	7.17 (6.94)	7.52	7.36	7.34	8.68	7.70
<b>b</b>	5.69 (5.65)	5.86	5.78	5.79	6.29	5.83
<b>c</b>	18.84 (20.26)	19.22	19.05	19.08	19.81	19.19
<b>Volume (Å<sup>3</sup>)</b>	769.27 (793.69)	847.16	810.05	810.10	1082.30	860.72

\* without dispersion correction

^) Parenthesis values were calculated with CRYSTAL program

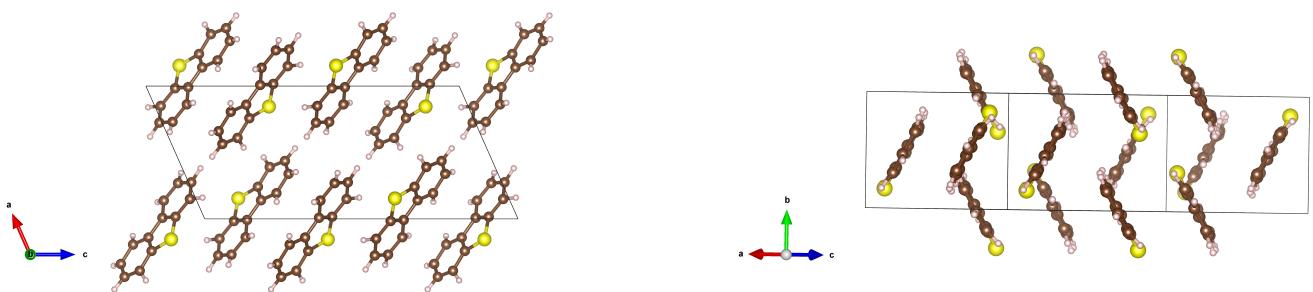
Figure 6.17 – The *ac* plane and other view of molecular arrangement of Dibenzothiophene crystal

Table 6.24 – Lattice parameters of Dibenzothiophene molecule calculated in VASP code

	<b>D2</b>	<b>D3</b>	<b>TS</b>	<b>TS-SCS</b>	<b>PBE*</b>	<b>Exp ref[226]</b>
<b>a</b>	8.26 (8.59)	8.56	8.36	8.54	9.83	8.67
<b>b</b>	5.86 (5.68)	6.03	5.95	6.05	6.48	6.00
<b>c</b>	17.94 (17.83)	18.44	18.04	18.39	20.95	18.70
<b>β</b>	111.54 (112.66)	112.07	111.87	111.79	119.32	113.90
<b>Volume (Å<sup>3</sup>)</b>	808.18 (802.53)	882.92	847.06	883.16	1163.07	889.36

\* without dispersion correction

^) Parenthesis values were calculated with CRYSTAL program

Table 6.25 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Carbazole system).

Monomer		Dimer	Experimental Fig $\nu(\text{cm}^{-1})$	VASP/CRYSTAL $\omega(\text{cm}^{-1})$ Int(km/mol)
Assignment	$\nu(\text{cm}^{-1})$ Int(km/mol)	$\nu(\text{cm}^{-1})$ Int(km/mol)		
		<i>26 (0.05)</i>		
		<i>31 (0.03)</i>		
		<i>43 (0.02)</i>		
		<i>65 (1.32)</i>	71 (w)	<i>66 (0.13)</i>
		<i>70 (0.01)</i>		<i>72 (13.88)</i>
		<i>95 (0.51)</i>		
$\nu_1$	103 (5.54)	115 (2.85) 158 (0.11)	125 (m)	112 (0.86) 114 (3.31) 118 (26.62)
$\nu_2$	150 (0.00)	153 (0.12) 168 (0.04)	146 (w) 190 (vw)	142 (8.41) 185 (0.49)
$\nu_3$	222 (0.38)	221 (0.47) 221 (0.35)	221 (m)	221 (0.26) 222 (0.52)
$\nu_4$	284(15.47)	301 (0.74) 305 (0.01)	294 (w)	283 (0.43)
$\nu_5$	297 (0.00)	321 (0.02) 325 (2.83)	312 (w)	303 (9.96) 309 (12.45)
$\nu_6$	344(61.71)			
$\nu_7$	434(11.32)	431 (7.50) 438 (11.67)	423 (s)	418 (63.87) 419 (6.27)
$\nu_8$	439 (1.44)	435 (5.64) 436 (77.40)		428 (0.78)
		446 (203.89)		434(43.81)
		447 (30.18)		
		457 (13.23)	445 (s)	
$\nu_5$	453 (0.00)	458 (0.02)		441 (1.34)

Italic: Intermolecular modes

Table 6.26 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Dibenzofuran system).

Monomer		Dimer	Experimental ref [227]	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(km/mol)
		<i>12 (0.03)</i>		
		<i>27 (0.04)</i>		
		<i>37 (0.09)</i>		
		<i>63 (0.06)</i>		<i>53 (1.83)</i>
		<i>78 (0.26)</i>		<i>78 (0.01)</i>
		<i>88 (0.08)</i>		
$\nu_1$	105 (2.18)	121 (4.03) 126 (0.04)	103	<i>106 (3.86)</i> <i>108 (0.75)</i> 117 (1.09)
$\nu_2$	152 (0.00)	163 (0.04) 167 (0.06)	150	140 (0.47) 152 (0.14)
$\nu_3$	222 (1.19)	222 (0.34) 222 (1.52)	217	217 (0.42) 221 (2.86)
$\nu_4$	294 (0.00)	299 (0.05) 302 (0.11)	287	286 (0.01)
$\nu_5$	317 (0.01)	321 (0.28) 323 (0.02)	310	314 (0.22) 317 (1.27)
$\nu_6$	434 (6.24)	434 (10.51) 437 (2.71)	419	413 (1.65) 434 (26.52)
$\nu_7$	434 (1.88)	436 (1.41) 436 (1.47)	424	425 (6.91)

Italic: Intermolecular modes

Table 6.27 – Calculated vibrational frequencies ( $\text{cm}^{-1}$ ) of the monomer, dimer and solid-state (PBE Dibenzothiophene system).

Monomer Assignment	$\nu(\text{cm}^{-1})$ Int(km/mol)	Dimer $\nu(\text{cm}^{-1})$ Int(km/mol)	Experimental Table $\nu(\text{cm}^{-1})$	VASP/CRYSTAL $\omega(\text{cm}^{-1})$ Int(km/mol)
		<i>18 (0.05)</i>		
		<i>28 (0.01)</i>		
		<i>36 (0.09)</i>		<i>49 (0.38)</i>
		<i>70 (0.32)</i>		<i>80 (1.86)</i>
		<i>89 (0.01)</i>		<i>84 (0.07)</i>
		<i>93 (0.56)</i>		<i>91 (2.43)</i>
				<i>99 (0.47)</i>
				<i>108 (0.60)</i>
				<i>113 (2.61)</i>
$\nu_1$	102 (1.54)	118 (2.14) 128 (0.06)	135 (m)	138 (0.03) 140 (1.14)
$\nu_2$	135 (0.00)	147 (0.04) 162 (0.10)		163 (0.45) 172 (0.39)
$\nu_3$	217 (0.85)	217 (0.29) 217 (1.09)	213 (m)	213 (2.52) 215 (5.04)
$\nu_4$	221 (1.46)	231 (3.05) 235 (0.07)	228 (m) 269 (vw)	229 (1.93) 230 (3.94)
$\nu_5$	279 (0.00)	286 (0.01) 289 (0.13)	282 (vw)	278 (0.39) 283 (0.11)
$\nu_6$	416 (2.83)	416 (0.24) 416 (4.21)	404 (s)	404 (8.59) 406 (0.39)
$\nu_7$	428 (0.15)	428 (0.24) 429 (0.08)	419 (sh)	415 (6.30) 416 (0.60) 418 (0.24)
$\nu_8$	433 (7.15)	433 (16.66) 435 (0.23)	425 (s)	423 (18.48) 435 (3.29)

Italic: Intermolecular modes

## Conclusions

The DFT calculations presented in this work provide detailed and accurate descriptions of experimental photoacoustic far IR spectra for acene family and for far IR spectra of the various aromatic compounds as a whole. Inter- and intramolecular vibrations 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. 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 spectra.

The intermolecular vibrations for aromatic compounds are found in the same wavelength range of the far infrared spectrum. The specificity of the aromatics  $\pi$ -stacking interactions, reinforced by calculations with distorted dimers and solid-state calculations, allows not only the identification of this family of molecules, but also highlights the presence of  $\pi$ -stacking interaction in a mixture as a whole. As illustration, the example of the carbazole is reported in figure 6.18.

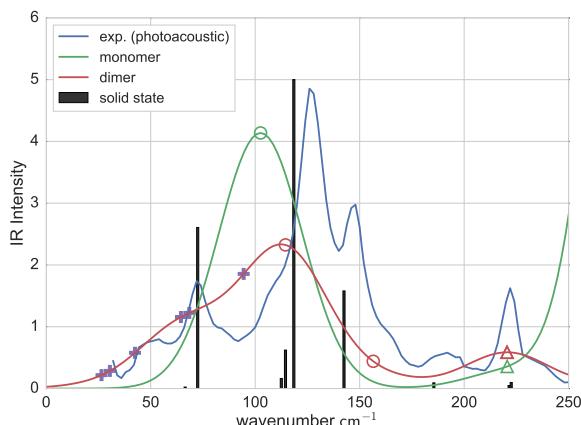


Figure 6.18 – Final assignment of the experimental infrared transitions for carbazole. Experimental values are indicated for comparison, and represented with blue lines

If our previsions turn out to be generalizable for all kind of molecular families and molecules, this could give experimentalists an additional tool of characterisation to identify the presence of systems more or less associated and interacting by means of  $\pi$ - $\pi$  or  $\pi$ -H interactions in the solid state. This work is a complement of the results performed in the next chapter concerning the characterisation of the H-bond interactions in asphaltene systems [228, 194]. In this way, we believe to be henceforth capable of disposing all the vibrational pieces of information concerning the three potential types of interactions present in the asphaltenic medium and provide, in this way, experimentalists a way to find out the nature of the interactions present in petroleum fluids and their residues and, to a lesser extent, to identify the chemical nature of oil's constituents.

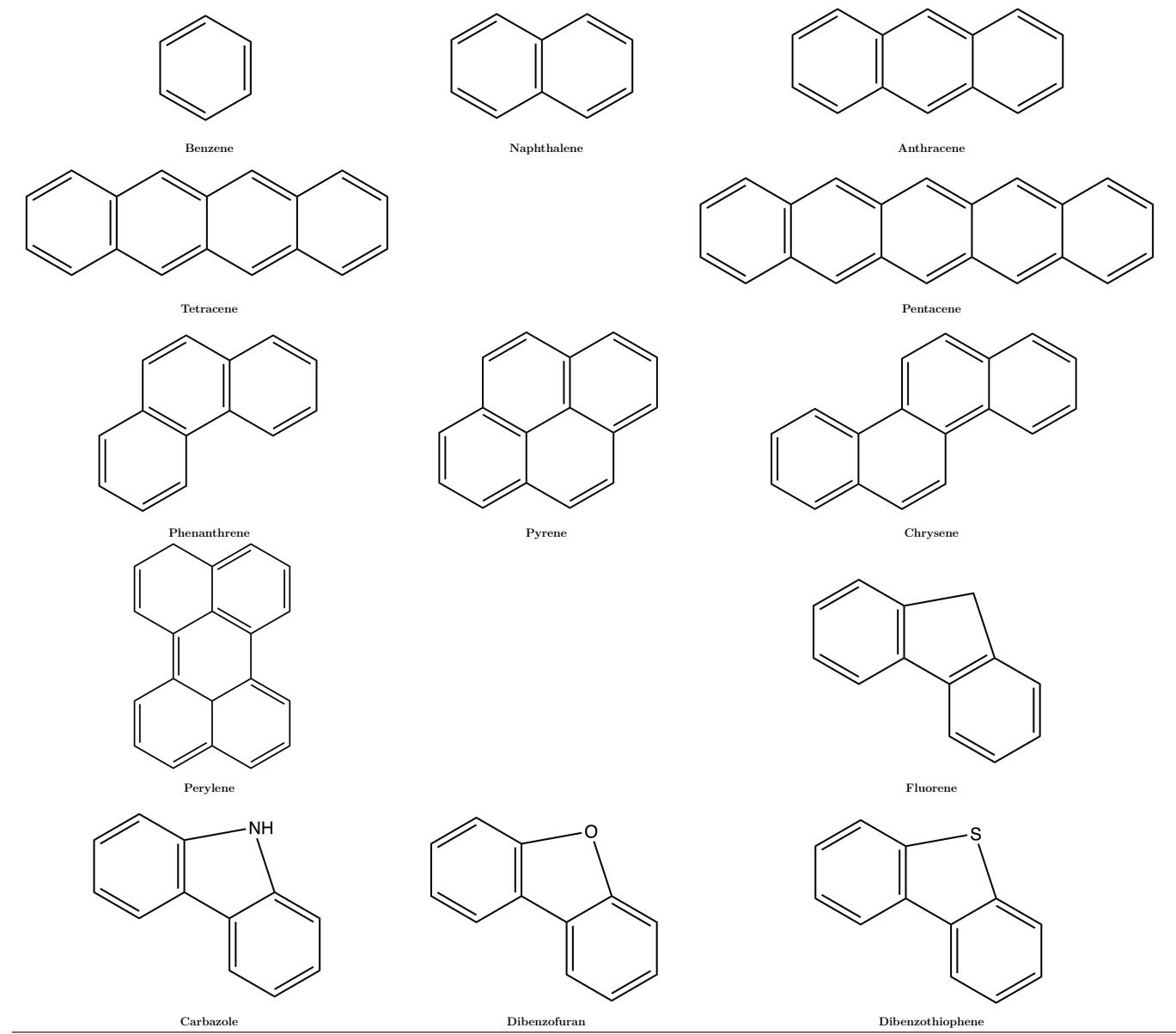


Figure 6.19 – Molecular structures of studied molecules

## Chapter 7

**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:<sup>1</sup>

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.*<sup>2</sup>

These simulations are based on recent X-ray (SAXS) and neutron (SANS) small-angle scattering experiments from Eyssautier *et al.*<sup>3</sup> 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.

---

<sup>1</sup> Guevara Level *et al.* *Energy & Fuel*, **30**(6), 4758-4766, 2016

<sup>2</sup> *JACS*, **2015**, 137, 31, 9870

<sup>3</sup> *The J. of Phys. Chem. B*, **2011**, 115, 21, 6827

## 7.1 Introduction

Asphaltenes are the oil's insoluble phase in light *n*-alkanes and soluble in aromatic solvents like benzene or toluene[229]. Their presence is critical and can cause problems in several fields of petroleum utilization, from oil recovery and transportation to refining.[230] 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.) [231].

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.[232] 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[233].

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  $\pi$ -stacking of the conjugated cores, with heights varying from 1 to 10 nm [6, 234].

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.[235, 236, 237] 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[237, 238].

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-licpentyl)-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).

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

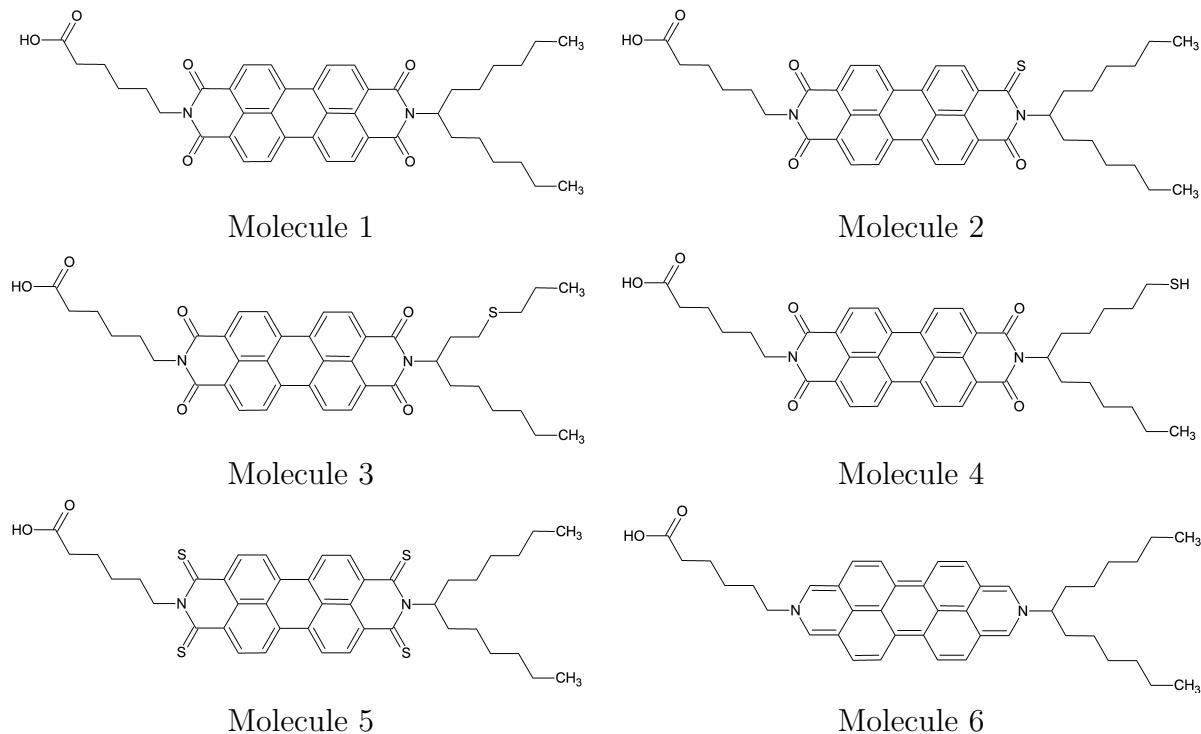


Figure 7.1 – C5Pe molecular models of asphaltenes.

envisioned: 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[239] 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 7.1, are originally labeled PA3 and CA22 after Schuler *et al.*,[239] both of them being of the continental type.[232, 229, 176] 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.

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.[240, 241] In the case of CA22, nitrogen was chosen

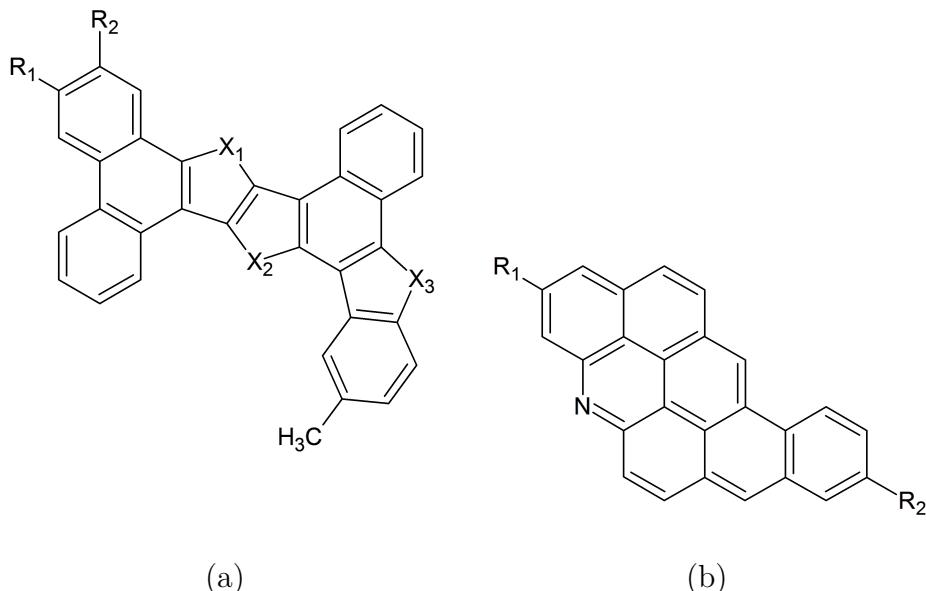


Figure 7.2 – Two different asphaltene structures found in the crude petroleum: (a) PA3 and (b) CA22.  $R_1$  and  $R_2$  are the  $n$ -hexyl lateral chains which can have  $-COOH$  or  $-CH_3$  as chain-ends.  $X_n$  ( $n=1-3$ ) can be O, S, or NH atom groups.

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 7.1. 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, [242, 243, 244] which indexes might be of use for further referring.

The DBE in this table was calculated as:

$$DBE = C - \frac{H}{2} + \frac{N}{2} + 1 \quad (7.1)$$

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;

Table 7.1 – 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 <sup>-1</sup> )
<b>A11</b>	O	O	O	COOH	COOH	45	27	O <sub>7</sub>	690.8
<b>A12</b>	O	O	O	CH <sub>3</sub>	COOH	45	26	O <sub>5</sub>	660.8
<b>A13</b>	O	O	O	CH <sub>3</sub>	CH <sub>3</sub>	45	25	O <sub>3</sub>	630.8
<b>A21</b>	S	S	S	COOH	COOH	45	27	S <sub>3</sub> O <sub>4</sub>	740.0
<b>A23</b>	S	S	S	CH <sub>3</sub>	CH <sub>3</sub>	45	25	S <sub>3</sub>	679.0
<b>A31</b>	NH	NH	NH	COOH	COOH	45	27	N <sub>3</sub> O <sub>4</sub>	687.8
<b>A33</b>	NH	NH	NH	CH <sub>3</sub>	CH <sub>3</sub>	45	25	N <sub>3</sub>	627.8
<b>A41</b>	O	O	S	COOH	COOH	45	27	O <sub>6</sub> S	706.8
<b>A43</b>	O	O	S	CH <sub>3</sub>	CH <sub>3</sub>	45	25	O <sub>2</sub> S	646.8
<b>A51</b>	O	S	S	COOH	COOH	45	27	O <sub>5</sub> S <sub>2</sub>	723.9
<b>A53</b>	O	S	S	CH <sub>3</sub>	CH <sub>3</sub>	45	25	OS <sub>2</sub>	662.9
<b>AAA</b>	-	-	-	COOH	COOH	37	22	N <sub>1</sub> O <sub>4</sub>	555.6
<b>AAB</b>	-	-	-	COOH	CH <sub>3</sub>	37	21	N <sub>1</sub> O <sub>2</sub>	525.7
<b>AAC</b>	-	-	-	CH <sub>3</sub>	CH <sub>3</sub>	37	20	N <sub>1</sub>	501.7

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,[228] 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.[245, 246, 247]

## 7.2 Simulation Method

The MD simulations were carried out using the GROMACS 5.1.1 software package.[248, 249, 250] The GROMOS96 force field[251] with the 53a6 parameter set[252] 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.[238, 253, 254, 255]

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[256] 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  $0.88 \text{ g.cm}^{-3}$ , in good agreement with the  $0.86 \text{ g/cm}^{-3}$  value extracted from NIST WebBook (National Institute of Standards and Technology).[257]

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 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.[258] 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 [136, 259, 260, 261, 262] using Orca 3.0.3 software[263].

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 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  $0.35 \text{ nm}$  in the  $\pi$ -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  $298 \text{ K}$  and  $1 \text{ bar}$  pressure during  $3 \text{ ns}$  using the V-rescale thermostat[264]<sup>4</sup> and Berendsen barostat.[265] The production phase consisted of  $60 \text{ ns}$  of dynamics<sup>5</sup> at  $298 \text{ K}$  and  $1 \text{ bar}$  of pressure also under the NPT ensemble but using the Nosé-Hoover thermostat[266, 267] and the Parrinello-Rahman[268] 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  $12 \text{ \AA}$  was used for both Coulomb and Van der Waals interactions. Periodic boundary conditions were thoroughly applied in the three directions. The Verlet integration algorithm[269] has been used with a  $2 \text{ fs}$  integration step. The electrostatic interactions were computed using the Particle-Mesh Ewald[270] summation method with a fast Fourier transform grid spacing of  $1.6 \text{ \AA}$  to account for long-range electrostatic interactions of the system. All the bond lengths have been constrained using the LINCS algorithm[271] 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  $12 \text{ \AA}$  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

<sup>4</sup>For 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.

<sup>5</sup>150 ns for the C5Pe derivatives systems.

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.

## 7.3 Results and Discussion

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

### 7.3.1 C5Pe-derivatives systems

Molecule 1 was already studied by MDS by Teklebrhan and co-workers[255, 238] and Gao and co-workers,[237] with results similar to the ones found in this study, *i.e.*, the dimers of these molecules are stable and keep interacting throughout the simulation. Figures 7.3, 7.4, 7.5 and 7.6 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 7.6. 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.

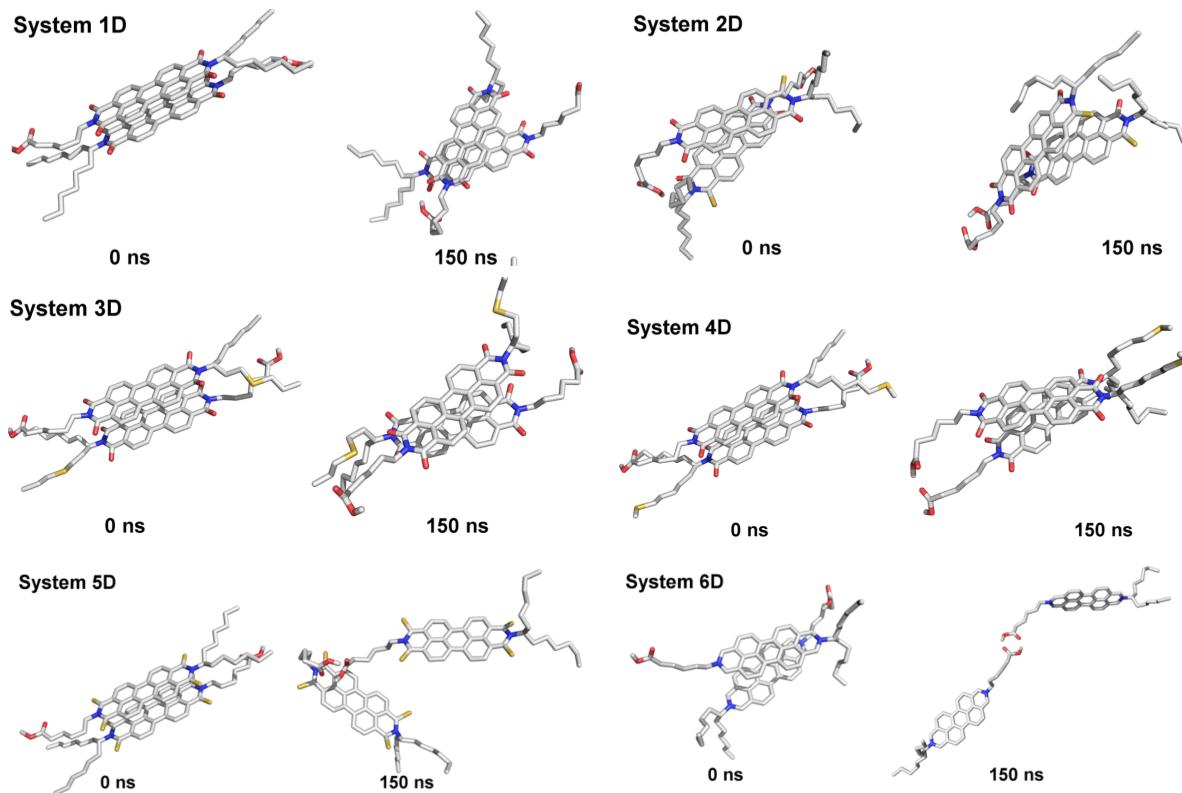


Figure 7.3 – Snapshots of the initial configuration and the final frame of the dimers simulations. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.

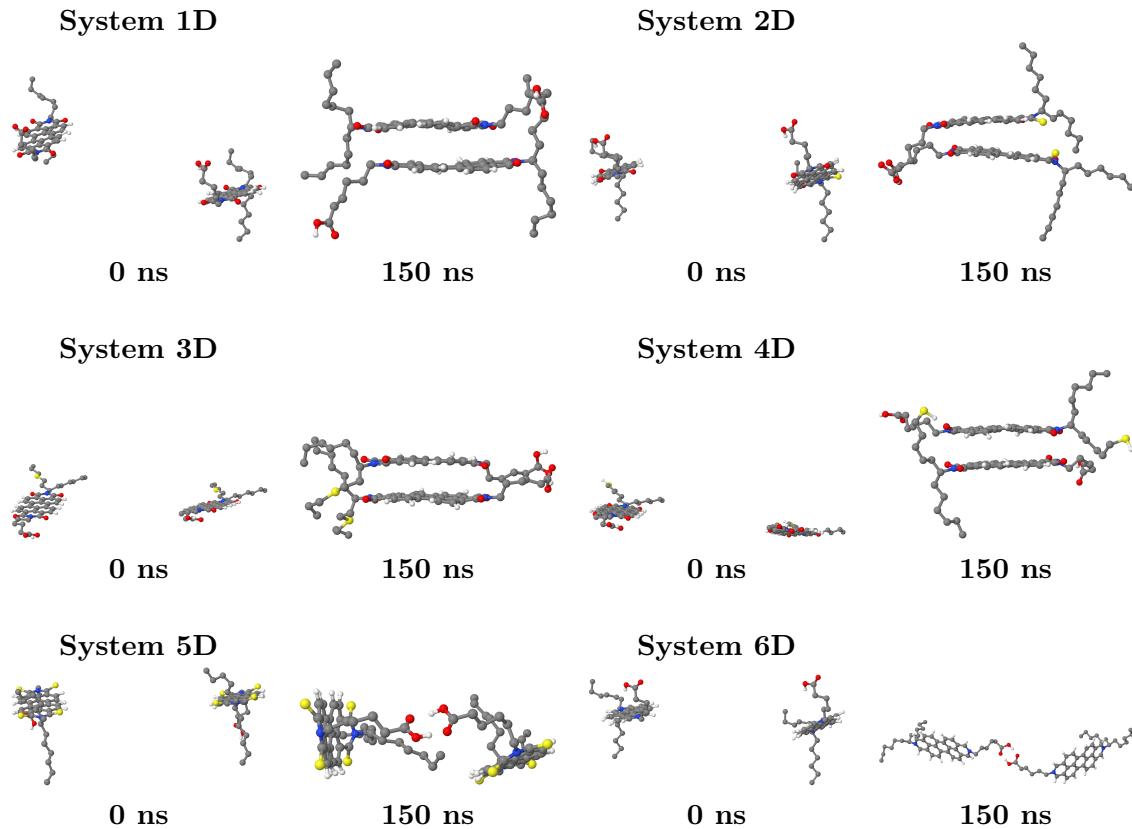


Figure 7.4 – Snapshots of the initial configuration and the final frame of the dimers simulations for the random disposition case. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.

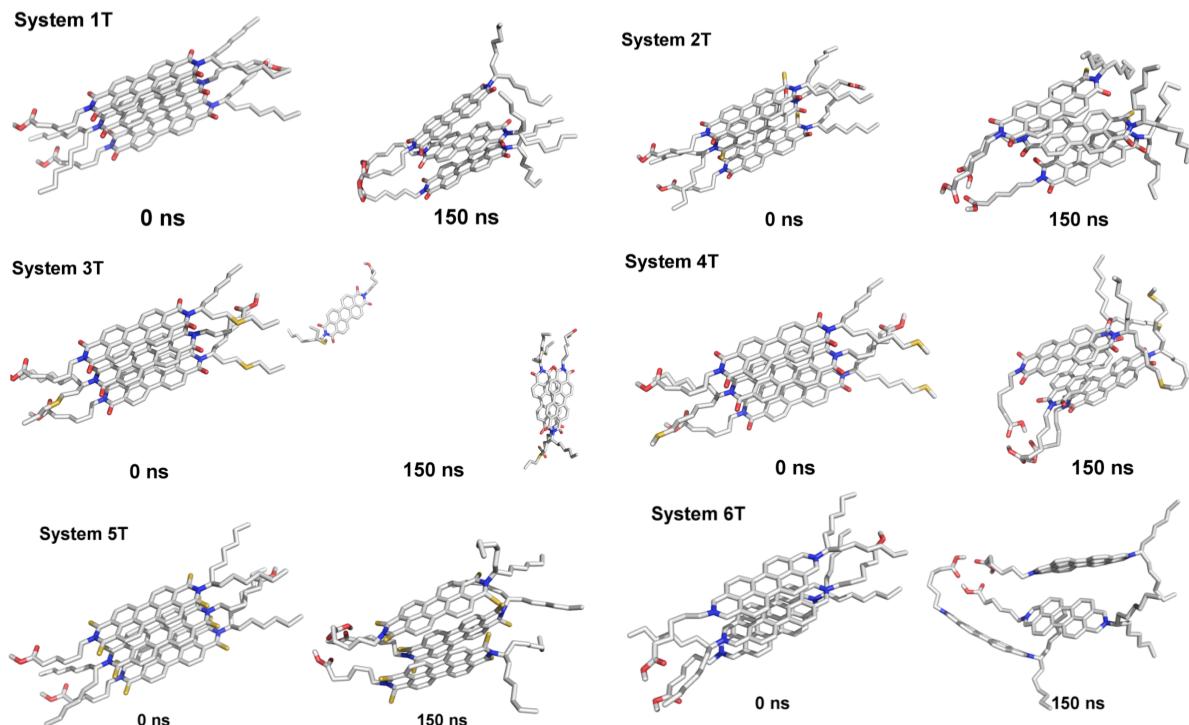


Figure 7.5 – Snapshots of the initial configuration and the final frame of the trimers simulations. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.

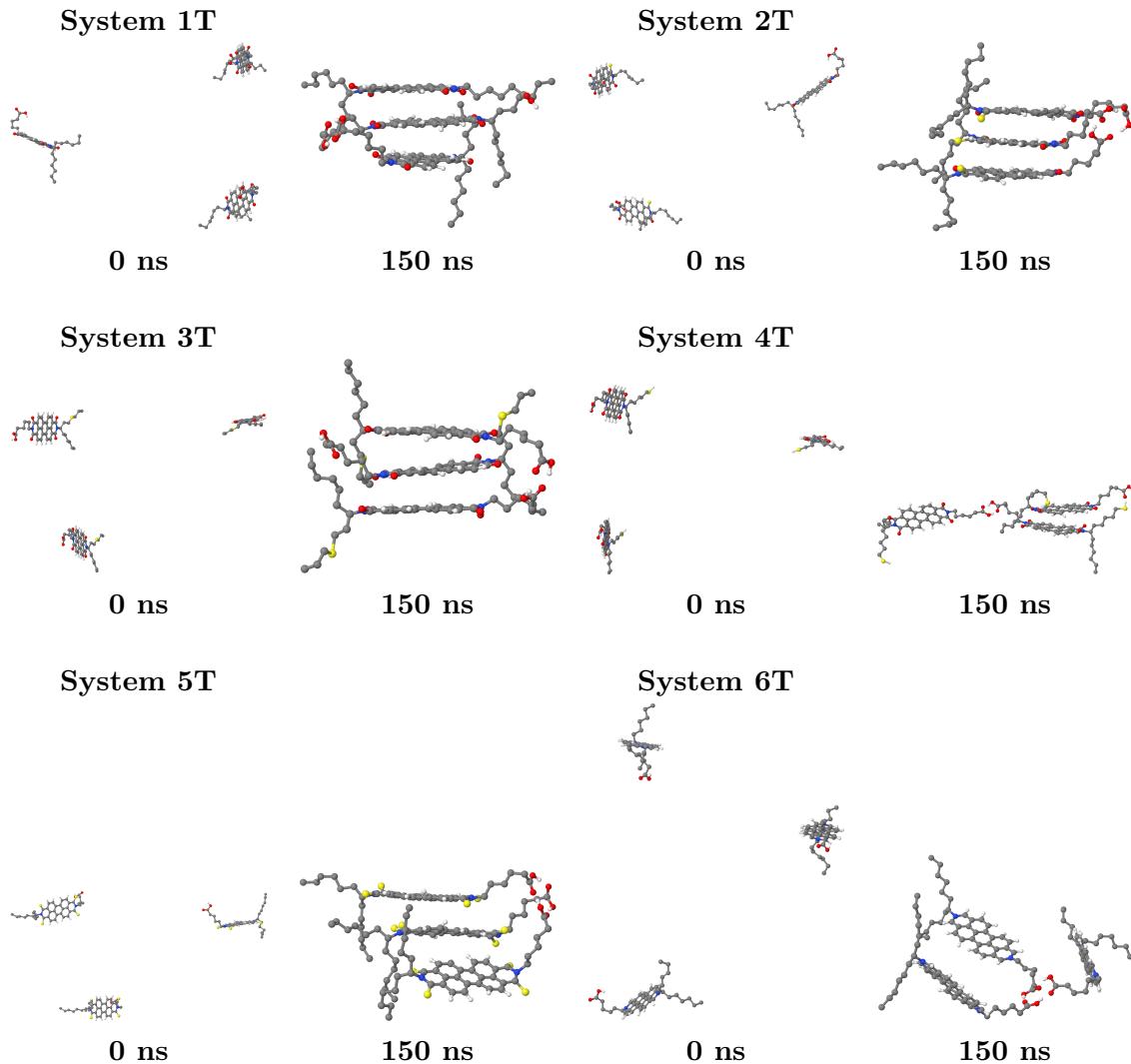


Figure 7.6 – Snapshots of the initial configuration and the final frame of the trimers simulations for the random disposition case. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.

**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 7.7 and 7.8). 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 7.7). The molecules are close to each other ( $< 0.4$  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.

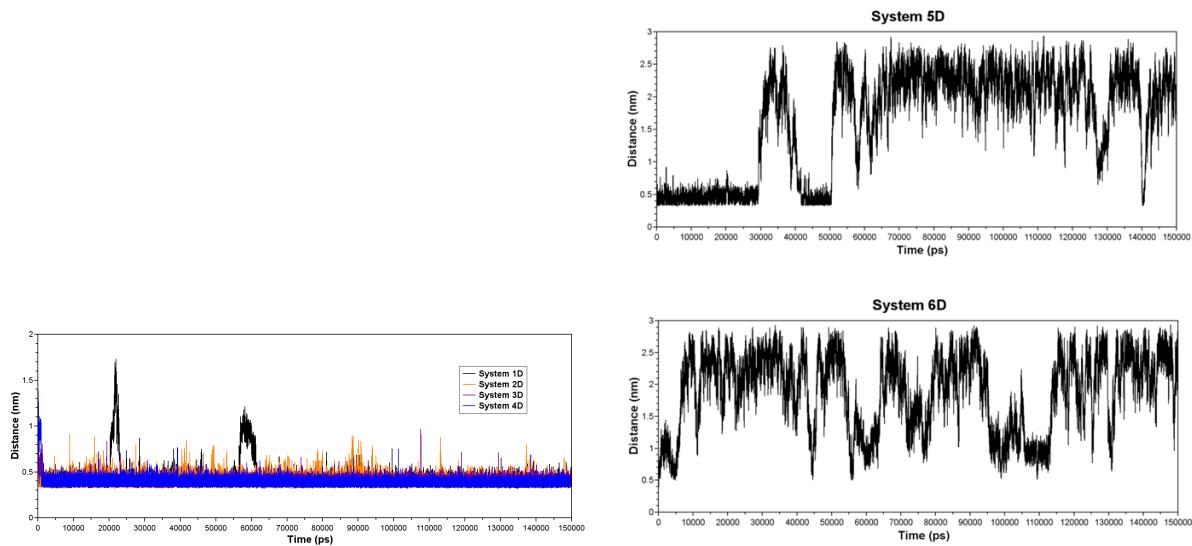


Figure 7.7 – Distances between monomers for “organized” starting point.

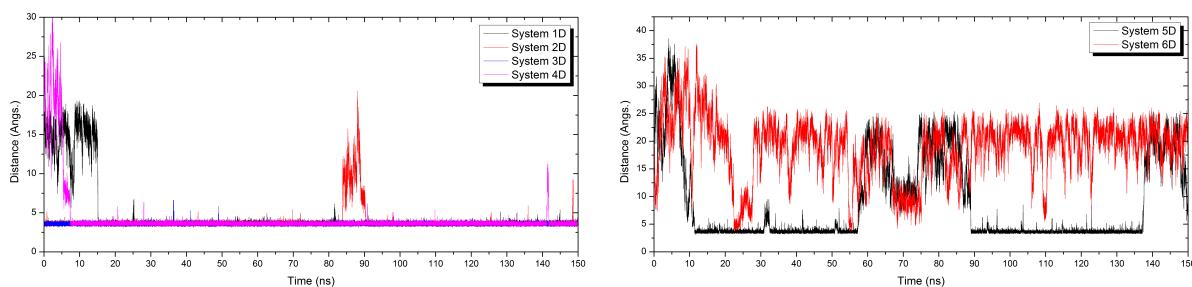


Figure 7.8 – Distances between monomers for “random” starting point.

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 7.9. 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.[255] 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 7.10.

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.[255] 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,[6] which suggest a height of 0.67 nm. Systems 2T and 4T (after 33 ns) exhibit similar results.

As it can be seen from Figures 7.11, 7.12, 7.13 and 7.14, 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 7.12 and 7.14). For all trimer systems, the planar configuration is more likely than others, namely the T-shape. System 3T, in the “organized” starting point (Figure 7.12) and system 4T in the random starting point (Figure 7.14) 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 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 ( $<$

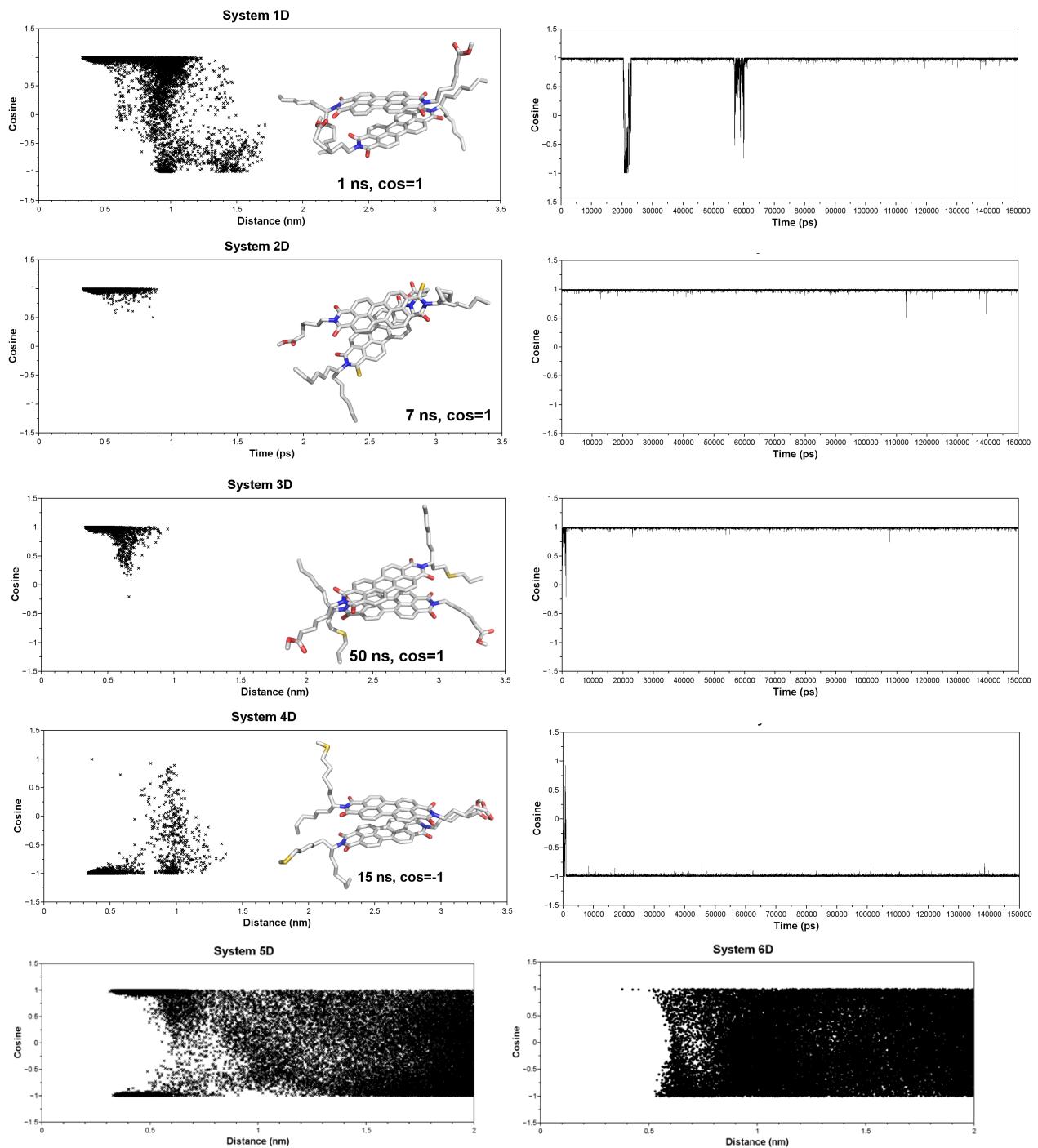


Figure 7.9 – Cosine of angle between aromatic cores in function of distance between dimers monomers for the organized starting point. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.

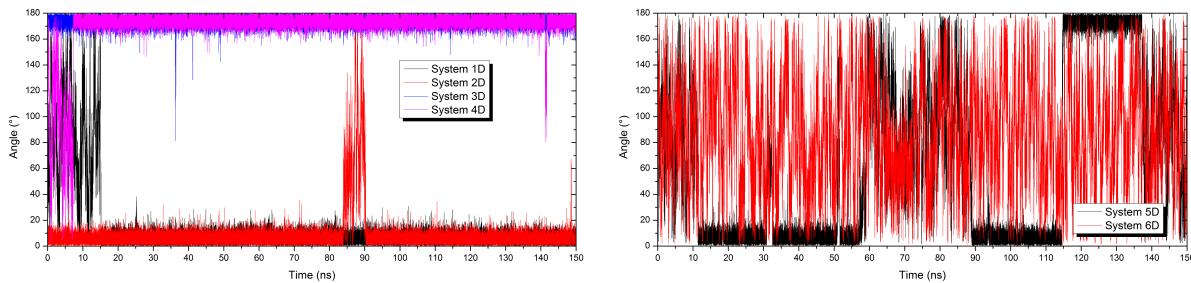


Figure 7.10 – Cosine of angle between aromatic cores in function of distance between dimers monomers for the random starting point. Carbon atoms are in grey, oxygen in red, nitrogen in blue and sulfur in yellow. Hydrogen atoms are omitted for clarity.

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,[272] sometimes strong enough to lock conformations in place. These graphics can be found in Figure 7.16 for the dimer and 7.17 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

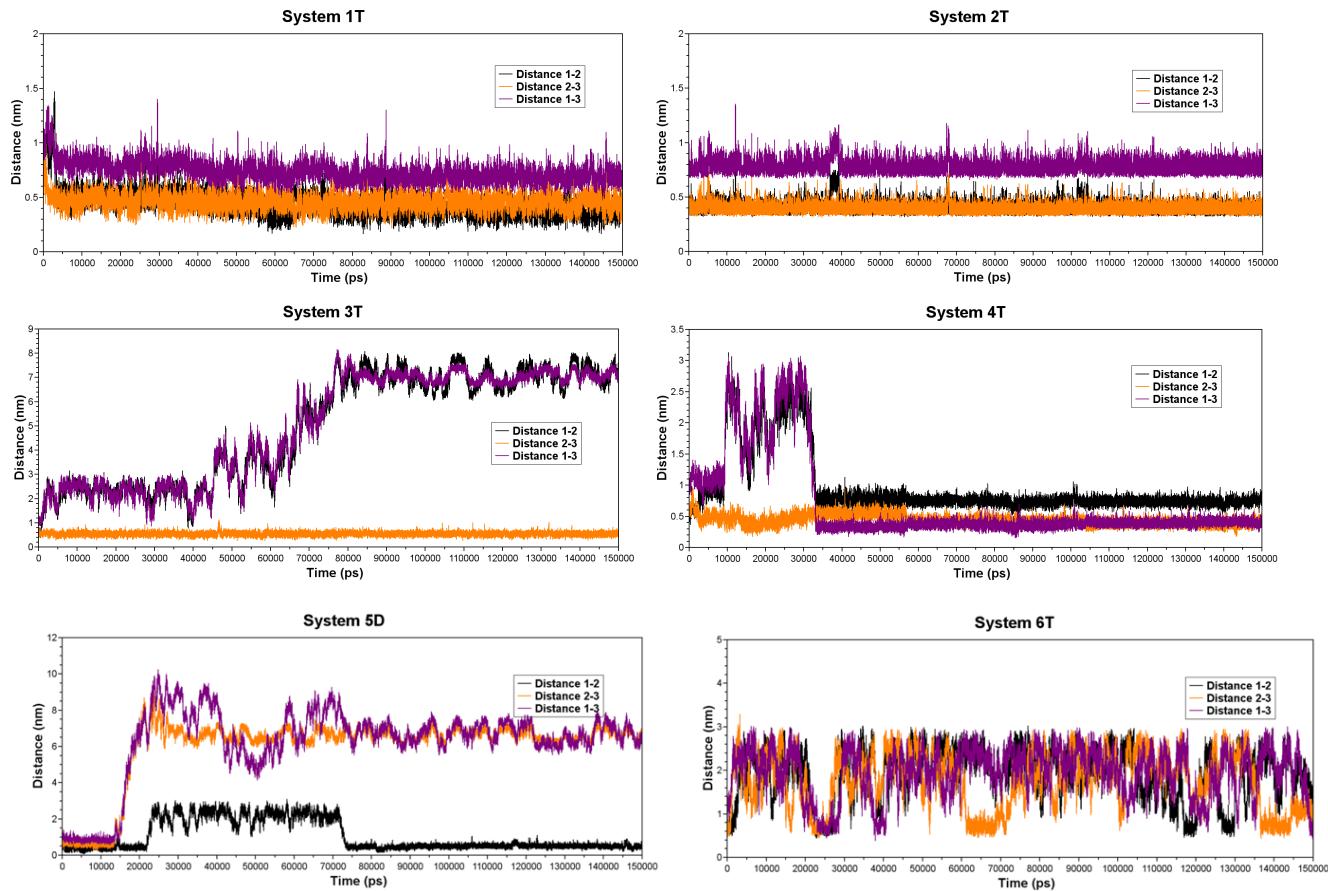


Figure 7.11 – Distances between monomers for “organized” starting point.

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 previous 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 7.9), 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 7.10).

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.

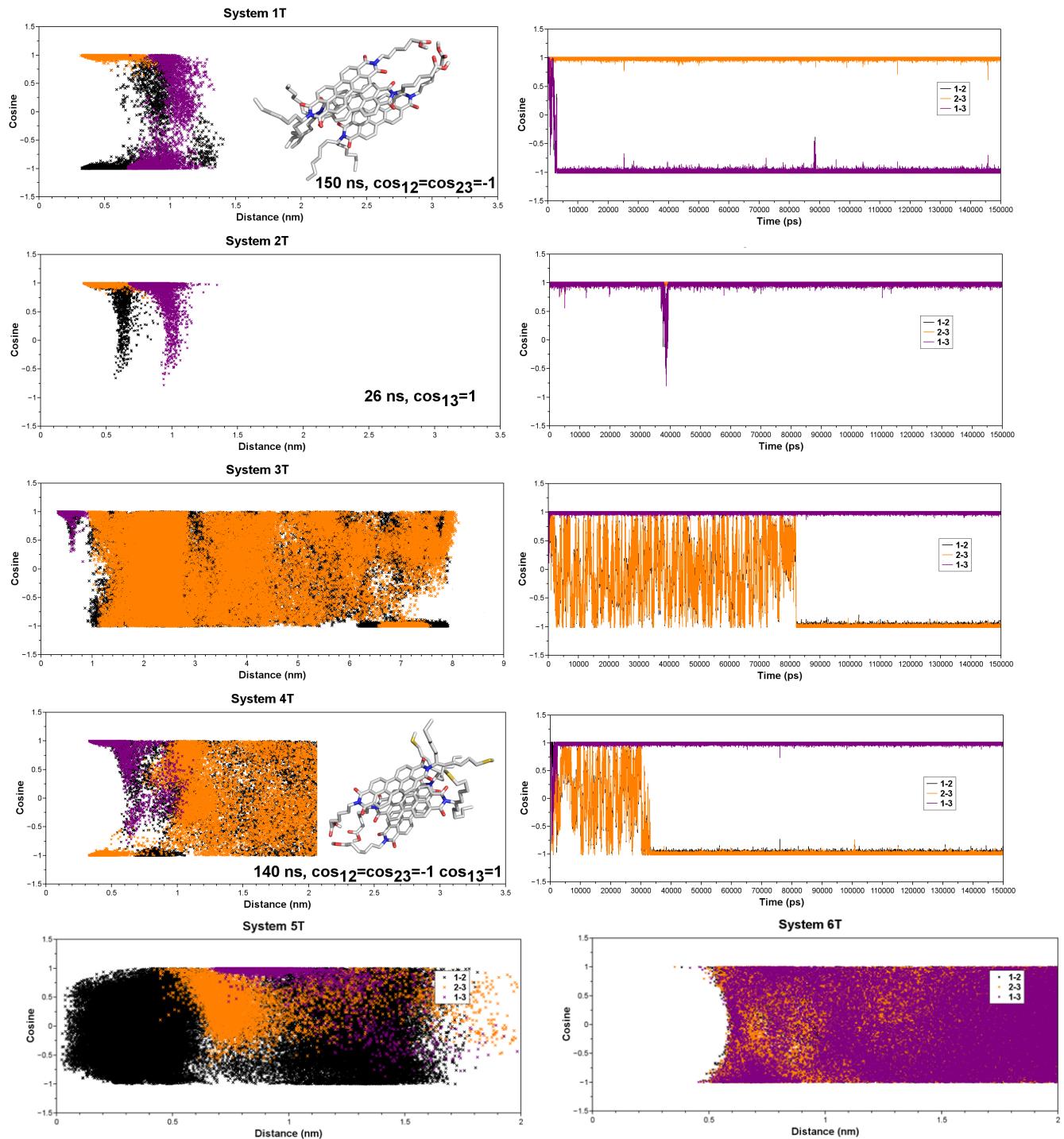


Figure 7.12 – Angle between monomers for “organized” starting point.

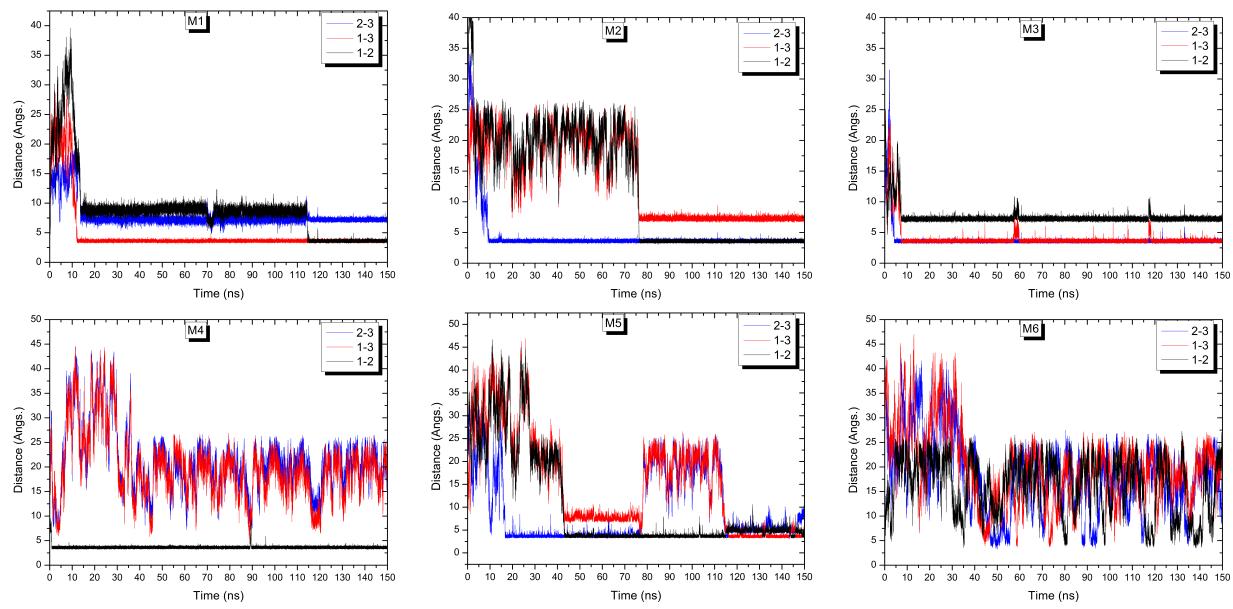


Figure 7.13 – Distances between monomers for “random” starting point.

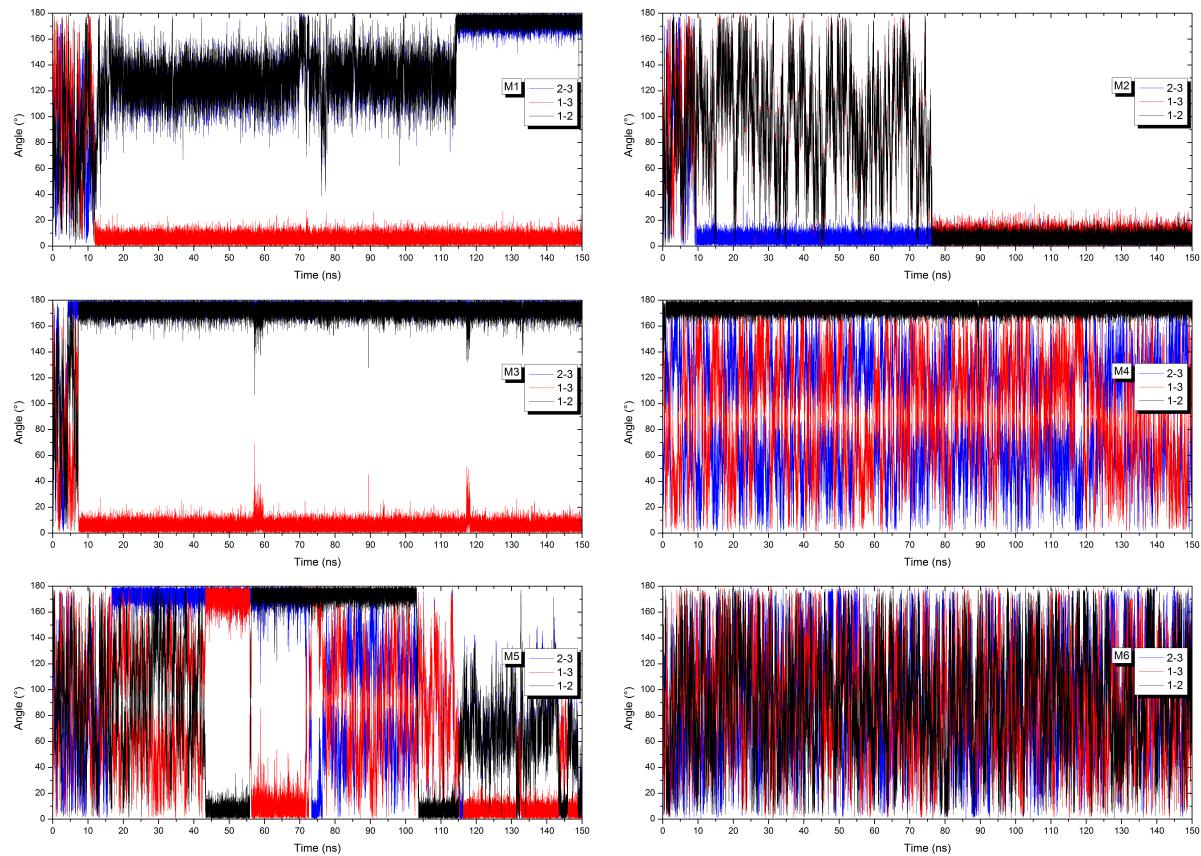


Figure 7.14 – Angle between monomers for “random” starting point.

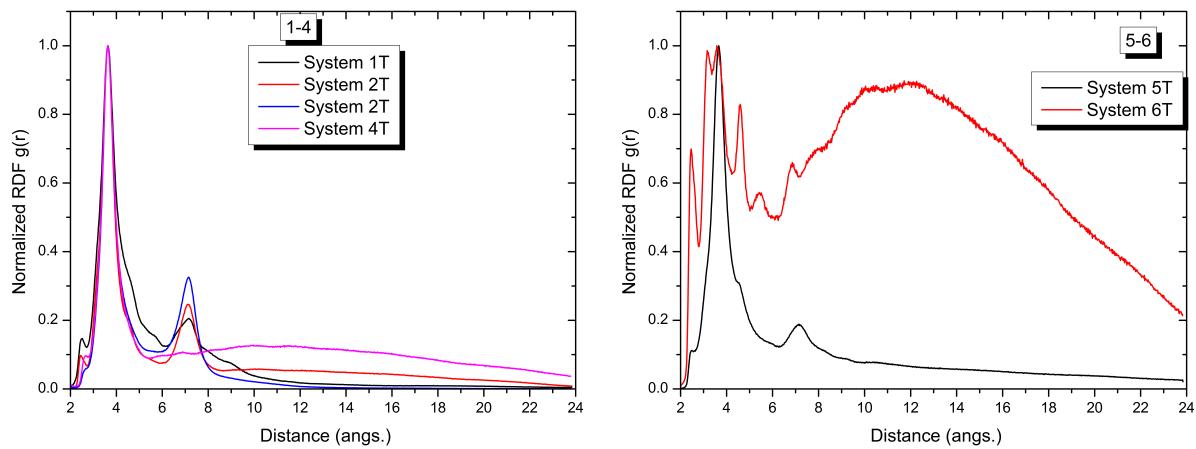


Figure 7.15 – Normalized Radial Distribution Functions for the trimer systems.

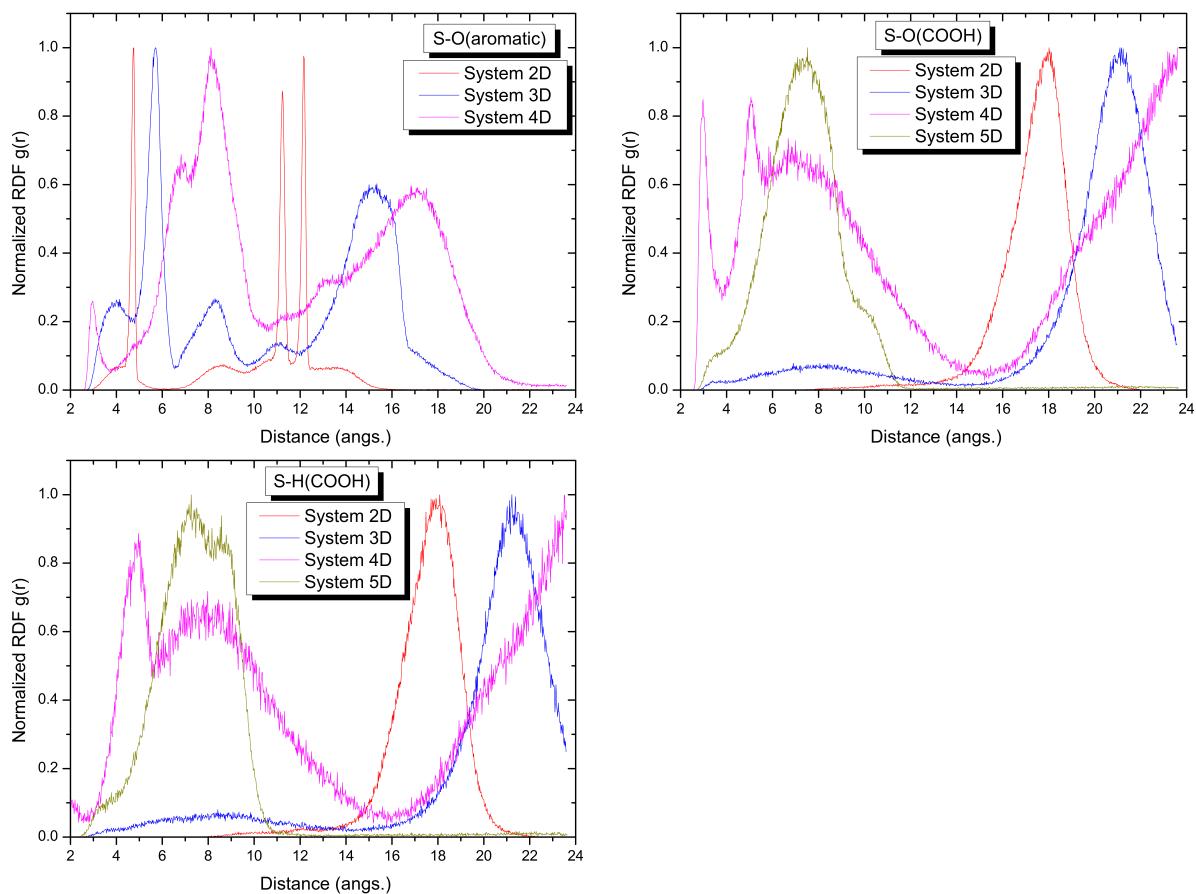


Figure 7.16 – Normalized Radial Distribution Functions for the S–X pairwise interactions, where X can be the oxygen atom of the aromatic core, the oxygen atom of the carboxylic group (COOH) or its hydrogen for the dimer system.

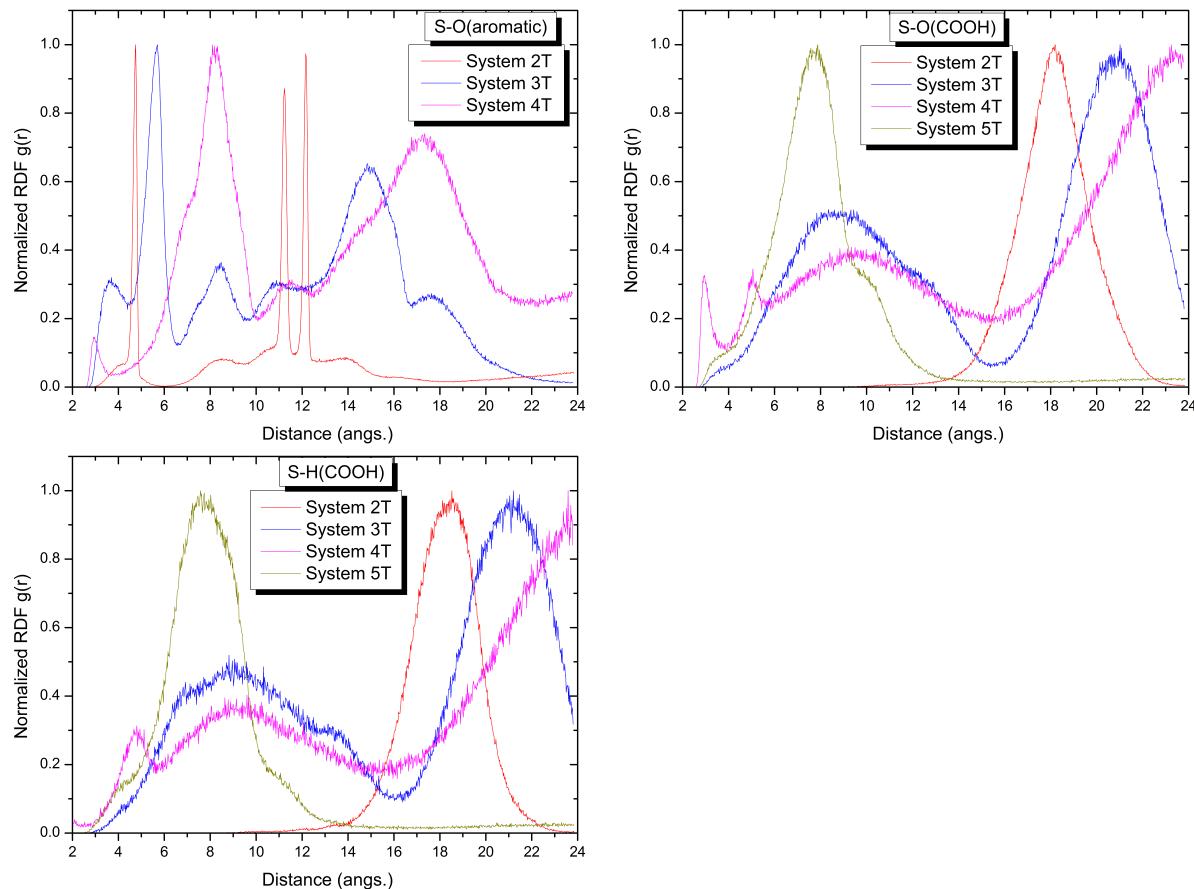


Figure 7.17 – Normalized Radial Distribution Functions for the S–X pairwise interactions, where X can be the oxygen atom of the aromatic core, the oxygen atom of the carboxylic group (COOH) or its hydrogen for the trimer system.

### 7.3.2 Schuler's-derivatives systems

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

#### 7.3.2.1 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 7.18, one can find the normalized RDF functions of these systems under 298 K and 1 bar conditions.

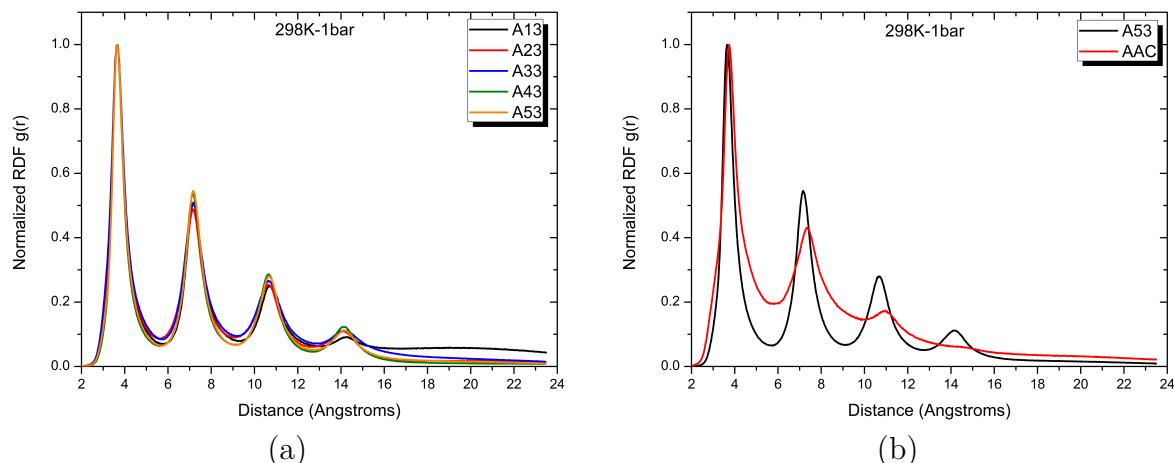


Figure 7.18 – Aggregation mechanism with  $-\text{CH}_3$  ended chains: (a) effect of the heteroatom (A13, A23, A33, A43, A53) and (b) effect of the asphaltene type (A53/AAC). Each peak of the RDFs corresponds to a neighbor molecule in the  $\pi$ -stacking structure presented in Figure 7.3.

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  $\sim 3.9$  Å stands for the interactions with the first neighbors generated by the  $\pi$ -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 7.8 Å.

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 7.19.

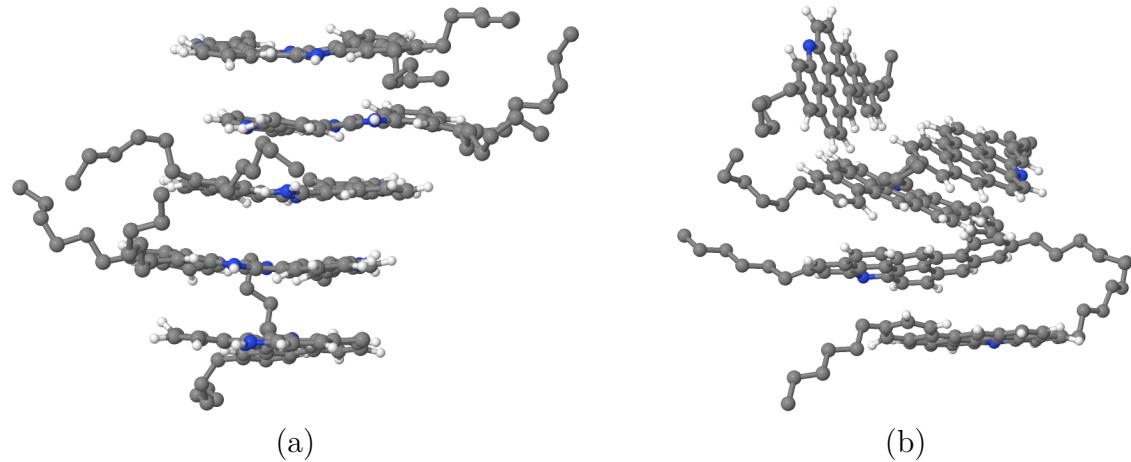


Figure 7.19 – 60 ns-snapshots of 5 molecules of each asphaltene in toluene at 298K and 1 bar conditions: (a) A33 and (b) AAC. Carbon atoms are presented in gray and nitrogen atoms in blue. Aromatic or polar hydrogen atoms are presented in white. Toluene molecules are omitted for clarity.

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:

$$\langle E_i \rangle = -(\langle E_d^p \rangle - 2 \cdot \langle E_m^p \rangle) \quad (7.2)$$

$\langle E_d^p \rangle$  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.*[236] The interaction energies of PA3 and CA22 dimers in vacuum, reported in Table 7.2 (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<sup>-1</sup> ( $3.8 \times 10^{-19}$ J) for the PA3 class of molecules. For the CA22 molecule type, the interaction energy is lowered by around 10 kcal.mol<sup>-1</sup>. The out-of-diagonal values reported in this Table will be discussed in the next sections.

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 C<sub>6</sub> and C<sub>12</sub> 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.

	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 7.2 – Interaction energies ( $\langle E_i \rangle$ ) in kcal.mol $^{-1}$  calculated for PA3- and CA22-type dimer molecules.

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.

### 7.3.2.2 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 7.20.

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  $\pi$ -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  $\pi$ -stacking aggregation. This can be visualized in the snapshots of A31, AAA, A33 and AAC presented in Figures 7.19 and 7.21.

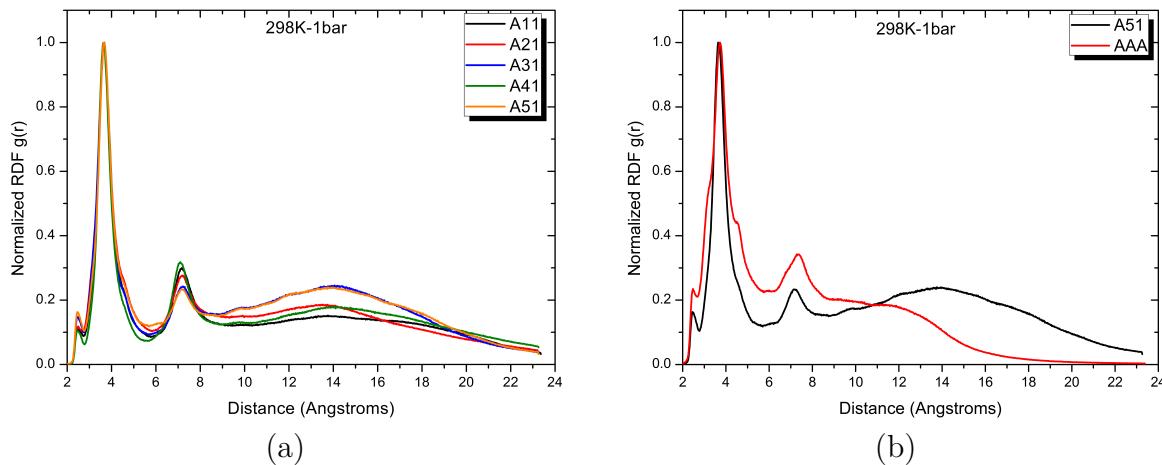


Figure 7.20 – Aggregation mechanism with  $-COOH$  ended chains: (a) effect of the heteroatom (A11, A21, A31, A41, A51) and (b) effect of the asphaltene type (A51/AAA). Each peak of the RDFs corresponds to a neighbor molecule in the  $\pi$ -stacking structure presented in Figure 7.19.

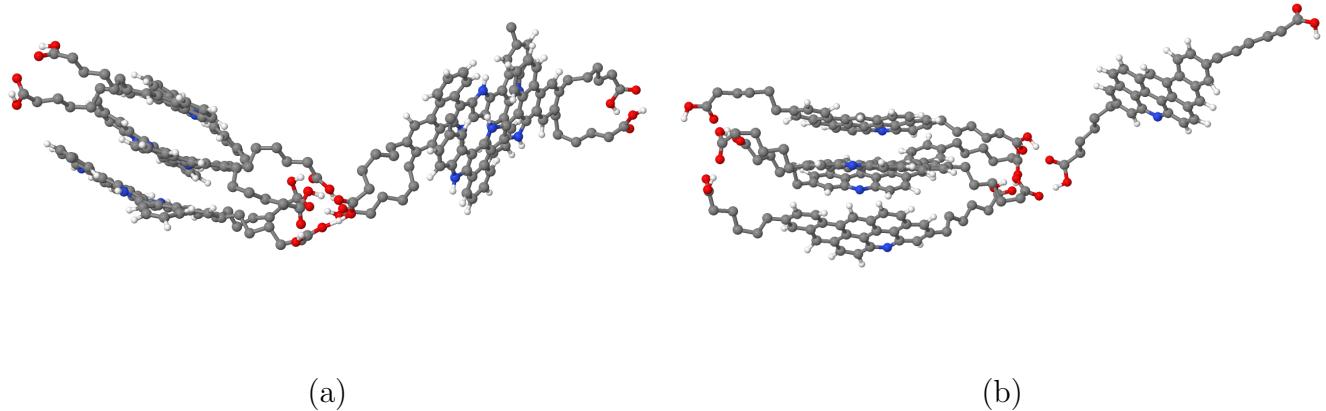


Figure 7.21 – 60 ns-snapshots of 5 molecules of each asphaltene in toluene at 298K and 1 bar conditions: (a) A31 and (b) AAA.

Interestingly, the formation of H-bonds are stable enough under these conditions to avoid the self-stacking as it is the case of the  $-CH_3$ -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  $22 \text{ kcal.mol}^{-1}$  for each pair of  $-COOH$  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  $75 \text{ kcal.mol}^{-1}$ , 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,[15] 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  $\pi$ -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.*[232, 229, 176]

### 7.3.2.3 Aggregation within heterogeneous boxes

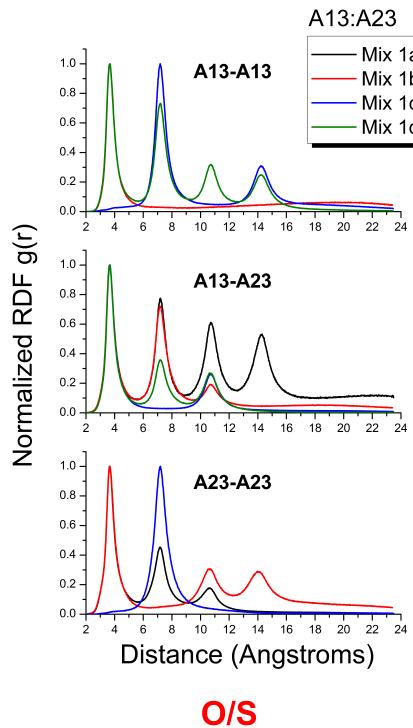
Next, we will present the results of aggregation for boxes containing different asphaltenes, be it the same family or not. For the simulation of the mixtures between asphaltenes of the same family, we continued on using the same conditions but varied the relative concentration between the constituents.

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 7.22. A13:A33 and A23:A33 mixtures can be found in Figure S1.

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



O/S

Figure 7.22 – 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.

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 7.2 (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.

## 7.4 Conclusions

Several asphaltene structural models have been proposed in the literature, such as the C<sub>5</sub>Pe molecule. Based on experimental work of Eyssautier and co-workers,[6] 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.[239] 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<sup>-1</sup>. 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.

# **Conclusion Générale**

Le cœur du présent travail de thèse consistait en une caractérisation théorique des spectres IR et Raman d'une série d'hétérocycles aromatiques définis comme modèles moléculaires en raison de leur représentativité de la diversité des systèmes asphalténiques. Ces derniers étant un enjeu conséquent pour l'industrie du pétrole, un effort de recherche se concentre en effet sur la compréhension de leur structure chimique et des interactions qui en découlent, responsables des propriétés physico-chimiques observées (agrégation en premier lieu, mais également précipitation et flocculation).

Pour répondre au double objectif de la résolution de la nature chimique des asphaltènes, d'une part, et des interactions à l'origine du phénomène d'agrégation, d'autre part, J. Shaw et K. Michaelian ont proposé un schéma de travail reposant sur la spectroscopie infra-rouge. Ce travail de thèse s'inscrit dans la mise en œuvre de leur méthodologie, en fournissant des données théoriques fiables à même de guider les expérimentateurs dans l'interprétation et la caractérisation de leurs spectres. L'idée générale de nos travaux est de caractériser les signatures vibrationnelles des modes inter et intra-moléculaires, de même que leurs modifications sous l'effet de l'environnement ou encore de la position et du nombre d'hétéroéléments. Pour peu que l'on parvienne à dégager de nos calculs des tendances, sinon des conclusions, quant à la fréquence et à la position des signatures vibrationnelles de ces modèles, en fonction de l'environnement et de la composition chimique, ces tendances serviront de guide à l'expérimentateur pour l'analyse des spectres réels.

La mise en place d'une stratégie calculatoire à même de répondre à ces problématiques s'est révélée ardue, du fait de la dimension des systèmes étudiés et de la nécessité d'incorporer à nos calculs des interactions longue portée, encore mal décrites par les méthodes de type DFT. Une part conséquente du présent travail a donc été consacrée au choix des outils méthodologiques les plus adaptés, ainsi qu'au choix d'approximations judicieuses, permettant de diminuer le coût calculatoire sans entamer la qualité des simulations. Sur les sept chapitres regroupés dans ce manuscrit, trois sont ainsi consacrés au détail des méthodes et approximations employées, tandis qu'un chapitre à part entière présente les résultats de calculs préliminaires, menés sur des molécules test, dans le but de démontrer la pertinence de la stratégie calculatoire mise en place et son aptitude à traduire, avec une précision satisfaisante, les propriétés ciblées. La part importante de ces chapitres préliminaires se veut représentative de l'effort méthodologique réalisé tout au long de ce travail de thèse, essentiel pour répondre au sujet ambitieux que représente la structure chimique des asphaltènes. Ces réflexions méthodologiques s'étendent à toutes les échelles de calcul envisagées, puisque les problématiques poursuivies sont multiples. Rappelons qu'il s'agit de déterminer l'influence

- du nombre et de la position des hétéroéléments,
- de l'environnement (état gazeux, liquide ou solide),
- et des interactions à l'origine des processus d'agrégation

sur les modes de vibration des systèmes asphalténiques.

Par soucis de clarté, nous avons fait le choix d'associer à chacune de ces problématiques un chapitre spécifique, mais il est bien évident que les résultats obtenus, bien qu'à différents niveaux calculatoires (moléculaire, solide ou dynamique moléculaire), demeurent intimement liés. Ainsi, notre chapitre cinq se concentrerait sur l'influence des hétéroatomes, et présentait

donc des résultats de calculs moléculaires. Dans cette partie, nos calculs portaient sur une série de six motifs moléculaires, contenant ou non une proportion d'hétéroéléments (N, S ou O), d'abord étudiés isolément puis sous forme de dimères (de sorte à simuler l'effet de l'interaction). Il est à noter que nos résultats se révèlent en bon accord avec les spectres expérimentaux, démontrant une fois encore la pertinence de notre prise en compte des anharmonicités électrique et mécanique pour la REVS. Toutefois, nos calculs ont révélé que les signatures vibrationnelles caractéristiques des modes intermoléculaires associés aux dimères se trouvaient aux bas nombre d'ondes ( $0\text{--}150\text{ cm}^{-1}$ ), soit dans une zone traditionnellement mal décrites par les calculs moléculaires. Ces conclusions se déduisent de la confrontation des spectres vibrationnels des monomères et des dimères, qui montrent, pour les systèmes contenant de l'oxygène, du soufre et du carbone, une correspondance parfaite pour l'ensemble des bandes, à l'exception des bandes induites par les interactions intermoléculaires, que l'on voit apparaître aux bas nombre d'ondes sur les spectres des dimères. Le cas des motifs contenant de l'azote se révèle plus spécifique, puisque monomères et dimères présentent des différences dans la signature des modes de vibration intramoléculaires dans la région en-dessous de  $600\text{ cm}^{-1}$ . De fait, les systèmes contenant de l'azote peuvent être plus aisément caractérisés en tant que systèmes isolés ou en tant que systèmes en interaction, au travers de l'analyse de deux zones spectrales distinctes. Pour l'ensemble des autres systèmes, la conclusion quant à la présence ou l'absence d'interactions se déduit d'une analyse aux bas nombres d'onde. Du point de vue structural, et non plus strictement vibrationnel, les configurations envisagées pour les dimères modèles étudiés révèlent qu'il n'existe pas d'orientation privilégiée des hétéroéléments, de sorte que l'interaction prépondérante demeure le  $\pi$ -stacking.

Le passage de l'état gazeux, simulé au travers de calculs moléculaires, à l'état solide, présenté dans le chapitre six, avait pour objectif de comprendre l'effet de l'environnement en caractérisant la modification des signatures des modes de vibration précédemment analysés. Les modes de réseau calculés, par la prise en compte d'un environnement plus représentatif des systèmes réels, ont ainsi été comparés aux modes de vibration intermoléculaires présentés au chapitre cinq. Cette comparaison révèle que, pour l'ensemble des systèmes aromatiques considérés, une bande active caractéristique des interactions intermoléculaires se retrouve systématiquement aux alentours de  $100\text{ cm}^{-1}$ , quels que soient la taille du système, la nature du noyau aromatique (benzénique ou non) ou encore le type d'hétéroélément présent.

Enfin, les simulations en dynamique moléculaire menées dans le but d'analyser l'influence des hétéroéléments et des proportions de mélanges des espèces modèles sur le comportement de nano-agrégats, ont fait l'objet d'une analyse au chapitre sept. De ces modélisations, il ressort que les noyaux polyaromatiques, au cœur de la structure chimique des asphaltènes, jouent un rôle prépondérant dans l'agrégation de ces systèmes. La position des hétéroatomes au sein des modèles se révèle influer sur les interactions de  $\pi$ -stacking à l'origine du phénomène. Toutefois, bien que la proportion d'hétéroatomes soit conséquente au sein des asphaltènes – ce qui porte à penser qu'ils possèdent une influence remarquable dans les processus de nano-agrégation –, ce sont la taille et la nature des ramifications qui jouent véritablement un rôle décisif dans la description du phénomène.

En conclusion générale, les modes de vibration caractérisés tout au long de ce travail permettent de fournir aux expérimentateurs des critères d'analyse pour l'interprétation des spectres IR et Raman. Le résultat principal de la présente étude réside dans l'interprétation des modes

de vibration intermoléculaires, caractéristiques des interactions de type  $\pi - \pi$  et  $\pi - H$ , identifiables aux bas nombres d'onde (soit aux alentours de  $100\text{ cm}^{-1}$ ). La présence, sur les spectres expérimentaux, de bande(s) dans cette zone révèle sans ambiguïté l'existence de systèmes en interaction. Au-delà de ce résultat, au demeurant directement utilisable par les spectroscopistes, l'ensemble de nos calculs représente une avancée vers la compréhension de la structure et des propriétés des asphaltènes. Notamment, la méthodologie calculatoire poursuivie, de par sa fiabilité, ainsi que les différentes échelles de calcul utilisées, font de cette exploration un véritable travail de développement. Pour la rendre complète, la finalisation de cette étude devra toutefois passer par :

- le calcul de spectres au bas nombre d'ondes, sur des systèmes pour lesquels on ne dispose pas de données de référence, dans le but de valider nos conclusions
- une analyse, par exemple, des composantes principales (ACP) des spectres vibrationnels, afin d'étudier la distribution des échantillons (technique déjà employée en archéologie et en pharmacologie), problématique cruciale dès que l'on dispose d'un trop grand nombre de données à analyser et/ou à hiérarchiser comme dans le cas des asphaltènes.

# Appendices

# **Appendix A**

## **Calcul de la polarisabilité**

## A.1 Calcul de la polarisabilité en fonction de la fréquence imaginaire

En l'absence de champ électrostatique, les interactions entre deux systèmes, sont décrites en fonction de leurs moments permanents et des polarisabilités statiques. Pour un champ périodique de pulsation  $\omega$ , les interactions sont modifiées et les polarisabilités dépendent de cette pulsation. L'étude des forces de dispersion à longue distances nécessite la connaissance d'intégrales qui tiennent compte de l'ensemble de fréquences  $\omega$ . Pour des raisons mathématiques simples, les intégrations de fonctions discontinues restent difficiles. On doit se ramener alors à la connaissance de la variation des polarisabilités dépendantes des fréquences imaginaires de chacun des deux systèmes pris isolément puisque ces fonctions sont continues.

Un champ oscillant de fréquence  $\omega$  provoque sur un système donné une distorsion de la distribution des charges. Si  $\omega$  est loin d'une résonance, les charges induites peuvent facilement s'interpréter à partir des moments électriques du système. Si  $\omega$  est proche d'une résonance, les moments électriques induits sont très importants est seule une émission où une absorption d'un photon, associée à un changement d'état quantique du système peut se produire. Nous n'aborderons pas ici le problème dans son détail et renverrons le lecteur au travail effectué par Ayma et al [273].

L'énergie d'induction résulte de la distorsion d'un système (se trouvant dans un état  $m_b$ ) sous l'effet d'un champ  $F_\alpha^b$  et d'un gradient de champ  $F_{\beta\gamma}^b$ , induite par ses plus proches voisins (système  $a$ ). Elle dépend directement des propriétés statiques :

$$E_{ind} = -\frac{1}{2}\alpha_{\alpha\beta}^{mb}(0)F_\alpha^b F_\beta^b - \frac{1}{3!}A\alpha, \beta\gamma(0)F_\alpha^b F_{\beta\gamma}^b - \dots \quad (\text{A.1})$$

Elle est rarement la source prépondérante d'une attraction entre deux système et peut être souvent négligée en comparaison des énergies électrostatiques et des dispersion.

Le calcul de l'énergie de dispersion nécessite quant à lui la connaissance des polarisabilités dynamiques qui représentent la réponse du système à une variation dépendante du temps provoquée par le champ électrique externe.

L'expression des coefficients de dispersion déterminée à partir de l'énergie au second ordre de perturbation.

$$C_{2k} = \sum_{r \neq 0} \sum_{t \neq 0} 4\pi \sum_{m=-l}^l d_m(l, l) \frac{|\langle \psi_r(a) | \sum_{k_a} r_{k_a} Y_l^m | \psi_0(a) \rangle \langle \psi_t(b) | \sum_{k_b} r_{k_b} Y_l^{-m} | \psi_0(b) \rangle|^2}{\Delta E_{\gamma_0 L \rightarrow \gamma_r L'} + \Delta E_{\gamma_0 L \rightarrow \gamma_t L''}} \quad (\text{A.2})$$

s'exprime facilement en fonction des forces d'oscillateurs ( $l+1$ )-polaires :

$$C_{2k} = \sum_{r \neq 0} \sum_{t \neq 0} 4\pi \sum_{m=-l}^l d_m(l, l) \frac{f_{\gamma_0 L \rightarrow \gamma_r L'}^{l_a} f_{\gamma_0 L \rightarrow \gamma_t L''}^{l_b}}{\Delta E_{\gamma_0 L \rightarrow \gamma_r L'} \Delta E_{\gamma_0 L \rightarrow \gamma_t L''} (\Delta E_{\gamma_0 L \rightarrow \gamma_r L'} + \Delta E_{\gamma_0 L \rightarrow \gamma_t L'})} \quad (\text{A.3})$$

Chaque terme de cette somme fait intervenir des grandeurs dépendantes de  $a$  et de  $b$ . On peut transformer ce problème à deux centres, en deux problèmes à un centre, en introduisant la polarisabilité à la fréquences imaginaires :

$$\alpha_l(i\omega) = \frac{8\pi}{(2l+1)} \sum_s \frac{(E_s - E_0) |\langle \psi_0 | \sum_{k_a} r_{k_a}^l Y_l^0 | \psi_s \rangle|^2}{\Delta E^2 + \omega^2} \quad (\text{A.4})$$

et en utilisant l'astuce mathématique proposée par Dalgarno et al[274, 275, 276, 277] où :

$$\frac{1}{X+Y} = \frac{2}{\pi} \int_0^\infty \frac{XY}{(X^2+u^2)(Y^2+u^2)} du \quad \text{avec } X, Y > 0 \quad (\text{A.5})$$

L'expression classique des coefficients de dispersion s'écrit :

$$C_{2k} = \sum_{(l_a, l_b)}^{k-1} \frac{d_m(l_a, l_b)=0}{2\pi} \int_0^\infty \alpha^{l_a}(i\omega) \alpha^{l_b}(i\omega) d\omega$$

dans laquelle :

$$\alpha_{LL'}^{l_a}(i\omega) = \sum_{\gamma L} \frac{f_{\gamma_0 L \rightarrow \gamma L'}^{l_a}}{\Delta E_{\gamma_0 L \gamma_r L'} + \omega^2}$$

et où  $\gamma$  représente l'ensemble des nombres quantiques excepté le nombre quantique orbitalaire L.

Le calcul de  $\alpha(i\omega)$  sous cette forme infinie, semble à priori impossible puisque nous ne disposons que de très peu de données expérimentales et théoriques sur les forces d'oscillateurs pour les états très excités. Le calcul des polarisabilités imaginaires est alors effectué à l'aide de la méthode time dependent gauge invariant (TDGI). La gamme de fréquences généralement utilisée est comprise entre 0 et 3 u.a. puisque, au delà, les valeurs de la polarisabilité deviennent très faibles et négligeables.

A partir de ces valeurs dynamiques, le coefficients de dispersion d'ordre 6 dont l'expression à depuis longtemps établie par Casimir- Polder [278], est la suivante :

$$C_6 = \frac{3}{\pi} \int_0^\infty \alpha_{LL'_a}^1(i\omega) \times \alpha_{LL''_b}^1(i\omega) d\omega \quad (\text{A.6})$$

Dans le cas de deux atomes  $a$  et  $b$  se trouvant dans leur état fondamental, le terme en  $C_6$  est obtenu par un simple ajustement de l'expression A.6.

Le terme en  $C_8$  obtenu pour deux systèmes isotropes est calculé de la même façon.

$$\bar{C}_8 = \frac{15}{\pi} \int_0^\infty [\bar{\alpha}_{LL'_a}^1(i\omega) \bar{C}_{LL''_b}(i\omega) + \bar{C}_{LL'_a}(i\omega) \bar{\alpha}_{LL''_b}^1(i\omega)] d\omega \quad (\text{A.7})$$

où  $C_{LL'}(i\omega) = \frac{1}{3} \alpha_{LL'}^2(i\omega)$  [279] représente la polarisabilité quadrupolaire obtenue en remplaçant l'opérateur  $\vec{r}$  de perturbation par  $\vec{r}^2$ .

# **Appendix B**

## **Calcul de Phonon**

## Introduction

La spectroscopie vibrationnelle fournit de nombreuses informations sur la structure des matériaux. Dans le cas, des structures cristallines où l'interprétation des spectres expérimentaux est très difficile en raison de l'influence du champ cristallin sur les motifs, c'est possible éliminer ces difficultés par l'utilisation en parallèle de deux méthodes spectroscopiques: les spectroscopies Raman et Infrarouge.

En effet, la spectroscopie Raman et la spectroscopie Infrarouge sont des techniques expérimentales complémentaires, flexibles et puissantes qui sont utilisées dans la caractérisation de matériaux. Elles donnent des informations directes sur la surface d'énergie potentielle au voisinage de la position d'équilibre et permet ainsi un examen des caractéristiques structurales des systèmes étudiés. Cependant, dans le pratique, l'analyse et l'attribution des données du spectre sont difficiles, ce qui conduit souvent l'expérimentateur à se focaliser sur les caractéristiques de groupes fonctionnels et des familles de bandes associées. Une perte de certaines informations données par le spectre est donc inéluctable.

Les méthodes ab initio apportent une solution à ce problème dans une certaine mesure, en déterminant précisément le spectre vibrationnelle d'un système. Il est alors possible d'attribuer sans équivoque les données issues du spectre aux modes normaux d'oscillation correspondant, rendant ainsi possible une meilleure compréhension du spectre vibrationnel, ainsi qu'une meilleure caractérisation des propriétés chimiques du système considéré. Des alors qu'il sera possible de réaliser de manière routinière des calculs de type ab initio de haute qualité sur des systèmes de plus en plus complexes, cela permettra potentiellement aux expérimentateurs de réaliser des analyses complémentaires de plus en plus fiables du spectre obtenu.

## B.1 Dynamique du réseau cristallin: Hypothèses fondamentales

Comme point de départ, nous considérons une maille élémentaire d'un cristal parfait, n'étant soumis à aucune vibration. Le réseau de Bravais correspondant est constitué de tous les points décrits par les vecteurs  $\vec{R}$  tel que:

$$\vec{R} = n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3 \quad (\text{B.1})$$

Où,  $\vec{a}_1, \vec{a}_2, \vec{a}_3$  sont les vecteurs primitifs de la maille élémentaire et  $n_i \in Z$ . La position  $\vec{r}_j$  d'un atome j à sa position d'équilibre dans la cellule primitive spécifiée par  $\vec{R}$  est donnée par:

$$\vec{r}_j(\vec{R}) = \vec{R} + \vec{d}_j \quad (\text{B.2})$$

La position de l'atome j soumis à des vibrations est donnée par:

$$\vec{r}_j(\vec{R}) = \vec{R} + \vec{d}_j + \vec{u}_j(\vec{R}) \quad (\text{B.3})$$

Où  $\vec{u}_j(\vec{R})$  représente le déplacement de l'atome j par rapport à la position d'équilibre L'hypothèse de l'amplitude de déplacement  $\vec{u}(\vec{R})$  faible sera admise dans tout ce qui va suivre, ceci permettant de considérer l'approximation harmonique. On précise cependant qu'une telle

approche exclut l'étude de certaines propriétés comme par exemple, la diffusion d'un ion dans un cristal ou le comportement des solides à des températures proches de leur point de fusion. De la même façon, certaines propriétés, telles que la dilatation thermique et la conductibilité thermique, ne peuvent s'expliquer qu'en introduisant des termes anharmoniques. Considerons, un cristal dans une base monoatomique, dans lequel on peut décrire l'énergie potentielle d'interaction entre les ions comme une somme d'interactions de paires. Soit  $\phi(\vec{x})$  le potentiel d'interaction entre 2 ions séparés par le vecteur  $\vec{x}$ . Ce potentiel ne dépend que de la position relative des ions. En tenant compte des vibrations, on a:

$$\vec{x} = \vec{R} - \vec{R}' + \vec{u}(\vec{R}) - \vec{u}(\vec{R}') \quad (\text{B.4})$$

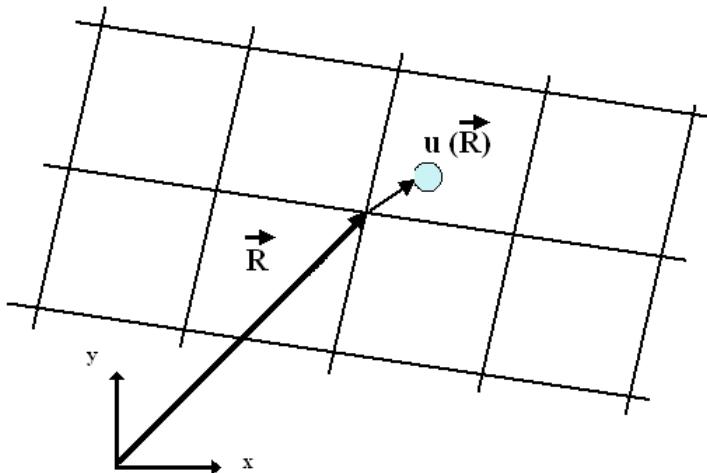


Figure B.1 – Réseau de Bravais et vecteur de déplacement  $u(\vec{R})$  pour une base monoatomique.

L'énergie potentielle totale s'écrit donc:

$$U = \frac{1}{2} \sum_{\substack{\vec{R}, \vec{R}' \\ \vec{R} \neq \vec{R}'}} \phi[\vec{R} - \vec{R}' + \vec{u}(\vec{R}) - \vec{u}(\vec{R}')] \quad (\text{B.5})$$

Dans l'hypothèse où les déplacements  $\vec{u}(\vec{R})$  sont faibles, on peut développer  $\phi(\vec{x})$  autour de  $(\vec{R} - \vec{R}')$ , et on obtient (avec  $\alpha, \beta = x, y, z$ ):

$$U = \frac{1}{2} \sum_{\substack{\vec{R}, \vec{R}' \\ \vec{R} \neq \vec{R}'}} \left\{ \phi(\vec{R} - \vec{R}') + \sum_{\alpha} [u_{\alpha}(\vec{R}) - u_{\alpha}(\vec{R}')] \frac{\partial \phi}{\partial x_{\alpha}} \Big|_{\vec{R} - \vec{R}'} \right. \\ \left. + \frac{1}{2} \sum_{\alpha, \beta} [u_{\alpha}(\vec{R}) - u_{\alpha}(\vec{R}')] [u_{\beta}(\vec{R}) - u_{\beta}(\vec{R}')] \frac{\partial^2 \phi}{\partial x_{\alpha} \partial x_{\beta}} \Big|_{\vec{R} - \vec{R}'} + \dots \right\} \quad (\text{B.6})$$

Le premier terme de l'équation (B.6) correspond au potentiel sans tenir compte des vibrations (réseau statique), il s'écrit:

$$U_{stat} = \frac{1}{2} \sum_{\substack{\vec{R}, \vec{R}' \\ \vec{R} \neq \vec{R}'}} \phi(\vec{R} - \vec{R}') = \frac{N}{2} \sum_{\vec{R} \neq 0} \phi(\vec{R}) \quad (\text{B.7})$$

Le terme linéaire de (B.6) s'annule, le coefficient de  $u_\alpha(\vec{R})$  correspondant à la somme des forces qui s'exercent sur l'ion R à l'équilibre

On a :

$$\frac{1}{2} \sum_{\vec{R}'} \frac{\partial \phi}{\partial x_\alpha} \Big|_{\vec{R}-\vec{R}'} = \frac{\partial U_{stat}}{\partial R_\alpha} = 0 \quad (\text{B.8})$$

L'approximation harmonique revient à négliger dans le développement (B.6) tous les termes d'ordre supérieur à deux. Il vient :

$$U = U_{stat} + U_{harm} \quad (\text{B.9})$$

avec

$$U_{harm} = \frac{1}{4} \sum_{\substack{\vec{R}, \vec{R}' \\ \vec{R} \neq \vec{R}'}} \sum_{\alpha, \beta} [u_\alpha(\vec{R}) - u_\alpha(\vec{R}')] \phi_{\alpha\beta}(\vec{R} - \vec{R}') [u_\beta(\vec{R}) - u_\beta(\vec{R}')] \quad (\text{B.10})$$

où

$$\phi_{\alpha\beta}(x) = \frac{\partial^2 \phi}{\partial x_\alpha \partial x_\beta} \quad (\text{B.11})$$

Le potentiel harmonique peut alors s'écrire :

$$U_{harm} = \frac{1}{2} \sum_{\vec{R}, \vec{R}'} \sum_{\alpha, \beta} u_\alpha(\vec{R}) D_{\alpha\beta}(\vec{R} - \vec{R}') u_\beta(\vec{R}') \quad (\text{B.12})$$

On peut vérifier que (B.10) s'exprime sous la forme générale (B.12) si :

$$D_{\alpha\beta}(\vec{R} - \vec{R}') = \delta_{\vec{R}, \vec{R}'} \sum_{\vec{R}''} \phi_{\alpha\beta}(\vec{R} - \vec{R}'') - \phi_{\alpha\beta}(\vec{R} - \vec{R}') \quad (\text{B.13})$$

Nous indiquerons que l'on peut aussi, du point de vue quantique, décomposer les vibrations d'un solide en une somme de modes propres, chaque mode étant régi par une équation de type oscillateur harmonique.

## B.2 Quantification des ondes élastiques La notion de phonos

Pour déterminer les niveaux d'énergie d'un cristal harmonique (avec une base monoatomique) formé de N ions, il faut déterminer les valeurs propres de l'hamiltonien quantique correspondant à l'hamiltonien classique (afin d'alléger les notations les  $\hat{}$  des opérateurs sont omis):

$$H = \frac{1}{2m} \sum_R p^2(R) + U_{harm} \quad (\text{B.14})$$

Où  $U_{harm}$  est donné par (B.12).

On peut montrer, dans le cas d'une chaîne linéaire, que le Hamiltonien qui est l'équivalent à une dimension de (B.14), peut être exprimé comme une somme de N Hamiltoniens découpés de type oscillateur harmonique. Chaque Hamiltonien est associé à un mode propre de vibration du cristal.

$$H|n_1, \dots, n_v, \dots, n_{N-1}\rangle = E|n_1, \dots, n_v, \dots, n_{N-1}\rangle \quad (\text{B.15})$$

où

$$E = \sum_{v=1}^{N-1} \hbar\omega_v \left( n_v + \frac{1}{2} \right) \quad (\text{B.16})$$

Ce résultat se généralise à 3 dimensions et l'on peut écrire (B.14) sous la forme de 3 N Hamiltoniens correspondant à des oscillateurs harmoniques découpés, les fréquences de ces oscillateurs correspondant aux 3 N modes normaux classiques. La contribution à l'énergie totale d'un mode normal particulier, de pulsation  $\omega_s(k_v)$ , ne peut prendre que l'ensemble discret de valeurs :

$$\left( n_{k_{v,s}} + \frac{1}{2} \right) \hbar\omega_s(k_v) \quad (\text{B.17})$$

où  $n_{k_{v,s}}$ , appelé nombre d'occupation du mode normal v,s, prend les valeurs 0,1,2,... L'énergie totale est la somme des énergies des modes normaux :

$$E = \sum_{k_{v,s}} \left( n_{k_{v,s}} + \frac{1}{2} \right) \hbar\omega_s(k_v) \quad (\text{B.18})$$

Nous avons décrit le résultat B.18 en terme de nombre d'occupation des modes normaux de vecteur d'onde  $k_v$  et d'indices s, où s caractérise la polarisation et la branche (acoustique ou optique) du mode normal considéré. En général, le langage des modes normaux est remplacé par une description de type corpusculaire, équivalente à la terminologie utilisée dans la description quantique du champ électromagnétique (E.M). Dans cette théorie, les énergies des modes normaux de la radiation E.M. dans une cavité sont données par  $(n + 1/2)\hbar\omega$  où  $\omega$  est la fréquence angulaire du mode. Dans ce cas, on ne parle pas du nombre d'occupation n du mode de fréquence  $\omega$ , mais du nombre n de phonons de fréquence  $\omega$ .

De la même manière, au lieu de parler du nombre d'occupation  $n_{k_{v,s}}$  du mode normal de fréquence  $\omega_s(k_v)$ , on dit qu'il y a  $n_{k_{v,s}}$  phonons de types s, de vecteur d'onde  $k_v$ , présents dans le cristal. Cette terminologie est particulièrement utile lorsqu'on examine les processus d'échange d'énergie entre modes normaux ou entre une onde E.M et une vibration du réseau. Cependant, il faut bien réaliser qu'un phonon qui est délocalisé sur l'ensemble du cristal.

## B.3 Théorie classique du cristal harmonique

### B.3.1 Approximation adiabatique

L'état fondamental d'un système est décrit classiquement par l'équation de Schrödinger stationnaire :

$$H(\{r_i\}, \{R_\kappa\})\Psi(\{r_i\}, \{R_\kappa\}) = E\Psi(\{r_i\}, \{R_\kappa\}) \quad (\text{B.19})$$

Où  $\{r_i\}$  est l'ensemble des coordonnées des électrons,  $i=1, \dots, N_e$ , où  $N_e$  est le nombre d'électrons présents dans le système,  $\{R_\kappa\}$  est l'ensemble des coordonnées des noyaux,  $\kappa=1, \dots, N_i$ , avec  $N_i$  le nombre de noyaux présents dans le système. Pour simplifier, nous désignerons dorénavant  $\{r_i\}$  et  $\{R_\kappa\}$  respectivement par  $r$  et  $R$  (même chose pour la notation des vecteurs). Le hamiltonien global s'écrit :

$$H(r, R) = T_i(R) + U_{ii}(R) + T_e(r) + U_{ee}(r) + U_{ie}(r, R) \quad (\text{B.20})$$

Où  $T_i(R)$  et  $U_{ii}(R)$  représentent les opérateurs d'énergie cinétique et d'énergie potentielle des ions,  $T_e(r)$  et  $U_{ii}(R)$  les opérateurs d'énergie cinétique et d'énergie potentielle des électrons et  $U_{ie}(r, R)$  l'opérateur d'interaction électron-ion

Ils sont définis respectivement par :

$$T_i(R) = - \sum_{\kappa} \frac{\hbar^2}{2M_{\kappa}} \nabla_{R_{\kappa}}^2 \quad (\text{B.21})$$

$$U_{ii}(R) = + \sum_{\kappa < \kappa'} \frac{Z_{\kappa} Z_{\kappa'} e^2}{|R_{\kappa} - R_{\kappa'}|} \quad (\text{B.22})$$

$$T_e(r) = - \sum_i \frac{\hbar^2}{2m_e} \nabla_{r_i}^2 \quad (\text{B.23})$$

$$U_{ee}(r) = + \sum_{i < j} \frac{e^2}{|r_i - r_j|} \quad (\text{B.24})$$

$$U_{ie}(r, R) = - \sum_{i, \kappa} \frac{Z_{\kappa} e^2}{|r_i - R_{\kappa}|} \quad (\text{B.25})$$

L'approximation adiabatique consiste à considérer le terme cinétique des noyaux,  $T_i(R)$ , comme une perturbation:

$$H(r, R) = H_{BO}(r, R) + T_i(R) \quad (\text{B.26})$$

Ainsi, le Hamiltonien de Born-Oppenheimer,  $H_{BO}$ , ne présente plus d'opérateur différentiel par rapport aux positions nucléaires. Par conséquent, ces dernières peuvent être considérées comme des variables classiques et non plus quantiques. Les positions nucléaires se comportent alors comme des paramètres. D'un point de vue physique, on considère que les électrons s'adaptent instantanément à la position des noyaux. Dans l'approximation de Born- Oppenheimer, l'état fondamental du système est alors solution du problème aux valeurs propres:

$$H_{BO}(r, R)\phi(r, R) = E_{BO}(R)\phi(r, R) \quad (\text{B.27})$$

### B.3.2 Cas du cristal harmonique

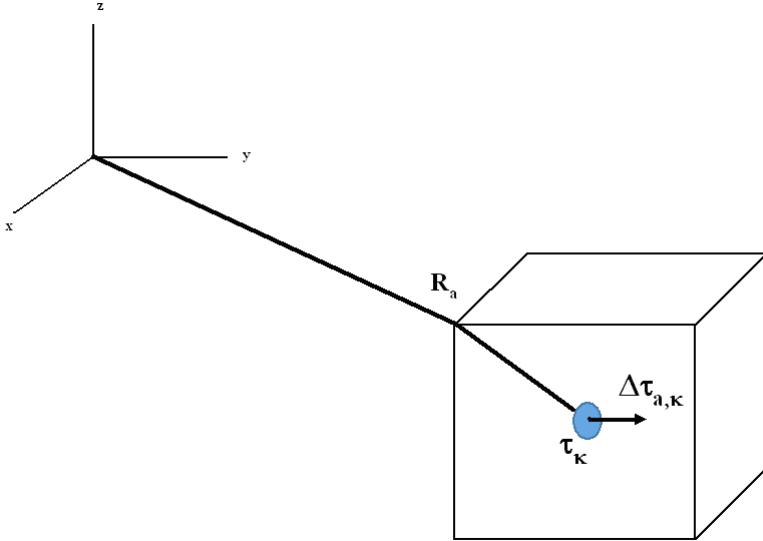


Figure B.2 – Chaque maille unitaire est repérée par un vecteur  $R_a$ , où  $a=1,2, \dots, N$ . Les atomes de la maille sont repérés par  $\tau_\kappa$  et leur déplacement par  $\Delta\tau_\kappa$  ( $\kappa$  variant de 1 à  $r$ )

Intéressons-nous maintenant au cas particulier d'un cristal, c'est-à-dire d'un solide périodique composé de  $N$  mailles élémentaires contenant  $r$  atomes par cellule unité. Alors, l'indice  $\kappa$  utilisé jusqu'à présent est remplacé par le couple  $(a,\kappa)$ . Ainsi, la position d'équilibre de l'atome  $\kappa$  dans la maille  $a$  est donnée par :

$$R_{a\kappa}^0 = R_a + \tau_\kappa \quad (\text{B.28})$$

où  $R_a = a_1 x_1 + a_2 x_2 + a_3 x_3$  est un vecteur de translation du réseau cristallin (les vecteurs  $x_i$  étant les vecteurs primitifs du réseau direct) et  $\kappa = 1, 2, \dots, r$  désignant les atomes de la maille unitaire. Chaque atome peut se déplacer de sa position d'équilibre d'une quantité  $\Delta\tau_{a,\kappa}$  (dépendant à la fois de  $\kappa$  et de  $a$ ) de sorte que sa position instantanée est donnée par :

$$R_{a,\kappa} = R_a + \tau_\kappa + \Delta\tau_{a,\kappa} = R_{a,\kappa}^0 + \Delta\tau_{a,\kappa} \quad (\text{B.29})$$

Pour calculer les phonos, c'est-à-dire les fréquences propres de vibration, nous ferons l'hypothèse d'un cristal périodique infini. En d'autres termes, nous négligerons les effets de surface. Un problème survient cependant : une infinité de valeurs seront nécessaires à la description des propriétés du cristal. Nous serons ainsi amené à imposer des conditions limites périodiques, nous permettant de traiter un cristal de volume fini répété de manière périodique.

Le problème n'est plus alors de résoudre une infinité d'équations du mouvement mais un ensemble de  $3r$  équations homogènes dont les inconnues sont les vecteurs propres de polarisation  $\eta_{\kappa,\alpha}$  (définis plus loin) et ce, pour un ensemble infini de vecteurs  $q$  de la première zone de Brillouin. En pratique, on doit se restreindre à un ensemble fini de vecteurs  $q$ . Cela correspond à imposer des conditions limites périodiques de Born von Karman, non plus sur un cristal infini mais sur un cristal fini :

$$\tau_\kappa(R_a + N_i x_i) = \tau_\kappa(R_a) \quad (\text{B.30})$$

pour chacun des trois vecteurs primitifs  $x_i$  et où les  $N_i$  sont des nombres entiers satisfaisant la relation  $N = N_1 N_2 N_3$ , limitant de ce fait le nombre de vecteurs  $\mathbf{q}$  permis dans la zone de Brillouin :

$$\mathbf{q} = \frac{n_1}{N_1} \mathbf{b}_1 + \frac{n_2}{N_2} \mathbf{b}_2 + \frac{n_3}{N_3} \mathbf{b}_3, \quad n_i \text{ entiers} \quad (\text{B.31})$$

En pratique on se limite à  $N_i = 2$  ou  $3$ . Nous avons vu que la dynamique du réseau était déterminée par l'énergie de Born- Oppenheimer  $E_{BO}(\{R_{a,\kappa}\})$ . Dorénavant, nous désignerons cette énergie par  $E_{BO}(\{\Delta\tau_{a,\kappa}\})$  étant donné que les positions d'équilibre des atomes donnent une contribution constante à l'énergie et que nous ne nous intéressons qu'aux propriétés dynamiques qui ne font intervenir que les déplacements. Nous supposerons les déplacements atomiques suffisamment faibles de manière à pouvoir utiliser l'approximation harmonique[1]. En absence de champ électrique macroscopique, l'énergie de BO s'écrit:

$$E_{BO}^{harm}(\Delta\tau) = \sum_{\alpha\kappa} \left( \frac{\partial E_{BO}}{\partial \tau_{\kappa\alpha}} \right)_0 \Delta\tau_{\kappa\alpha} + \frac{1}{2} \sum_{a,\kappa,\alpha} \sum_{b,\kappa',\beta} \left( \frac{\partial^2 E_{BO}}{\partial \tau_{a,\kappa,\alpha} \partial \tau_{b,\kappa',\beta}} \right)_0 \Delta\tau_{a,\kappa,\alpha} \Delta\tau_{b,\kappa',\beta} \quad (\text{B.32})$$

où  $\alpha$  et  $\beta$  sont des indices se rapportant aux trois directions de l'espace. Le terme du premier ordre est nul car nous nous sommes placés aux positions d'équilibre. La dérivée seconde est évaluée aux positions d'équilibre (indice 0). En reprenant les notations introduites précédemment, on définit alors:

$$C_{\kappa,\alpha\kappa'\beta}(a, a') = \frac{\partial^2 E_{BO}}{\partial \tau_{a,\kappa,\alpha} \partial \tau_{b,\kappa',\beta}} \quad (\text{B.33})$$

Ces coefficients sont appelés les constantes de force interatomiques.

### B.3.3 Calcul du spectre de phonons au point $\Gamma$

Dans tout ce qui va suivre, et quelque soit le système considéré, dans le cadre de cette thèse, nous avons calculé le spectre de phonons uniquement au centre de la zone de Brillouin (point  $\Gamma$ ). Il est par conséquent utile de préciser les propriétés de ce point.

Les caractéristiques du problème au point  $\Gamma$  ( $\vec{q} = 0$ ) sont intéressantes et spécifiques par certains aspects :

- La matrice  $W(\vec{q})$  qui est la matrice regroupant l'ensemble des valeurs propres  $\omega_m(\vec{q})$  est plus simple à calculer au point  $\Gamma$  qu'aux autres points du réseau réciproque, toutes les opérations de symétrie étant présentes en ce point. Ce qui permet:
  - de réduire le nombre de calculs explicites de dérivées à un minimum;
  - de factoriser la matrice  $W(\Gamma) = W(0)$  avant la diagonalisation;
  - d'éliminer les trois modes acoustiques.
- En effet, une propriété unique au point  $(\Gamma)$ , qui n'a pas d'équivalence dans la théorie des bandes, est que trois des modes ont une fréquence nulle, car ils correspondent aux translations pures du cristal entier. En fait, il est possible de les négliger (il est à noter

que le fait d'imposer des conditions aux limites périodiques exclut la possibilité de décrire les modes rotationnels de l'ensemble du cristal).

- $W(0)$  possédant toute la symétrie ponctuelle du cristal, elle peut alors être factorisée selon les représentations irréductibles du groupe ponctuel correspondant.
- Les modes du point ( $\Gamma$ ) sont les seuls qui peuvent interagir avec les ondes électromagnétiques, permettant d'atteindre les spectres IR ou Raman.
- Un traitement spécifique est requis pour le calcul des fréquences en ce point pour les semi-conducteurs polaires et les isolants.
- A l'aide de l'approche en super maille, la résolution du problème au point  $\Gamma$  permet l'étude des vibrations aux autres points  $\vec{q}$ . Cependant, cet aspect n'ayant pas été abordé lors de ma thèse, il ne sera pas traité ici.

# **Appendix C**

## **Database**

## C.1 Monomers

Table C.1 – Low wavenumber Raman and PA infrared spectra of Indole (Cs).

Observed			Predicted <sup>a</sup>	
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
	207 <sup>e</sup>	$\nu_1$ (A'')	214 (13.32) [0.14]	209 (10.02)
232 vw	241 <sup>e</sup>	$\nu_2$ (A'')	246 (0.04) [0.10]	239 (0.06)
262 vw				
	387 <sup>e</sup>	$\nu_3$ (A'')	399 (79.31) [0.47]	395 (54.39)
398 vw		$\nu_4$ (A')	406 (4.59) [0.35]	398 (4.05)
429 vw	419 <sup>e</sup>	$\nu_5$ (A'')	435 (2.84) [0.24]	425 (1.32)
545 w		$\nu_6$ (A')	555 (0.08) [6.55]	550 (0.04)
576 vw	579 m	$\nu_7$ (A'')	590 (0.83) [0.04]	575 (0.34)
610 w	610 m	$\nu_8$ (A'')	622 (7.08) [0.44]	609 (4.06)
		$\nu_9$ (A')	624 (1.37) [8.33]	616 (1.39)

<sup>a</sup> (km/mole) [Å<sup>4</sup>/AMU]

<sup>b</sup> ref [280]. Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder. <sup>c</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*. <sup>d</sup> Anharmonics ωB97X-D/6-311++G\*\* calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.2 – Raman and PA infrared spectra of Indole (Cs), 700–2000 cm<sup>1</sup>

Observed			Predicted <sup>a</sup>	
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
732 vw	732 m	$\nu_{10}$ (A'')	737 (28.08) [0.71]	734 (15.70)
761 m	759 m	$\nu_{11}$ (A'')	763 (107.60) [0.24]	758 (56.38)
	772 sh	$\nu_{12}$ (A')	782 (3.00) [25.50]	766 (3.18)
		$\nu_{13}$ (A'')	790 (19.58) [0.10]	775 (12.75)
	862 w	$\nu_{14}$ (A'')	869 (0.87) [0.06]	860 (0.33)
878 vw	875 sh	$\nu_{15}$ (A'')	893 (1.02) [1.79]	873 (0.63)
		$\nu_{16}$ (A')	894 (0.52) [1.91]	882 (0.32)
897 m	898 m	$\nu_{17}$ (A')	919 (6.75) [6.58]	905 (4.79)
937 vw	932 m	$\nu_{18}$ (A'')	961 (2.13) [0.17]	940 (2.20)
1005 s	1007 s	$\nu_{19}$ (A'')	997 (0.03) [0.10]	1003 (0.30)
1011 s		$\nu_{20}$ (A')	1044 (7.38) [26.11]	1021 (4.80)
1060 s	1060 m	$\nu_{21}$ (A')	1096 (5.59) [19.66]	1070 (4.71)
1094 vw	1091 m	$\nu_{22}$ (A')	1122 (31.78) [2.07]	1096 (27.43)
1120 w	1119 m	$\nu_{23}$ (A')	1153 (0.33) [8.11]	1128 (1.42)
1153 w	1150 w	$\nu_{24}$ (A')	1181 (1.71) [2.37]	1159 (1.60)

Observed			Predicted <sup>a</sup>	
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
1249 w	1247 vs	$\nu_{26}$ (A')	1276 (10.37) [10.38]	1255 (9.22)
1278 m	1278 s	$\nu_{27}$ (A')	1315 (11.49) [21.44]	1286 (10.54)
1338 s	1341 sh	$\nu_{28}$ (A')	1379 (32.50) [35.20]	1339 (13.46)
1355 m	1355 vs	$\nu_{29}$ (A')	1391 (5.98) [18.36]	1363 (12.70)
1416 m	1415 vs	$\nu_{30}$ (A')	1466 (21.86) [35.59]	1422 (16.80)
1456 m	1456 vs	$\nu_{31}$ (A')	1502 (29.55) [18.91]	1462 (18.53)
1488 m	1488 s	$\nu_{32}$ (A')	1543 (5.27) [10.38]	1500 (5.21)
1507 vs	1506 s	$\nu_{33}$ (A')	1577 (12.39) [103.43]	1520(12.05)
1578 m	1577 m	$\nu_{34}$ (A')	1652 (0.66) [12.79]	1590 (0.73)
1617 m	1615 s	$\nu_{35}$ (A')	1695 (3.18) [22.34]	1627 (2.80)

<sup>a</sup> (km/mole) [ $\text{A}^4/\text{AMU}$ ]    <sup>b</sup> Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder.    <sup>c</sup> Unscaled DFT frequencies calculated using  $\omega\text{B97X-D}/6-311++\text{G}^{**}$ .    <sup>d</sup> Anharmonics  $\omega\text{B97X-D}/6-311++\text{G}^{**}$  calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.3 – Low wavenumber Raman ad PA infrared spectra of Carbazole ( $\text{C}_2\text{v}$ ).

Observed			Predicted <sup>a</sup>	
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
71				
125		$\nu_1$ (b <sub>1</sub> )	103 (5.54) [0.18]	110 (6.20)
146		$\nu_2$ (a <sub>2</sub> )	150 (0.00) [0.02]	148 (0.00)
221 m	221	$\nu_3$ (a <sub>1</sub> )	222 (0.38) [1.52]	217 (0.52)
299 m		$\nu_5$ (a <sub>2</sub> )	297 (0.00) [2.79]	291 (0.00)
	294	$\nu_4$ (b <sub>1</sub> )	284 (15.47) [0.81]	299 (13.20)
305 sh	312	$\nu_6$ (b <sub>1</sub> )	343 (61.71) [0.00]	310 (38.15)
431 s	423	$\nu_7$ (b <sub>1</sub> )	434 (11.32) [0.01]	430 (9.21)
		$\nu_9$ (a <sub>2</sub> )	453 (0.00) [0.05]	430 (0.00)
	445	$\nu_8$ (a <sub>1</sub> )	439 (1.44) [9.21]	434 (1.08)
	507	$\nu_{10}$ (b <sub>2</sub> )	520 (4.76) [1.22]	508 (3.23)
552 m		$\nu_{11}$ (b <sub>2</sub> )	569 (0.22) [10.91]	555 (0.23)
568 vw	573 vs	$\nu_{12}$ (b <sub>1</sub> )	582 (14.77) [0.02]	575 (15.68)
	602 w	$\nu_{13}$ (a <sub>2</sub> )	592 (0.00) [0.02]	590 (0.00)
	619 m	$\nu_{14}$ (b <sub>2</sub> )	636 (6.07) [0.00]	629 (5.84)
656 w	656 m	$\nu_{15}$ (a <sub>1</sub> )	675 (0.49) [3.03]	661 (0.58)

<sup>a</sup> (km/mole) [ $\text{A}^4/\text{AMU}$ ]    <sup>b</sup> Estimated uncertainties in band positions are within  $\pm 3 \text{ cm}^{-1}$ , Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder.    <sup>c</sup> Unscaled DFT frequencies calculated using  $\omega\text{B97X-D}/6-311++\text{G}^{**}$ .    <sup>d</sup> Anharmonics  $\omega\text{B97X-D}/6-311++\text{G}^{**}$  calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.4 – Raman and PA infrared spectra of Carbazole ( $C_2V$ ), 700–2000  $\text{cm}^{-1}$ 

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
	718 s	$\nu_{16}$ (b <sub>1</sub> )	749 (101,36) [0,02]	724 (54,66)
731 sh		$\nu_{17}$ (a <sub>2</sub> )	762 (0,00) [1,09]	739 (0.00)
743 s	742 m	$\nu_{18}$ (ab <sub>1</sub> )	766 (0.004) [32.29]	751 (0.02)
		$\nu_{19}$ (b <sub>1</sub> )	776 (66.02) [0.002]	752 (38.64)
	758 w			
774 vw	775 w	$\nu_{20}$ (a <sub>2</sub> )	795 (0.00) [0.75]	782 (0.00)
	809 w			
	843 sh			
	855 s	$\nu_{22}$ (b <sub>2</sub> )	873 (3.02) [0.96]	858 (3.09)
		$\nu_{21}$ (b <sub>1</sub> )	869 (0.46) [0.20]	863 (0.51)
863 vw	879 sh	$\nu_{23}$ (a <sub>2</sub> )	874 (0.00) [0.21]	870 (0.00)
886 vw		$\nu_{24}$ (a <sub>1</sub> )	898 (0.36) [23.08]	893 (0.31)
	917 s	$\nu_{26}$ (a <sub>2</sub> )	963 (0.00) [0.16]	929 (0.00)
		$\nu_{25}$ (b <sub>1</sub> )	961 (2.32) [0.07]	930 (3.57)
	969 vw	$\nu_{27}$ (a <sub>2</sub> )	997 (0.00) [0.05]	977 (0.00)
		$\nu_{28}$ (b <sub>1</sub> )	998 (0.00) [0.01]	979 (0.00)
	997 m	$\nu_{29}$ (b <sub>2</sub> )	1028 (7.00) [0.61]	993 (6.90)
	1009 m	$\nu_{30}$ (a <sub>1</sub> )	1053 (7.60) [0.00]	1007 (6.73)
1012 vs		$\nu_{31}$ (b <sub>2</sub> )	1049 (0.85) [64.17]	1020 (0.22)
	1072 w			
1106 w	1108 m	$\nu_{32}$ (a <sub>1</sub> )	1142 (5.84) [18.04]	1110 (6.01)
1121 vw				
1149 vw	1141 s	$\nu_{33}$ (b <sub>2</sub> )	1153 (7.45) [1.08]	1138 (7.20)
1162 w	1159 w	$\nu_{34}$ (a <sub>1</sub> )	1179 (7.55) [6.35]	1164 (5.03)
	1176 vw	$\nu_{35}$ (b <sub>2</sub> )	1188 (1.56) [3.00]	1171 (1.06)
1207 w	1207 s	$\nu_{36}$ (b <sub>2</sub> )	1243 (0.16) [3.32]	1204 (0.07)
		$\nu_{37}$ (a <sub>1</sub> )	1244 (5.92) [46.13]	1210 (5.54)
1239 vw	1235 s	$\nu_{38}$ (b <sub>2</sub> )	1276 (82.07) [11.07]	1240 (45.68)
	1255 vw			
1291 s	1288 m	$\nu_{39}$ (a <sub>1</sub> )	1328 (2.15) [122.15]	1295 (2.26)
1313 s		$\nu_{40}$ (a <sub>1</sub> )	1347 (1.35) [168.50]	1320 (1.20)
1326 vw				
1337 m	1330 vs	$\nu_{41}$ (b <sub>2</sub> )	1365 (77.48) [2.73]	1334 (37.61)
	1362 sh	$\nu_{42}$ (a <sub>1</sub> )	1372 (13.00) [16.37]	1360 (11.23)
	1394 m	$\nu_{43}$ (b <sub>2</sub> )	1440 (13.73) [3.13]	1405 (10.82)
	1427 sh			
	1446 s	$\nu_{44}$ (a <sub>1</sub> )	1500 (40.25) [44.07]	1460 (38.36)
1453 w				
1484 w		$\nu_{45}$ (b <sub>2</sub> )	1508 (33.32) [0.44]	1477 (21.15)
	1494 m	$\nu_{46}$ (a <sub>1</sub> )	1538 (1.34) [67.63]	1505 (1.28)
		$\nu_{47}$ (b <sub>2</sub> )	1551 (58.07) [2.20]	1510 (34.90)
1574 m				

<sup>a</sup> (km/mole) [A<sup>4</sup>/AMU]    <sup>b</sup> Estimated uncertainties in band positions are within  $\pm 3$   $\text{cm}^{-1}$ , Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder.    <sup>c</sup> Unscaled DFT frequencies calculated using  $\omega\text{B97X-D/6-311++G**}$ .    <sup>d</sup> Anharmonics  $\omega\text{B97X-D/6-311++G**}$  calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.5 – Low wavenumber Raman ad PA infrared spectra of 4,5-iminophenanthrene ( $C_2V$ ).

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
124 m		$\nu_1$ (b <sub>1</sub> )	92 (7.03) [0.04]	100 (5.02)
139 vw		$\nu_2$ (a <sub>2</sub> )	195 (0.00) [0.006]	193 (0.00)
210 vw		$\nu_3$ (b <sub>1</sub> )	211 (3.98) [0.01]	208 (2.13)
227 vw		$\nu_4$ (b <sub>1</sub> )	287 (30.72) [1.34]	276 (20.14)
		$\nu_5$ (a <sub>2</sub> )	307 (0.00) [0.55]	
315 w		$\nu_6$ (b <sub>1</sub> )	348 (60.77) [0.04]	313 (32.55)
346 vw		$\nu_7$ (a <sub>1</sub> )	351 (3.25) [0.34]	343 (2.98)
435 w		$\nu_8$ (a <sub>2</sub> )	444 (0.00) [1.60]	439 (0.00)
449 s		$\nu_9$ (a <sub>1</sub> )	457 (0.04) [25.02]	453 (0.03)
481 w		$\nu_{10}$ (b <sub>2</sub> )	491 (2.86) [3.06]	485 (2.06)
		$\nu_{11}$ (b <sub>1</sub> )	508 (4.09) [0.04]	498 (3.64)
537 w		$\nu_{12}$ (b <sub>2</sub> )	549 (0.22) [2.00]	540 (0.20)
562 w		$\nu_{13}$ (b <sub>2</sub> )	580 (0.01) [2.73]	
		$\nu_{14}$ (a <sub>2</sub> )	585 (0.00) [0.04]	581 (0.00)
		$\nu_{15}$ (b <sub>1</sub> )	613 (0.002) [0.01]	595 (0.001)
605 s	604 m	$\nu_{16}$ (a <sub>1</sub> )	620 (3.78) [33.39]	610 (3.58)
613 sh				
	623 vw			
	668 w	$\nu_{17}$ (a <sub>2</sub> )	664 (0.00) [0.17]	
678 m	677 w	$\nu_{18}$ (a <sub>1</sub> )	696 (0.86) [19.24]	684 (0.82)
	693 w	$\nu_{19}$ (b <sub>2</sub> )	710 (0.10) [0.05]	703 (0.06)

<sup>a</sup> (km/mole) [A<sup>4</sup>/AMU]    <sup>b</sup> Estimated uncertainties in band positions are within  $\pm 3 \text{ cm}^{-1}$ , Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder.    <sup>c</sup> Unscaled DFT frequencies calculated using  $\omega\text{B97X-D}/6-311++\text{G}^{**}$ .    <sup>d</sup> Anharmonics  $\omega\text{B97X-D}/6-311++\text{G}^{**}$  calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.6 – Raman and PA infrared spectra of 4,5-iminophenanthrene ( $C_2V$ ), 700–2000  $\text{cm}^{-1}$ 

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
717 vw	715 vs	$\nu_{20}$ (b <sub>1</sub> )	740 (61.10) [0.01]	719 (40.72)
755 vw	752 vs	$\nu_{21}$ (b <sub>1</sub> )	774 (27.27) [0.001]	750 (44.35)
766 vw		$\nu_{22}$ (a <sub>2</sub> )	776 (0.00) [1.41]	769 (0.00)
	797 sh	$\nu_{23}$ (b <sub>2</sub> )	821 (0.18) [1.01]	804 (0.17)
818 w	820 vs	$\nu_{26}$ (b <sub>1</sub> )	845 (89.63) [0.39]	822 (82.81)

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
	847 sh	$\nu_{25}$ (a <sub>2</sub> )	841 (0.71) [10.92]	838 (0.70)
		$\nu_{24}$ (a <sub>2</sub> )	828 (0.00) [1.31]	860 (0.00)
871 vw	872 s	$\nu_{28}$ (b <sub>1</sub> )	904 (2.33) [0.29]	880 (1.78)
		$\nu_{27}$ (a <sub>2</sub> )	896 (0.00) [0.25]	880 (0.00)
962 vw		$\nu_{30}$ (b <sub>1</sub> )	987 (1.49) [0.00]	970 (1.60)
969 vw		$\nu_{29}$ (a <sub>2</sub> )	986 (0.00) [0.05]	972 (0.00)
1000 vw	999 w	$\nu_{31}$ (a <sub>2</sub> )	1002 (0.00) [2.00]	998 (0.00)
1016 m	1018 vw	$\nu_{32}$ (b <sub>2</sub> )	1030 (1.20) [1.17]	1017 (1.27)
1030 m	1029 w	$\nu_{33}$ (a <sub>1</sub> )	1049 (0.46) [1.33]	1030 (0.10)
		$\nu_{34}$ (a <sub>1</sub> )	1058 (1.30) [50.56]	1034 (1.47)
1071 w	1069 m	$\nu_{35}$ (b <sub>2</sub> )	1076 (0.56) [4.20]	1070 (0.45)
	1084 sh	$\nu_{36}$ (a <sub>2</sub> )	1103 (49.30) [8.89]	1088 (45.49)
1135 w	1132 m	$\nu_{37}$ (b <sub>1</sub> )	1165 (1.36) [8.47]	1140 (1.14)
1174 sh	1167 sh	$\nu_{38}$ (b <sub>2</sub> )	1203 (4.35) [0.30]	1180 (4.17)
1179 w	1177 s	$\nu_{39}$ (a <sub>1</sub> )	1214 (5.61) [12.35]	1188 (5.38)
1214 s	1212 m	$\nu_{41}$ (a <sub>1</sub> )	246 (3.66) [76.83]	1222 (3.51)
1228 w	1227 w	$\nu_{40}$ (b <sub>2</sub> )	1245 (0.54) [3.89]	1226 (0.52)
1246 m	1244 w	$\nu_{42}$ (b <sub>2</sub> )	1262 (12.34) [4.69]	1241(11.18)
		$\nu_{43}$ (a <sub>1</sub> )	1279 (3.04) [34.76]	1250 (2.90)
1328 m	1325 vs	$\nu_{45}$ (a <sub>1</sub> )	1364 (22.45) [88.61]	1330 (21.49)
		$\nu_{44}$ (b <sub>2</sub> )	1361 (52.78) [0.34]	1332 (50.61)
1355 w				
1381 vw	1379 m	$\nu_{46}$ (a <sub>1</sub> )	1394 (0.21) [16.30]	1371 (0.20)
	1395 sh	$\nu_{47}$ (b <sub>2</sub> )	1434 (4.48) [0.00]	1402 (4.08)
1405 s		$\nu_{48}$ (b <sub>2</sub> )	1447 (3.06) [53.67]	1416 (3.00)
1440 m	1437 vs	$\nu_{49}$ (b <sub>2</sub> )	1485 (5.36) [0.80]	1447 (5.10)
1468 vs	1466 w	$\nu_{50}$ (a <sub>1</sub> )	1490 (26.18) [43.74]	1461 (25.02)
1480 sh				
1502 vw	1500 s	$\nu_{51}$ (a <sub>1</sub> )	1532 (1.14) [393.19]	1508 (1.10)
		$\nu_{52}$ (b <sub>2</sub> )	1555 (2.09) [0.01]	1515 (1.99)
		$\nu_5$ (a <sub>1</sub> )	1565 (25.86) [9.36]	1528 (24.72)
1595 sh	1596 m	$\nu_{55}$ (b <sub>2</sub> )	1672 (120.44) [0.07]	1625 (115.12)
1634 m	1633 m	$\nu_{54}$ (a <sub>1</sub> )	1672 (0.18) [29.07]	1630 (0.17)
1664 vw		$\nu_{56}$ (b <sub>2</sub> )	1707 (3.81) [128.95]	1674 (3.13)
		$\nu_{57}$ (a <sub>1</sub> )	1738 (0.66) [6.79]	1705 (1.46)

<sup>a</sup> (km/mole) [Å<sup>4</sup>/AMU]    <sup>b</sup> Estimated uncertainties in band positions are within ±3 cm<sup>-1</sup>, Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder.    <sup>c</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*.    <sup>d</sup> Anharmonics ωB97X-D/6-311++G\*\* calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.7 – Low wavenumber Raman and PA infrared spectra of 1-methylcarbazole (Cs).

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
191 m	239 w	$\nu_1$ (A'')	93 (3.87) [0.22]	89 (2.90)
		$\nu_2$ (A'')	121 (0.56) [0.46]	118 (0.15)
		$\nu_3$ (A'')	134 (2.24) [0.34]	140 (2.02)
		$\nu_4$ (A')	185 (0.24) [0.69]	184 (0.63)
		$\nu_5$ (A'')	220 (0.02) [0.62]	216 (1.21)
	438 s	$\nu_6$ (A'')	270 (27.07) [0.57]	267 (20.32)
		$\nu_7$ (A'')	296 (0.02) [2.66]	293 (0.05)
		$\nu_8$ (A')	301 (0.08) [4.58]	302 (0.15)
		$\nu_9$ (A'')	347 (48.61) [0.18]	342 (27.08)
		$\nu_{10}$ (A'')	440 (4.85) [0.003]	430 (5.20)
439 m	507 m	$\nu_{11}$ (A')	447 (1.66) [4.24]	437 (2.27)
512 s		$\nu_{12}$ (A')	514 (0.36) [9.49]	505 (0.38)
555 w		$\nu_{13}$ (A'')	522 (3.46) [0.11]	516 (2.96)
573 sh	553 m	$\nu_{14}$ (A')	567 (6.92) [2.72]	560 (6.10)
584 m	574 vs	$\nu_{15}$ (A'')	585 (9.23) [0.003]	574 (14.04)
639 m	635 vw	$\nu_{16}$ (A'')	595 (0.00) [0.06]	588 (0.01)
		$\nu_{17}$ (A')	597 (3.33) [9.86]	590 (3.73)
		$\nu_{18}$ (A')	656 (1.62) [4.82]	644 (0.56)

<sup>a</sup> (km/mole) [A<sup>4</sup>/AMU]    <sup>b</sup> Estimated uncertainties in band positions are within $\pm 3 \text{ cm}^{-1}$ , Relative intensities: vw = very weak, w = weak, m = medium, s =strong, vs = very strong, sh = shoulder.    <sup>c</sup> Unscaled DFT frequencies calcu-lated using  $\omega\text{B97X-D}/6-311++\text{G}^{**}$ .    <sup>d</sup> Anharmonics  $\omega\text{B97X-D}/6-311++\text{G}^{**}$ calculation using the P\_VMWCI<sub>2</sub> algorithmTable C.8 – Raman and PA infrared spectra of 1-methylcarbazole (Cs), 700–2000 cm<sup>-1</sup>

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
705 s	704 w	$\nu_{19}$ (A')	724 (0.48) [18.53]	713 (0.46)
737vw	734 vs	$\nu_{20}$ (A'')	756 (37.58) [0.67]	739 (41.72)
751 vw	751 vs	$\nu_{21}$ (A'')	771 (40.47) [0.16]	753 (38.95)
756 vw		$\nu_{22}$ (A'')	774 (55.84) [0.05]	756 (32.48)
788 w	787 s	$\nu_{23}$ (A'')	811 (5.95) [1.08]	798 (8.11)
835 w	834 vw	$\nu_{24}$ (A')	862 (1.07) [7.31]	844 (0.68)
846 vw	851 m	$\nu_{25}$ (A'')	871 (0.23) [0.16]	859 (0.29)
	882 m	$\nu_{26}$ (A')	880 (2.21) [10.68]	875 (2.28)
		$\nu_{27}$ (A'')	923 (0.00) [0.15]	903 (0.00)
914 m	918 m	$\nu_{28}$ (A')	926 (2.48) [13.84]	914 (2.38)

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
936 vw	930 m	$\nu_{29}$ (A'')	961 (1.06) [0.21]	942 (1.25)
	950 w	$\nu_{30}$ (A'')	986 (0.40) [0.07]	961 (0.42)
	970 vw	$\nu_{31}$ (A'')	996 (0.003) [0.01]	972 (0.01)
990 m	990 w	$\nu_{32}$ (A')	1011 (2.23) [22.06]	985 (2.19)
1014 s	1012 w	$\nu_{33}$ (A')	1050 (6.99) [30.51]	1023 (4.56)
1038 vw	1037 m	$\nu_{34}$ (A'')	1065 (1.20) [0.01]	1041 (1.63)
1060 vw	1059 s	$\nu_{35}$ (A')	1095 (12.91) [1.12]	1068 (11.90)
1091 w	1092 m	$\nu_{36}$ (A')	1118 (2.11) [21.30]	1094 (2.54)
1122 vw	1120 m	$\nu_{37}$ (A')	1153 (12.36) [0.78]	1131 (8.93)
1160 vw	1157 m	$\nu_{38}$ (A')	1184 (3.39) [4.75]	1166 (2.60)
	1179 w	$\nu_{39}$ (A')	1195 (0.12) [0.48]	1174 (0.27)
1214 w	1213 m	$\nu_{40}$ (A')	1241 (1.42) [21.26]	1225 (1.32)
	1235 s	$\nu_{41}$ (A')	1271 (56.32) [6.44]	1245 (47.88)
1249 vw	1247 sh	$\nu_{42}$ (A')	1282 (5.05) [5.66]	1254 (3.96)
1282 m	1281 m	$\nu_{43}$ (A')	1321 (6.28) [37.72]	1288 (6.06)
1313 vs	1313 sh	$\nu_{44}$ (A')	1349 (22.82) [266.80]	1320 (9.80)
1326 m	1325 s	$\nu_{45}$ (A')	1360 (67.03) [29.73]	1335 (55.86)
1343 w	1340 s	$\nu_{46}$ (A')	1372 (16.15) [8.50]	1346 (12.26)
1365 vw	1364 w	$\nu_{47}$ (A')	1421 (4.76) [10.16]	1370 (2.29)
	1375 w			
1387 w				
1393 sh	1392 w			
1418 w	1418 s	$\nu_{48}$ (A')	1441 (16.33) [24.00]	1420 (14.68)
		$\nu_{49}$ (A')	1466 (27.70) [29.28]	1430 (18.64)
1452 vw	1451 s	$\nu_{50}$ (A'')	1491 (9.40) [9.75]	1448 (7.77)
1460 sh		$\nu_{51}$ (A')	1503 (31.09) [27.83]	1470 (29.33)
		$\nu_{52}$ (A')	1511 (31.42) [1.68]	1480 (18.26)
1501 vw	1500s	$\nu_{53}$ (A')	1541 (10.53) [35.17]	1507 (9.20)
		$\nu_{54}$ (A')	1559 (43.09) [3.75]	1527 (23.27)
1579 m				
1592 w	1591 m	$\nu_{55}$ (A')	1662 (1.50) [58.26]	1603 (1.60)
1605 vw	1605 s	$\nu_{56}$ (A')	1669 (21.36) [36.31]	1619 (18.53)
1624 m	1622 sh	$\nu_{57}$ (A')	1684 (23.74) [16.61]	1640 (13.40)

<sup>a</sup> (km/mole) [Å<sup>4</sup>/AMU]    <sup>b</sup> Estimated uncertainties in band positions are within ±3 cm<sup>-1</sup>, Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder.

<sup>c</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*.

<sup>d</sup> Anharmonics ωB97X-D/6-311++G\*\* calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.9 – Low wavenumber Raman ad PA infrared spectra of 1,8-dimethylcarbazole ( $C_2V$ ).

Predicted <sup>a</sup>			Predicted <sup>a</sup>		
Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>
$\nu_1$ (a <sub>2</sub> )	91 (0.00) [0.57]	90 (0.00)	$\nu_{12}$ (b <sub>1</sub> )	354 (43.97) [0.46]	353 (49.16)
$\nu_2$ (b <sub>1</sub> )	93 (4.10) [0.11]	106 (6.16)	$\nu_{13}$ (a <sub>1</sub> )	475 (2.09) [1.27]	475 (3.08)
$\nu_3$ (b <sub>1</sub> )	133 (3.94) [0.25]	232 (0.07)	$\nu_{14}$ (b <sub>2</sub> )	491 (0.002) [4.23]	487 (0.17)
$\nu_4$ (a <sub>2</sub> )	135 (0.00) [0.16]	241 (0.00)	$\nu_{15}$ (b <sub>1</sub> )	508 (2.97) [0.18]	494 (3.00)
$\nu_5$ (a <sub>1</sub> )	158 (0.18) [0.01]	179 (19.22)	$\nu_{16}$ (a <sub>2</sub> )	534 (0.00) [0.003]	529 (0.00)
$\nu_6$ (b <sub>1</sub> )	196 (0.20) [1.25]	201 (0.59)	$\nu_{17}$ (a <sub>1</sub> )	565 (0.34) [23.58]	562 (0.04)
$\nu_7$ (a <sub>2</sub> )	239 (0.00) [0.64]	243 (0.00)	$\nu_{18}$ (b <sub>1</sub> )	589 (5.40) [0.00]	579 (3.22)
$\nu_8$ (b <sub>1</sub> )	255 (29.26) [0.24]	313 (12.24)	$\nu_{19}$ (b <sub>2</sub> )	589 (11.45) [0.93]	592 (6.60)
$\nu_9$ (b <sub>2</sub> )	273 (0.04) [2.40]	320 (0.14)	$\nu_{20}$ (a <sub>2</sub> )	598 (0.00) [0.10]	586 (0.00)
$\nu_{10}$ (a <sub>2</sub> )	296 (0.00) [2.50]	290 (0.00)	$\nu_{21}$ (b <sub>2</sub> )	602 (2.22) [7.05]	603 (0.01)
$\nu_{11}$ (a <sub>1</sub> )	317 (0.02) [5.95]	325 (0.22)			

<sup>a</sup> (km/mole) [ $A^4/AMU$ ]    <sup>b</sup> Unscaled DFT frequencies calculated using  $\omega$ B97X-D/6-311++G\*\*.<sup>c</sup> Anharmonics PT2 level using  $\omega$ B97X-D/6-311++G\*\*Table C.10 – Raman and PA infrared spectra of 1,8-dimethylcarbazole ( $C_2V$ ), 700–2000  $cm^{-1}$ 

Predicted <sup>a</sup>			Predicted <sup>a</sup>		
Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>
$\nu_{22}$ (a <sub>1</sub> )	713 (0.04) [13.67]	706 (0.13)	$\nu_{44}$ (b <sub>2</sub> )	1264 (29.72) [1.65]	1276 (157.09)
$\nu_{23}$ (a <sub>2</sub> )	766 (0.00) [0.54]	752 (0.00)	$\nu_{45}$ (a <sub>1</sub> )	1278 (0.23) [16.08]	1253 (0.44)
$\nu_{24}$ (b <sub>1</sub> )	770 (32.85) [0.10]	758 (37.61)	$\nu_{46}$ (b <sub>2</sub> )	1284 (13.33) [0.16]	1279 (29.43)
$\nu_{25}$ (b <sub>1</sub> )	786 (82.35) [0.50]	773 (42.18)	$\nu_{47}$ (a <sub>1</sub> )	1313 (0.10) [10.39]	1292 (0.06)
$\nu_{26}$ (b <sub>2</sub> )	816 (0.10) [0.11]	808 (0.09)	$\nu_{48}$ (b <sub>2</sub> )	1354 (108.32) [6.62]	1328 (56.40)
$\nu_{27}$ (a <sub>2</sub> )	823 (0.00) [1.12]	803 (0.00)	$\nu_{49}$ (a <sub>1</sub> )	1356 (7.10) [352.11]	1329 (0.15)
$\nu_{28}$ (b <sub>2</sub> )	881 (4.73) [1.75]	888 (6.98)	$\nu_{50}$ (a <sub>1</sub> )	1372 (9.46) [11.13]	1344 (6.19)
$\nu_{29}$ (a <sub>1</sub> )	881 (0.03) [15.77]	873 (0.00)	$\nu_{51}$ (b <sub>2</sub> )	1420 (1.46) [1.60]	1390 (0.46)
$\nu_{30}$ (b <sub>1</sub> )	921 (0.01) [0.00]	905 (0.00)	$\nu_{52}$ (a <sub>1</sub> )	1421 (8.57) [17.44]	1392 (1.96)
$\nu_{31}$ (a <sub>2</sub> )	925 (0.00) [0.37]	909 (0.00)	$\nu_{53}$ (b <sub>2</sub> )	1441 (12.01) [7.87]	1436 (4.12)
$\nu_{32}$ (a <sub>1</sub> )	972 (5.34) [45.89]	963 (2.79)	$\nu_{54}$ (a <sub>1</sub> )	1463 (31.26) [77.86]	1433 (21.13)
$\nu_{33}$ (a <sub>2</sub> )	986 (0.00) [0.08]	975 (0.00)	$\nu_{55}$ (b <sub>2</sub> )	1471 (35.37) [0.46]	1442 (13.63)
$\nu_{34}$ (b <sub>1</sub> )	988 (0.68) [0.06]	975 (1.22)	$\nu_{56}$ (a <sub>2</sub> )	1490 (0.00) [9.49]	1460 (0.00)
$\nu_{35}$ (b <sub>2</sub> )	1004 (0.70) [0.04]	1001 (0.38)	$\nu_{57}$ (b <sub>1</sub> )	1491 (19.29) [9.45]	1464 (10.73)
$\nu_{36}$ (a <sub>1</sub> )	1053 (0.14) [6.18]	1038 (0.23)	$\nu_{58}$ (a <sub>1</sub> )	1509 (5.02) [7.43]	1473 (2.80)
$\nu_{37}$ (a <sub>2</sub> )	1064 (0.00) [0.00]	1039 (0.00)	$\nu_{59}$ (b <sub>2</sub> )	1512 (38.75) [0.06]	1480 (27.68)
$\nu_{38}$ (b <sub>1</sub> )	1065 (2.32) [0.03]	1040 (2.60)	$\nu_{60}$ (a <sub>1</sub> )	1550 (2.98) [17.22]	1518 (2.95)
$\nu_{39}$ (b <sub>2</sub> )	1103 (15.14) [0.53]	1089 (15.64)	$\nu_{61}$ (b <sub>2</sub> )	1563 (49.41) [1.23]	1534 (23.75)
$\nu_{40}$ (a <sub>1</sub> )	1108 (15.19) [22.18]	1095 (10.73)	$\nu_{62}$ (b <sub>2</sub> )	1667 (26.11) [23.32]	1628 (9.28)
$\nu_{41}$ (b <sub>2</sub> )	1151 (7.53) [3.77]	1163 (1.31)	$\nu_{63}$ (a <sub>1</sub> )	1669 (5.32) [94.50]	1626 (1.89)
$\nu_{42}$ (a <sub>1</sub> )	1192 (0.03) [0.59]	1178 (0.02)	$\nu_{64}$ (b <sub>2</sub> )	1685 (12.12) [0.53]	1654 (2.23)
$\nu_{43}$ (b <sub>2</sub> )	1197 (0.15) [1.27]	1186 (0.28)	$\nu_{65}$ (a <sub>1</sub> )	1697 (1.53) [178.61]	1663 (0.02)

<sup>a</sup> (km/mole) [ $A^4/AMU$ ]    <sup>b</sup> Unscaled DFT frequencies calculated using  $\omega$ B97X-D/6-311++G\*\*.<sup>c</sup> Anharmonics PT2 level using  $\omega$ B97X-D/6-311++G\*\*

Table C.11 – Low wavenumber Raman ad PA infrared spectra of Benzofuran (Cs).

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
212 w	211 m	$\nu_1$ (A'')	218 (6.43) [0.21]	215 (6.13)
246 w	246 w	$\nu_2$ (A'')	254 (0.28) [0.29]	249 (0.27)
403 w	400 m	$\nu_3$ (A')	411 (5.68) [0.14]	407 (5.40)
	418 m	$\nu_4$ (A'')	433 (8.48) [0.38]	422 (8.90)
540 m	539 vw	$\nu_5$ (A')	553 (0.51) [6.04]	546 (0.36)
571 w	569 vw	$\nu_6$ (A'')	586 (0.14) [0.23]	573 (0.20)
586 w	584 vw	$\nu_7$ (A'')	605 (0.04) [1.58]	598 (0.03)
611 m	612 w	$\nu_8$ (A')	628 (1.33) [7.97]	620 (1.24)

<sup>a</sup> (km/mole) [A<sup>4</sup>/AMU]    <sup>b</sup> Ref [281]    <sup>c</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*.    <sup>d</sup> Anharmonics ωB97X-D/6-311++G\*\* calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.12 – Raman and PA infrared spectra of Benzofuran (Cs), 700–2000 cm<sup>1</sup>

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
736 w	732 m	$\nu_9$ (A'')	755 (0.60) [0.38]	738 (1.81)
	746 vs	$\nu_{10}$ (A'')	772 (107.05) [0.16]	755 (95.74)
763 s	763 w	$\nu_{11}$ (A')	788 (5.86) [22.38]	765 (5.13)
	768 m	$\nu_{12}$ (A'')	794 (34.92) [0.09]	773 (28.22)
	848 sh	$\nu_{14}$ (A')	881 (8.73) [3.47]	859 (7.12)
	855 w	$\nu_{13}$ (A'')	878 (0.55) [0.05]	866 (0.26)
862 w	861 m	$\nu_{15}$ (A'')	902 (0.05) [2.17]	869 (0.51)
900 w	900 m	$\nu_{16}$ (A')	922 (3.26) [5.26]	908 (2.22)
932 w	929 m	$\nu_{17}$ (A'')	965 (2.69) [0.04]	949 (2.70)
970 w	970 m	$\nu_{18}$ (A'')	997 (0.02) [0.20]	982 (0.07)
1006 s	1007 m	$\nu_{19}$ (A'')	1042 (6.65) [24.87]	1013 (5.84)
1036m	1036 s	$\nu_{20}$ (A')	1074 (13.65) [13.91]	1048 (10.37)
1106 vw	1107 s	$\nu_{21}$ (A')	1137 (14.94) [4.74]	1113 (12.76)
1130 m	1131 s	$\nu_{22}$ (A')	1167 (23.12) [6.05]	1146 (13.86)
1169 vw	1161 w	$\nu_{23}$ (A')	1183 (4.31) [2.99]	1159 (3.06)
1175 vw	1179 s	$\nu_{24}$ (A')	1212 (22.39) [1.23]	1188 (20.65)
1252 vs	1253 vs	$\nu_{25}$ (A')	1299 (8.68) [15.38]	1264 (9.96)
1264 m	1264 m	$\nu_{26}$ (A')	1301 (33.80) [34.53]	1274 (32.76)
1327 w	1329 m	$\nu_{27}$ (A')	1370 (4.97) [25.99]	1340 (4.95)
1344 m	1346 m	$\nu_{28}$ (A')	1393 (1.77) [15.79]	1355 (1.58)

Observed			Predicted <sup>a</sup>	
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
1456 w	1457 vs	$\nu_{29}$ (A')	1501 (40.40) [11.13]	1468 (34.46)
1475 vw	1479 m	$\nu_{30}$ (A')	1524 (4.32) [4.41]	1490 (3.90)
1543 vs	1543 s	$\nu_{31}$ (A')	1603 (17.91) [115.23]	1560 (7.89)
1595 vw	1595 w	$\nu_{32}$ (A')	1663 (0.13) [12.63]	1620 (0.40)
1616 s	1617 w	$\nu_{33}$ (A')	1691 (0.57) [52.11]	1636 (0.42)

<sup>a</sup> (km/mole) [A<sup>4</sup>/AMU]    <sup>b</sup> Ref [281]    <sup>c</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*.    <sup>d</sup> Anharmonics ωB97X-D/6-311++G\*\* calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.13 – Low wavenumber Raman ad PA infrared spectra of Dibenzofuran (C<sub>2</sub>v).

Observed			Predicted <sup>a</sup>	
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
104	103	$\nu_1$ (b <sub>1</sub> )	105 (2.18) [0.25]	104 (2.80)
	150	$\nu_2$ (a <sub>2</sub> )	152 (0.00) [0.02]	150 (0.00)
217	217	$\nu_3$ (a <sub>1</sub> )	222 (1.19) [0.77]	217 (1.21)
286	287	$\nu_4$ (a <sub>2</sub> )	294 (0.00) [3.60]	291 (0.00)
	310	$\nu_5$ (b <sub>1</sub> )	318 (0.01) [1.21]	317 (0.03)
	419	$\nu_6$ (b <sub>1</sub> )	434 (6.24) [0.003]	427 (6.93)
424	424	$\nu_7$ (a <sub>1</sub> )	435 (1.88) [8.04]	430 (1.78)
	440	$\nu_8$ (a <sub>2</sub> )	445 (0.00) [0.21]	442 (0.00)
516	516	$\nu_9$ (b <sub>2</sub> )	531 (0.16) [2.37]	523 (0.14)
555	556	$\nu_{10}$ (b <sub>2</sub> )	573 (3.06) [9.01]	560 (3.10)
	562	$\nu_{11}$ (b <sub>1</sub> )	580 (4.29) [0.24]	568 (5.83)
	576	$\nu_{12}$ (a <sub>2</sub> )	586 (0.00) [0.01]	580 (0.00)
	617	$\nu_{13}$ (b <sub>2</sub> )	634 (5.78) [0.00]	624 (5.74)
660	660	$\nu_{14}$ (a <sub>1</sub> )	679 (0.46) [0.95]	671 (0.42)

<sup>a</sup> (km/mole) [A<sup>4</sup>/AMU]    <sup>b</sup> Ref [227]    <sup>c</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*.    <sup>d</sup> Anharmonics ωB97X-D/6-311++G\*\* calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.14 – Raman and PA infrared spectra of Dibenzofuran ( $C_2V$ ), 700–2000  $\text{cm}^{-1}$ 

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
745	724	$\nu_{15}$ (b <sub>1</sub> )	747 (55.66) [0.004]	725 (28.54)
	746	$\nu_{17}$ (a <sub>1</sub> )	767 (0.02) [29.87]	755 (0.08)
	752	$\nu_{19}$ (b <sub>1</sub> )	779 (112.07) [0.02]	763 (67.49)
	769	$\nu_{16}$ (a <sub>2</sub> ) (a <sub>2</sub> )	764 (0.00) [0.02] 774 (0.00) [1.57]	762 (0.00) 770 (0.00)
	849	$\nu_{20}$ (b <sub>2</sub> ) $\nu_{23}$ (b <sub>1</sub> )	876 (23.98) [1.31] 878 (0.88) [0.11]	857 (18.10) 858 (0.54)
	850	$\nu_{22}$ (a <sub>1</sub> )	876 (7.30) [21.75]	861 (5.72)
	855			
		$\nu_{21}$ (a <sub>2</sub> )	876 (0.00) [0.10]	872 (0.00)
	890			
	930	$\nu_{24}$ (b <sub>1</sub> )	966 (3.64) [0.02]	942 (2.77)
1010	931	$\nu_{25}$ (a <sub>2</sub> )	966 (0.00) [0.07]	945 (0.00)
	967	$\nu_{26}$ (a <sub>2</sub> ) $\nu_{27}$ (b <sub>1</sub> )	999 (0.00) [0.28] 999 (1.28) [0.04]	1003 (0.00) 979 (2.34)
	1005	$\nu_{28}$ (b <sub>2</sub> )	1031 (1.78) [0.15]	1005 (1.03)
	1010	$\nu_{29}$ (a <sub>1</sub> )	1045 (0.64) [59.37]	1017 (0.52)
	1024	$\nu_{30}$ (b <sub>2</sub> )	1051 (14.22) [0.00]	1031 (8.32)
	1102	$\nu_{31}$ (a <sub>1</sub> )	1134 (10.74) [10.10]	1109 (7.61)
	1114	$\nu_{32}$ (b <sub>2</sub> )	1146 (8.85) [3.38]	1121 (7.40)
	1148	$\nu_{33}$ (a <sub>1</sub> )	1179 (2.67) [15.79]	1157 (2.54)
	1153	$\nu_{34}$ (b <sub>2</sub> )	1187 (0.07) [2.24]	1160 (0.05)
	1170	$\nu_{35}$ (a <sub>1</sub> )	1231 (0.48) [17.72]	1200 (0.25)
1242	1201	$\nu_{36}$ (b <sub>2</sub> )	1247 (161.74) [6.66]	1212 (110.56)
	1245	$\nu_{37}$ (a <sub>1</sub> )	1290 (9.07) [150.19]	1260 (8.14)
	1286	$\nu_{38}$ (b <sub>2</sub> )	1316 (22.32) [2.00]	1299 (13.58)
	1306	$\nu_{39}$ (a <sub>1</sub> )	1342 (2.56) [124.60]	1318 (1.46)
	1319	$\nu_{40}$ (b <sub>2</sub> )	1365 (14.72) [4.22]	1326 (10.74)
	1339	$\nu_{41}$ (a <sub>1</sub> )	1387 (5.81) [38.73]	1359 (3.63)
	1446	$\nu_{42}$ (a <sub>1</sub> )	1493 (41.84) [25.46]	1459 (25.64)
	1458	$\nu_{43}$ (b <sub>2</sub> )	1499 (31.37) [0.41]	1466 (20.78)
	1477	$\nu_{44}$ (b <sub>2</sub> )	1521 (49.24) [8.37]	1490 (29.11)
	1489	$\nu_{45}$ (a <sub>1</sub> )	1541 (1.49) [58.41]	1503 (0.77)
1633	1593	$\nu_{46}$ (b <sub>2</sub> )	1661 (8.69) [6.62]	1615 (5.49)
	1595	$\nu_{47}$ (a <sub>1</sub> )	1666 (1.42) [22.56]	1621 (0.96)
	1599	$\nu_{48}$ (b <sub>2</sub> )	1673 (10.97) [3.66]	1630 (8.42)
	1636	$\nu_{49}$ (a <sub>1</sub> )	1711 (0.67) [395.42]	1655 (0.58)

<sup>a</sup> (km/mole) [ $\text{A}^4/\text{AMU}$ ]    <sup>b</sup> Ref [227]    <sup>c</sup> Unscaled DFT frequencies calculated using  $\omega\text{B97X-D}/6-311++\text{G}^{**}$ .    <sup>d</sup> Anharmonics  $\omega\text{B97X-D}/6-311++\text{G}^{**}$  calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.15 – Low wavenumber Raman ad PA infrared spectra of Tribenzofuran ( $C_2V$ ).

Predicted <sup>a</sup>			Predicted <sup>a</sup>		
Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>
$\nu_1$ (b <sub>1</sub> )	83 (3.91) [0.09]	88 (4.58)	$\nu_{10}$ (b <sub>1</sub> )	503 (3.49) [0.09]	525 (3.55)
$\nu_2$ (a <sub>2</sub> )	197 (0.00) [0.00]	197 (0.00)	$\nu_{11}$ (b <sub>2</sub> )	558 (0.59) [3.77]	556 (0.57)
$\nu_3$ (b <sub>1</sub> )	212 (2.30) [0.00]	210 (2.65)	$\nu_{12}$ (a <sub>2</sub> )	575 (0.00) [0.18]	633 (0.00)
$\nu_4$ (a <sub>2</sub> )	299 (0.00) [1.02]	330 (0.00)	$\nu_{13}$ (b <sub>2</sub> )	586 (3.58) [1.01]	580 (4.19)
$\nu_5$ (b <sub>1</sub> )	324 (0.12) [1.91]	323 (0.11)	$\nu_{14}$ (b <sub>1</sub> )	610 (2.71) [0.05]	610 (3.02)
$\nu_6$ (a <sub>1</sub> )	347 (5.52) [0.33]	344 (5.34)	$\nu_{15}$ (a <sub>1</sub> )	622 (3.45) [32.59]	619 (2.96)
$\nu_7$ (a <sub>2</sub> )	434 (0.00) [1.86]	480 (0.00)	$\nu_{16}$ (a <sub>2</sub> )	660 (0.00) [0.11]	669 (0.00)
$\nu_8$ (a <sub>1</sub> )	457 (0.001) [23.05]	456 (0.01)	$\nu_{17}$ (a <sub>1</sub> )	690 (0.32) [12.35]	683 (0.48)
$\nu_9$ (b <sub>2</sub> )	490 (2.50) [3.08]	487 (2.41)	$\nu_{18}$ (b <sub>2</sub> )	706 (1.10) [0.00]	705 (0.65)

<sup>a</sup> (km/mole) [ $A^4/AMU$ ]    <sup>b</sup> Unscaled DFT frequencies calculated using  $\omega$ B97X-D/6-311++G\*\*.<sup>c</sup> Anharmonics PT2 level using  $\omega$ B97X-D/6-311++G\*\*Table C.16 – Raman ad PA infrared spectra of Tribenzofuran ( $C_2V$ ), 700 - 2000  $cm^{-1}$ 

Predicted <sup>a</sup>			Predicted <sup>a</sup>		
Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>
$\nu_{19}$ (b <sub>1</sub> )	733 (38.39) [0.13]	729 (54.45)	$\nu_{38}$ (a <sub>1</sub> )	1205 (0.09) [1.16]	1191 (0.28)
$\nu_{20}$ (a <sub>2</sub> )	763 (0.00) [0.01]	928 (0.00)	$\nu_{39}$ (a <sub>1</sub> )	1241 (10.71) [96.39]	1223 (3.01)
$\nu_{21}$ (b <sub>1</sub> )	778 (39.68) [0.01]	761 (28.8)	$\nu_{40}$ (a <sub>1</sub> )	1253 (2.89) [7.06]	1229 (2.81)
$\nu_{22}$ (a <sub>2</sub> )	803 (0.00) [1.76]	892 (0.00)	$\nu_{41}$ (b <sub>2</sub> )	1256 (22.50) [6.29]	1235 (25.32)
$\nu_{23}$ (b <sub>2</sub> )	815 (4.99) [1.26]	809 (5.73)	$\nu_{42}$ (b <sub>2</sub> )	1275 (35.29) [2.39]	1251 (0.90)
$\nu_{24}$ (a <sub>1</sub> )	838 (3.73) [14.15]	835 (4.01)	$\nu_{43}$ (a <sub>1</sub> )	1351 (20.17) [89.92]	1327 (7.21)
$\nu_{25}$ (b <sub>1</sub> )	845 (98.70) [0.26]	836 (65.69)	$\nu_{44}$ (a <sub>1</sub> )	1402 (4.04) [12.33]	1381 (4.87)
$\nu_{26}$ (a <sub>2</sub> )	910 (0.00) [0.39]	905 (0.00)	$\nu_{45}$ (b <sub>2</sub> )	1404 (18.31) [0.05]	1373 (10.86)
$\nu_{27}$ (b <sub>1</sub> )	916 (0.18) [0.18]	912 (0.17)	$\nu_{46}$ (b <sub>2</sub> )	1440 (0.53) [34.15]	1418 (0.98)
$\nu_{28}$ (a <sub>2</sub> )	988 (0.00) [0.01]	999 (0.00)	$\nu_{47}$ (b <sub>2</sub> )	1466 (30.32) [0.34]	1440 (30.42)
$\nu_{29}$ (b <sub>1</sub> )	989 (1.66) [0.00]	997 (2.90)	$\nu_{48}$ (a <sub>1</sub> )	1490 (20.14) [29.62]	1462 (16.82)
$\nu_{30}$ (a <sub>2</sub> )	1002 (0.00) [1.83]	992 (0.00)	$\nu_{49}$ (a <sub>1</sub> )	1533 (0.02) [386.40]	1505 (0.02)
$\nu_{31}$ (b <sub>2</sub> )	1025 (5.33) [3.29]	1023 (44.24)	$\nu_{50}$ (b <sub>2</sub> )	1541 (0.43) [0.97]	1507 (0.02)
$\nu_{32}$ (a <sub>1</sub> )	1025 (6.91) [0.49]	1015 (34.94)	$\nu_{51}$ (a <sub>1</sub> )	1562 (24.90) [2.51]	1528 (11.38)
$\nu_{33}$ (a <sub>1</sub> )	1045 (0.53) [47.16]	1029 (0.28)	$\nu_{52}$ (b <sub>2</sub> )	1657 (65.43) [1.97]	1623 (47.09)
$\nu_{34}$ (b <sub>2</sub> )	1072 (34.49) [4.59]	1059 (81.42)	$\nu_{53}$ (a <sub>1</sub> )	1697 (1.81) [44.60]	1664 (1.34)
$\nu_{35}$ (b <sub>2</sub> )	1075 (68.06) [1.25]	1089 (2.52)	$\nu_{54}$ (b <sub>2</sub> )	1721 (2.21) [112.66]	1690 (1.72)
$\nu_{36}$ (a <sub>1</sub> )	1164 (0.17) [4.28]	1150 (0.16)	$\nu_{55}$ (a <sub>1</sub> )	1750 (0.11) [38.19]	1714 (0.31)
$\nu_{37}$ (b <sub>2</sub> )	1197 (5.66) [0.23]	1180 (4.97)			

<sup>a</sup> (km/mole) [ $A^4/AMU$ ]    <sup>b</sup> Unscaled DFT frequencies calculated using  $\omega$ B97X-D/6-311++G\*\*.<sup>c</sup> Anharmonics PT2 level using  $\omega$ B97X-D/6-311++G\*\*

Table C.17 – Low wavenumber Raman ad PA infrared spectra of 4-methyldibenzofuran (Cs).

Predicted <sup>a</sup>			Predicted <sup>a</sup>		
Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>
$\nu_1$ (A'')	88 (0.07) [0.21]	200 (64.55)	$\nu_{10}$ (A')	444 (2.40) [4.21]	441 (2.27)
$\nu_2$ (A'')	99 (2.60) [0.07]	113 (1.42)	$\nu_{11}$ (A'')	517 (0.21) [9.41]	513 (0.27)
$\nu_3$ (A'')	125 (0.14) [0.68]	130 (0.40)	$\nu_{12}$ (A')	517 (1.70) [0.22]	551 (2.13)
$\nu_4$ (A')	182 (0.89) [0.69]	180 (0.85)	$\nu_{13}$ (A')	580 (0.95) [1.66]	577 (0.74)
$\nu_5$ (A'')	224 (1.63) [0.51]	227 (1.50)	$\nu_{14}$ (A'')	583 (1.74) [0.10]	580 (1.95)
$\nu_6$ (A'')	295 (0.07) [3.18]	298 (0.05)	$\nu_{15}$ (A'')	589 (0.01) [0.04]	600 (0.21)
$\nu_7$ (A')	301 (0.49) [3.36]	300 (0.48)	$\nu_{16}$ (A')	597 (6.95) [8.78]	593 (6.46)
$\nu_8$ (A'')	328 (0.56) [1.40]	329 (0.79)	$\nu_{17}$ (A')	658 (1.69) [3.97]	653 (1.27)
$\nu_9$ (A'')	440 (3.26) [0.03]	442 (3.54)			

<sup>a</sup> (km/mole) [A<sup>4</sup>/AMU]    <sup>b</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*.<sup>c</sup> Anharmonics PT2 level using ωB97X-D/6-311++G\*\*Table C.18 – Raman ad PA infrared spectra of 4-methyldibenzofuran (Cs), 700 - 2000 cm<sup>-1</sup>

Predicted <sup>a</sup>			Predicted <sup>a</sup>		
Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>
$\nu_{18}$ (A')	721 (0.32) [18.49]	714 (0.42)	$\nu_{38}$ (A')	1197 (0.06) [0.48]	1174 (0.07)
$\nu_{19}$ (A'')	756 (20.32) [0.27]	758 (14.79)	$\nu_{39}$ (A')	1241 (153.34) [3.34]	1216 (135.85)
$\nu_{20}$ (A'')	767 (2.91) [0.03]	788 (29.63)	$\nu_{40}$ (A')	1255 (7.92) [2.88]	1236 (2.71)
$\nu_{21}$ (A'')	777 (106.12) [0.08]	771 (126.6)	$\nu_{41}$ (A')	1297 (6.22) [99.05]	1276 (4.91)
$\nu_{22}$ (A'')	801 (13.45) [1.13]	827 (7.30)	$\nu_{42}$ (A')	1306 (26.35) [25.62]	1283 (1.51)
$\nu_{23}$ (A')	859 (7.32) [14.37]	846 (9.67)	$\nu_{43}$ (A')	1345 (4.72) [173.57]	1314 (1.76)
$\nu_{24}$ (A')	872 (18.11) [6.47]	859 (12.42)	$\nu_{44}$ (A')	1364 (11.36) [10.71]	1335 (0.84)
$\nu_{25}$ (A'')	878 (0.37) [0.08]	885 (0.82)	$\nu_{45}$ (A')	1387 (4.79) [33.80]	1359 (2.80)
$\nu_{26}$ (A'')	921 (8.71) [9.69]	912 (6.58)	$\nu_{46}$ (A')	1425 (3.99) [9.49]	1399 (3.17)
$\nu_{27}$ (A')	922 (0.002) [0.15]	915 (0.00)	$\nu_{47}$ (A')	1459 (20.27) [25.41]	1430 (9.30)
$\nu_{28}$ (A'')	966 (1.69) [0.09]	964 (1.92)	$\nu_{48}$ (A'')	1485 (9.00) [9.58]	1459 (6.16)
$\nu_{29}$ (A'')	992 (0.39) [0.10]	988 (0.43)	$\nu_{49}$ (A')	1498 (45.19) [16.12]	1465 (14.98)
$\nu_{30}$ (A'')	998 (0.003) [0.09]	1026 (0.50)	$\nu_{50}$ (A')	1506 (26.95) [2.29]	1471 (15.95)
$\nu_{31}$ (A')	1015 (2.01) [21.55]	1002 (1.14)	$\nu_{51}$ (A')	1526 (23.32) [24.46]	1499 (8.05)
$\nu_{32}$ (A')	1046 (7.93) [26.83]	1024 (5.42)	$\nu_{52}$ (A')	1551 (7.49) [21.68]	1520 (8.84)
$\nu_{33}$ (A'')	1067 (1.57) [0.002]	1056 (1.53)	$\nu_{53}$ (A')	1661 (11.64) [15.22]	1623 (9.83)
$\nu_{34}$ (A')	1095 (10.33) [1.63]	1075 (7.55)	$\nu_{54}$ (A')	1669 (4.23) [5.71]	1633 (2.62)
$\nu_{35}$ (A')	1123 (14.02) [21.38]	1107 (11.8)	$\nu_{55}$ (A')	1684 (2.13) [13.84]	1644 (3.12)
$\nu_{36}$ (A')	1152 (6.53) [3.53]	1133 (6.3)	$\nu_{56}$ (A')	1708 (2.42) [396.11]	1672 (1.77)
$\nu_{37}$ (A')	1185 (0.71) [4.01]	1170 (0.92)			

<sup>a</sup> (km/mole) [A<sup>4</sup>/AMU]    <sup>b</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*.<sup>c</sup> Anharmonics PT2 level using ωB97X-D/6-311++G\*\*

Table C.19 – Low wavenumber Raman ad PA infrared spectra of 4,6-dimethyldibenzofuran ( $C_2V$ ).

Predicted <sup>a</sup>			Predicted <sup>a</sup>		
Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>
$\nu_1$ (b <sub>1</sub> )	86 (0.19) [0.37]	101 (1.42)	$\nu_{12}$ (a <sub>1</sub> )	472 (2.72) [1.79]	472 (2.46)
$\nu_2$ (a <sub>2</sub> )	86 (0.00) [0.07]	83 (0.11)	$\nu_{13}$ (b <sub>2</sub> )	493 (0.00) [3.97]	491 (0.02)
$\nu_3$ (a <sub>2</sub> )	96 (0.00) [0.55]	103 (0.00)	$\nu_{14}$ (b <sub>1</sub> )	504 (1.23) [0.25]	513 (1.14)
$\nu_4$ (b <sub>1</sub> )	104 (3.18) [0.06]	100 (2.72)	$\nu_{15}$ (a <sub>2</sub> )	531 (0.00) [0.05]	555 (0.00)
$\nu_5$ (a <sub>1</sub> )	155 (0.50) [0.07]	152 (0.44)	$\nu_{16}$ (a <sub>1</sub> )	571 (1.22) [23.04]	563 (1.09)
$\nu_6$ (b <sub>1</sub> )	201 (2.71) [0.91]	193 (2.36)	$\nu_{17}$ (b <sub>1</sub> )	587 (0.44) [0.03]	583 (0.15)
$\nu_7$ (a <sub>2</sub> )	240 (0.00) [0.60]	235 (0.00)	$\nu_{18}$ (a <sub>2</sub> )	591 (0.00) [0.09]	598 (0.01)
$\nu_8$ (b <sub>2</sub> )	266 (1.30) [2.66]	289 (0.86)	$\nu_{19}$ (b <sub>2</sub> )	597 (7.99) [1.59]	590 (6.00)
$\nu_9$ (a <sub>2</sub> )	297 (0.00) [2.90]	292 (0.00)	$\nu_{20}$ (b <sub>2</sub> )	606 (1.03) [5.33]	597 (0.54)
$\nu_{10}$ (a <sub>1</sub> )	319 (0.08) [3.97]	318 (0.01)	$\nu_{21}$ (a <sub>1</sub> )	710 (0.61) [13.68]	697 (0.68)
$\nu_{11}$ (b <sub>1</sub> )	337 (2.07) [1.58]	329 (2.75)			

<sup>a</sup> (km/mole) [ $A^4/AMU$ ]    <sup>b</sup> Unscaled DFT frequencies calculated using  $\omega$ B97X-D/6-311++G\*\*.<sup>c</sup> Anharmonics PT2 level using  $\omega$ B97X-D/6-311++G\*\*Table C.20 – Raman ad PA infrared spectra of 4,6-dimethyldibenzofuran ( $C_2V$ ), 700 - 2000  $cm^{-1}$ 

Predicted <sup>a</sup>			Predicted <sup>a</sup>		
Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>
$\nu_{22}$ (a <sub>2</sub> )	766 (0.00) [0.11]	770 (0.06)	$\nu_{43}$ (b <sub>2</sub> )	1240 (149.71) [2.17]	1212 (137.58)
$\nu_{23}$ (b <sub>1</sub> )	769 (17.48) [0.12]	763 (37.66)	$\nu_{44}$ (a <sub>1</sub> )	1261 (2.13) [2.66]	1238 (0.63)
$\nu_{24}$ (b <sub>1</sub> )	788 (102.32) [0.33]	778 (66.44)	$\nu_{45}$ (a <sub>1</sub> )	1297 (2.69) [44.94]	1278 (1.35)
$\nu_{25}$ (a <sub>2</sub> )	808 (0.00) [1.17]	840 (0.00)	$\nu_{46}$ (b <sub>2</sub> )	1304 (21.15) [2.10]	1288 (12.24)
$\nu_{26}$ (b <sub>2</sub> )	816 (2.00) [0.05]	807 (1.4)	$\nu_{47}$ (a <sub>1</sub> )	1347 (8.87) [26.63]	1321 (4.27)
$\nu_{27}$ (a <sub>1</sub> )	865 (3.32) [16.71]	856 (3.20)	$\nu_{48}$ (b <sub>2</sub> )	1362 (10.07) [3.49]	1334 (1.26)
$\nu_{28}$ (b <sub>2</sub> )	887 (25.33) [2.08]	870 (23.80)	$\nu_{49}$ (a <sub>1</sub> )	1385 (3.85) [29.32]	1363 (3.77)
$\nu_{29}$ (b <sub>1</sub> )	920 (0.003) [0.01]	909 (0.04)	$\nu_{50}$ (b <sub>2</sub> )	1424 (0.77) [0.45]	1388 (0.81)
$\nu_{30}$ (a <sub>2</sub> )	922 (0.00) [0.26]	918 (0.00)	$\nu_{51}$ (a <sub>1</sub> )	1425 (7.56) [18.01]	1390 (4.59)
$\nu_{31}$ (a <sub>1</sub> )	968 (10.39) [39.50]	956 (6.70)	$\nu_{52}$ (a <sub>1</sub> )	1453 (21.91) [46.21]	1425 (18.23)
$\nu_{32}$ (a <sub>2</sub> )	990 (0.00) [0.005]	997 (0.00)	$\nu_{53}$ (b <sub>2</sub> )	1465 (20.31) [4.22]	1436 (6.05)
$\nu_{33}$ (b <sub>1</sub> )	992 (0.80) [0.11]	991 (0.44)	$\nu_{54}$ (a <sub>2</sub> )	1485 (0.00) [8.55]	1446 (0.00)
$\nu_{34}$ (b <sub>2</sub> )	1012 (0.39) [0.08]	999 (0.05)	$\nu_{55}$ (b <sub>1</sub> )	1486 (18.52) [10.12]	1451 (10.00)
$\nu_{35}$ (a <sub>1</sub> )	1052 (0.22) [8.94]	1033 (0.37)	$\nu_{56}$ (a <sub>1</sub> )	1507 (10.27) [7.32]	1469 (11.18)
$\nu_{36}$ (a <sub>2</sub> )	1066 (0.00) [0.00]	1044 (0.02)	$\nu_{57}$ (b <sub>2</sub> )	1507 (47.61) [0.78]	1471 (42.33)
$\nu_{37}$ (b <sub>1</sub> )	1067 (3.08) [0.01]	1042 (2.00)	$\nu_{58}$ (b <sub>2</sub> )	1544 (19.87) [6.64]	1517 (13.60)
$\nu_{38}$ (b <sub>2</sub> )	1102 (20.85) [1.56]	1086 (15.62)	$\nu_{59}$ (a <sub>1</sub> )	1550 (2.77) [20.51]	1520 (3.65)
$\nu_{39}$ (a <sub>1</sub> )	1109 (11.35) [20.47]	1095 (8.68)	$\nu_{60}$ (b <sub>2</sub> )	1661 (10.82) [7.55]	1628 (3.09)
$\nu_{40}$ (b <sub>2</sub> )	1175 (7.49) [0.87]	1161 (5.15)	$\nu_{61}$ (a <sub>1</sub> )	1682 (4.54) [27.30]	1640 (3.11)
$\nu_{41}$ (a <sub>1</sub> )	1192 (0.00) [1.49]	1176 (0.07)	$\nu_{62}$ (b <sub>2</sub> )	1683 (0.90) [2.87]	1645 (0.81)
$\nu_{42}$ (b <sub>2</sub> )	1210 (10.84) [1.11]	1193 (9.58)	$\nu_{63}$ (a <sub>1</sub> )	1706 (2.50) [387.88]	1677 (1.10)

<sup>a</sup> (km/mole) [ $A^4/AMU$ ]    <sup>b</sup> Unscaled DFT frequencies calculated using  $\omega$ B97X-D/6-311++G\*\*.<sup>c</sup> Anharmonics PT2 level using  $\omega$ B97X-D/6-311++G\*\*

Table C.21 – Low wavenumber Raman ad PA infrared spectra of Indene (Cs).

Observed			Predicted <sup>a</sup>	
Raman <sup>b</sup>	Infrared <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>d</sup>	Anharmonic <sup>e</sup>
196	189	$\nu_1$ (A'')	195 (1.02) [0.52]	193 (1.57)
214	206	$\nu_2$ (A'')	211 (4.05) [0.11]	208 (4.00)
381	380	$\nu_3$ (A')	390 (1.50) [0.89]	384 (1.24)
	388	$\nu_4$ (A'')	397 (3.61) [0.54]	393 (3.09)
	415	$\nu_5$ (A'')	430 (9.64) [0.03]	422 (9.11)
534	534	$\nu_6$ (A')	546 (0.13) [7.22]	543 (0.11)
	549	$\nu_7$ (A'')	567 (5.87) [0.23]	555 (5.13)
591	592	$\nu_8$ (A')	610 (1.63) [6.57]	598 (1.68)
	690	$\nu_9$ (A'')	715 (13.12) [0.47]	694 (13.03)

<sup>a</sup> (km/mole) [A<sup>4</sup>/AMU]<sup>b</sup> Ref [282]    <sup>c</sup> Ref [283]    <sup>d</sup> Unscaled DFT frequencies calculated using  $\omega$ B97X-D/6-311++G\*\*.<sup>e</sup> Anharmonics  $\omega$ B97X-D/6-311++G\*\* calculation using the P\_VMWCI<sub>2</sub> algorithmTable C.22 – Raman and PA infrared spectra of Indene (Cs), 700–2000 cm<sup>1</sup>

Observed			Predicted <sup>a</sup>	
Raman <sup>b</sup>	Infrared <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>d</sup>	Anharmonic <sup>e</sup>
	718	$\nu_{10}$ (A'')	741 (31.30) [0.11]	722 (20.73)
731	731	$\nu_{11}$ (A')	751 (2.35) [12.98]	740 (2.06)
	766	$\nu_{12}$ (A'')	794 (67.06) [0.86]	772 (79.53)
838	834	$\nu_{13}$ (A')	852 (1.43) [10.01]	835 (1.81)
	854	$\nu_{14}$ (A'')	879 (3.07) [4.96]	864 (3.06)
862	862	$\nu_{15}$ (A'')	884 (0.02) [0.36]	871 (0.12)
	917	$\nu_{16}$ (A'')	948 (8.20) [0.54]	926 (10.37)
	926	$\nu_{17}$ (A'')	969 (11.30) [6.65]	940 (9.21)
	943	$\nu_{18}$ (A'')	970 (0.06) [0.37]	954 (0.02)
947	947	$\nu_{19}$ (A')	985 (0.81) [2.89]	966 (1.19)
	972	$\nu_{20}$ (A'')	1009 (0.02) [0.11]	986 (0.00)
1020	1020	$\nu_{21}$ (A')	1054 (5.61) [29.00]	1030 (4.40)
1069	1069	$\nu_{22}$ (A')	1096 (1.83) [9.30]	1075 (1.74)
1106	1111	$\nu_{23}$ (A')	1141 (0.75) [17.11]	1117 (0.58)
1131	1125	$\nu_{24}$ (A'')	1159 (2.01) [5.90]	1126 (2.13)
1154	1154	$\nu_{25}$ (A')	1184 (0.09) [4.29]	1165 (0.23)
	1168	$\nu_{26}$ (A')	1196 (1.32) [1.47]	1170 (0.57)
1206	1206	$\nu_{27}$ (A')	1240 (1.77) [30.07]	1212 (1.56)
1228	1228	$\nu_{28}$ (A')	1264 (3.51) [22.00]	1235 (1.97)
1289	1289	$\nu_{29}$ (A')	1325 (0.47) [9.11]	1295 (0.14)

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
	1316	$\nu_{30}$ (A')	1351 (3.06) [8.69]	1321 (3.16)
1364	1364	$\nu_{31}$ (A')	1399 (2.51) [31.97]	1372 (1.74)
1405	1404	$\nu_{32}$ (A')	1441 (14.05) [19.95]	1406 (10.86)
1460	1462	$\nu_{33}$ (A')	1506 (5.27) [5.11]	1478 (4.43)
1465	1469	$\nu_{34}$ (A')	1511 (10.77) [21.43]	1480 (7.13)
1557	1558	$\nu_{35}$ (A')	1634 (0.97) [97.55]	1585 (1.30)
1591	1592	$\nu_{36}$ (A')	1669 (0.56) [22.83]	1620 (0.59)
1613	1615	$\nu_{37}$ (A')	1690 (1.88) [91.95]	1649 (1.40)

<sup>a</sup> (km/mole) [Å<sup>4</sup>/AMU]    <sup>b</sup> Ref [282]    <sup>c</sup> Ref [283]    <sup>d</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*.    <sup>e</sup> Anharmonics ωB97X-D/6-311++G\*\* calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.23 – Low wavenumber Raman ad PA infrared spectra of Fluorene (C<sub>2</sub>v).

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
	108 m	$\nu_1$ (b <sub>1</sub> )	97 (0.73) [0.12]	101 (0.87)
	122 w			
	140 w	$\nu_2$ (a <sub>2</sub> )	136 (0.00) [0.005]	138 (0.00)
	151 vw			
	172 w			
	217 m	$\nu_3$ (a <sub>1</sub> )	217 (0.20) [0.40]	216 (0.21)
	256 m	$\nu_4$ (b <sub>1</sub> )	242 (7.78) [0.56]	244 (8.09)
	266 m			
	280 s	$\nu_5$ (a <sub>2</sub> )	278 (0.00) [4.37]	280 (0.00)
	302 w			
	367 vw			
	411 s	$\nu_7$ (b <sub>1</sub> )	425 (6.72) [0.005]	420 (6.89)
	417 s	$\nu_6$ (a <sub>1</sub> )	423 (0.41) [7.45]	421 (0.50)
	428 sh	$\nu_8$ (a <sub>2</sub> )	441 (0.00) [0.66]	437 (0.00)
	474 m	$\nu_9$ (b <sub>1</sub> )	482 (0.92) [0.01]	481 (1.30)
	487 w	$\nu_{10}$ (b <sub>2</sub> )	501 (0.45) [0.11]	495 (0.45)
	519 w			
	545 w	$\nu_{11}$ (b <sub>2</sub> )	557 (0.14) [9.59]	552 (0.12)
	562 vw			
	573 vw	$\nu_{12}$ (a <sub>2</sub> )	581 (0.00) [0.06]	575 (0.00)
	587 vw			
	623 w	$\nu_{13}$ (b <sub>2</sub> )	640 (7.27) [0.26]	633 (7.13)
	633 w	$\nu_{14}$ (a <sub>1</sub> )	649 (0.36) [0.18]	645 (0.35)
	650 vw			

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
	667 sh			
	697 s	$\nu_{15}$ (b <sub>1</sub> )	718 (5.67) [0.001]	705 (5.50)

<sup>a</sup> (km/mole) [Å<sup>4</sup>/AMU]

<sup>b</sup> Ref [223], Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder. <sup>c</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*. <sup>d</sup> Anharmonics ωB97X-D/6-311++G\*\* calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.24 – Raman and PA infrared spectra of Fluorene (C<sub>2</sub>v), 700–2000 cm<sup>-1</sup>

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
		$\nu_{16}$ (a <sub>2</sub> )	750 (0.00) [0.38]	733 (0.00)
	741 vs	$\nu_{17}$ (a <sub>1</sub> )	763 (140.76) [0.01]	749 (138.64)
741 vs		$\nu_{18}$ (b <sub>1</sub> )	763 (0.03) [27.38]	750 (0.02)
781 w	775 sh			
786 w	792 vw	$\nu_{19}$ (a <sub>2</sub> )	804 (0.00) [1.90]	795 (0.00)
	820 vw	$\nu_{20}$ (b <sub>2</sub> )	824 (0.27) [0.14]	818 (0.24)
842 m		$\nu_{21}$ (a <sub>1</sub> )	861 (0.08) [29.74]	852 (0.11)
852 w				
864 vw	859 m	$\nu_{22}$ (a <sub>1</sub> )	884 (0.77) [0.20]	866 (0.80)
	881 vw	$\nu_{23}$ (a <sub>2</sub> )	894 (0.00) [0.00]	881 (0.00)
917 vw	910 vw	$\nu_{24}$ (b <sub>1</sub> )	944 (0.17) [0.14]	925 (0.12)
955 vw	953 m	$\nu_{25}$ (a <sub>2</sub> )	973 (0.00) [0.01]	954 (0.00)
		$\nu_{26}$ (b <sub>1</sub> )	989 (4.37) [0.09]	971 (3.18)
973 vw	979 sh	$\nu_{27}$ (a <sub>2</sub> )	1009 (0.00) [0.45]	990 (0.00)
993 vw	997 m	$\nu_{28}$ (b <sub>1</sub> )	1011 (0.06) [0.04]	995 (0.06)
		$\nu_{29}$ (b <sub>2</sub> )	1033 (3.45) [0.48]	1018 (2.64)
1002 vw				
	1020 m	$\nu_{31}$ (b <sub>2</sub> )	1061 (7.81) [0.11]	1030 (6.15)
1020 vs		$\nu_{30}$ (a <sub>1</sub> )	1057 (1.00) [68.27]	1035 (0.92)
	1061 vw			
1092 vw	1091 m	$\nu_{32}$ (a <sub>1</sub> )	1128 (3.94) [8.12]	1105 (2.86)
1107 vw	1108 m	$\nu_{33}$ (b <sub>2</sub> )	1143 (0.12) [0.10]	1120 (0.30)
1115 vw				
1127 vw	1122 sh			
1146 w				
1153 m	1155 m	$\nu_{34}$ (a <sub>2</sub> )	1173 (0.00) [3.46]	1163 (0.00)
1169 vw		$\nu_{35}$ (b <sub>2</sub> )	1182 (1.13) [10.44]	1163 (0.35)
1188 w		$\nu_{36}$ (a <sub>1</sub> )	1187 (0.01) [12.68]	1170 (0.00)

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
1193 w	1188 m	$\nu_{37}$ (b <sub>2</sub> )	1201 (1.66) [0.45]	1190 (2.79)
1210 vw		$\nu_{38}$ (a <sub>1</sub> )	1221 (4.16) [5.97]	1205 (3.34)
1221 vw	1219 vw	$\nu_{39}$ (b <sub>2</sub> )	1235 (4.93) [17.59]	1210 (4.08)
1235 vs	1231 w	$\nu_{40}$ (a <sub>1</sub> )	1268 (1.58) [118.69]	1242 (1.68)
1268 vw				
1292 m	1300 sh	$\nu_{41}$ (a <sub>1</sub> )	1329 (0.50) [126.02]	1305 (0.40)
1300 w	1310 m	$\nu_{42}$ (b <sub>2</sub> )	1336 (3.68) [2.29]	1311 (4.14)
1344 w	1341 m	$\nu_{43}$ (b <sub>2</sub> )	1354 (6.32) [1.98]	1348 (7.01)
1351 sh		$\nu_{44}$ (a <sub>1</sub> )	1379 (0.38) [53.60]	1356 (0.16)
1369 vw				
1386 vw	1384 sh			
1399 w				
1403 sh	1401 s	$\nu_{45}$ (a <sub>1</sub> )	1455 (10.53) [19.96]	1415 (9.69)
1412 sh				
1429 vw				
1450 vw	1447 vs	$\nu_{46}$ (a <sub>1</sub> )	1494 (13.17) [5.44]	1465 (16.40)
1479 m	1477 m	$\nu_{47}$ (b <sub>2</sub> )	1498 (21.02) [0.76]	1471 (17.20)
1508 vw		$\nu_{48}$ (b <sub>2</sub> )	1528 (7.99) [7.37]	1499 (2.34)
		$\nu_{49}$ (a <sub>1</sub> )	1533 (0.38) [112.85]	1505 (0.21)
	1525 w			
1577 m	1573 m	$\nu_{50}$ (a <sub>1</sub> )	1653 (0.81) [39.68]	1600 (0.25)
	1597 m	$\nu_{51}$ (b <sub>2</sub> )	1660 (1.70) [0.17]	1621 (1.56)
1612 s		$\nu_{53}$ (a <sub>1</sub> )	1690 (0.01) [412.97]	1650 (0.02)
1625 w				
	1645 w	$\nu_{52}$ (b <sub>2</sub> )	1686 (2.23) [11.08]	1656 (1.10)

<sup>a</sup> (km/mole) [Å<sup>4</sup>/AMU]    <sup>b</sup> Ref [223] Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder.    <sup>c</sup> Unscaled DFT frequencies calculated using  $\omega$ B97X-D/6-311++G\*\*.    <sup>d</sup> Anharmonics  $\omega$ B97X-D/6-311++G\*\* calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.25 – Low wavenumber Raman ad PA infrared spectra of Cyclopentaphenanthrene ( $C_2V$ ).

Predicted <sup>a</sup>			Predicted <sup>a</sup>		
Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>
$\nu_1$ (b <sub>1</sub> )	94 (1.66) [0.01]	93 (1.67)	$\nu_{10}$ (b <sub>1</sub> )	495 (1.02) [3.09]	491 (1.00)
$\nu_2$ (a <sub>2</sub> )	186 (0.00) [0.04]	183 (0.00)	$\nu_{11}$ (b <sub>2</sub> )	508 (3.20) [0.03]	498 (3.32)
$\nu_3$ (b <sub>1</sub> )	202 (10.09) [0.01]	199 (9.89)	$\nu_{12}$ (a <sub>2</sub> )	534 (0.33) [0.72]	523 (0.29)
$\nu_4$ (a <sub>2</sub> )	258 (1.45) [0.83]	255 (1.29)	$\nu_{13}$ (b <sub>2</sub> )	560 (0.08) [2.88]	550 (0.12)
$\nu_5$ (b <sub>1</sub> )	288 (0.00) [1.38]	290 (0.00)	$\nu_{14}$ (b <sub>1</sub> )	578 (0.00) [0.05]	573 (0.00)
$\nu_6$ (a <sub>1</sub> )	346 (1.68) [0.38]	336 (1.66)	$\nu_{15}$ (a <sub>1</sub> )	604 (1.99) [21.65]	596 (1.24)
$\nu_7$ (a <sub>2</sub> )	435 (0.00) [2.59]	427 (0.00)	$\nu_{16}$ (a <sub>2</sub> )	631 (0.00) [0.02]	627 (0.00)
$\nu_8$ (a <sub>1</sub> )	446 (0.02) [22.93]	441 (0.03)	$\nu_{17}$ (a <sub>1</sub> )	671 (1.10) [27.51]	663 (1.12)
$\nu_9$ (b <sub>2</sub> )	480 (0.40) [0.18]	473 (0.21)			

<sup>a</sup> (km/mole) [ $A^4/AMU$ ]      <sup>b</sup> Unscaled DFT frequencies calculated using  $\omega$ B97X-D/6-311++G\*\*.<sup>c</sup> Anharmonics PT2 level using  $\omega$ B97X-D/6-311++G\*\*Table C.26 – Raman ad PA infrared spectra of Cyclopentaphenanthrene ( $C_2V$ ), 700- 2000 cm<sup>-1</sup>.

Predicted <sup>a</sup>			Predicted <sup>a</sup>		
Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>
$\nu_{18}$ (b <sub>2</sub> )	710 (0.01) [0.11]	702 (0.00)	$\nu_{39}$ (a <sub>1</sub> )	1195 (1.69) [0.42]	1177 (0.69)
$\nu_{19}$ (b <sub>1</sub> )	721 (16.00) [0.01]	706 (11.68)	$\nu_{40}$ (a <sub>1</sub> )	1197 (0.15) [0.85]	1182 (0.04)
$\nu_{20}$ (a <sub>2</sub> )	772 (53.93) [0.01]	761 (44.53)	$\nu_{41}$ (b <sub>2</sub> )	1213 (0.03) [10.78]	1196 (0.04)
$\nu_{21}$ (b <sub>1</sub> )	793 (0.00) [0.55]	794 (2.98)	$\nu_{42}$ (b <sub>2</sub> )	1241 (3.64) [37.21]	1220 (0.62)
$\nu_{22}$ (a <sub>2</sub> )	793 (8.19) [0.01]	779 (7.48)	$\nu_{43}$ (a <sub>1</sub> )	1247 (0.96) [13.48]	1223 (1.02)
$\nu_{23}$ (b <sub>2</sub> )	824 (0.00) [1.02]	857 (0.03)	$\nu_{44}$ (a <sub>1</sub> )	1256 (0.17) [3.67]	1232 (0.19)
$\nu_{24}$ (a <sub>1</sub> )	832 (0.34) [6.06]	821 (0.42)	$\nu_{45}$ (b <sub>2</sub> )	1304 (0.00) [0.17]	1278 (0.01)
$\nu_{25}$ (b <sub>1</sub> )	844 (75.45) [0.7]	827 (69.25)	$\nu_{46}$ (b <sub>2</sub> )	1350 (4.05) [56.79]	1330 (3.54)
$\nu_{26}$ (a <sub>2</sub> )	912 (6.81) [0.23]	896 (6.91)	$\nu_{47}$ (b <sub>2</sub> )	1392 (2.21) [31.99]	1362 (2.02)
$\nu_{27}$ (b <sub>1</sub> )	929 (0.00) [0.84]	913 (0.00)	$\nu_{48}$ (a <sub>1</sub> )	1404 (0.12) [3.56]	1370 (0.01)
$\nu_{28}$ (a <sub>2</sub> )	972 (1.04) [0.02]	951 (1.75)	$\nu_{49}$ (a <sub>1</sub> )	1444 (0.73) [24.02]	1412 (7.30)
$\nu_{29}$ (b <sub>1</sub> )	996 (0.00) [0.82]	981 (0.01)	$\nu_{50}$ (b <sub>2</sub> )	1455 (8.39) [11.73]	1418 (0.35)
$\nu_{30}$ (a <sub>2</sub> )	1001 (2.20) [0.15]	990 (0.80)	$\nu_{51}$ (a <sub>1</sub> )	1471 (11.68) [0.26]	1442 (9.21)
$\nu_{31}$ (b <sub>2</sub> )	1003 (0.04) [0.12]	988 (0.30)	$\nu_{52}$ (b <sub>2</sub> )	1493 (10.57) [3.69]	1462 (8.47)
$\nu_{32}$ (a <sub>1</sub> )	1005 (0.00) [0.93]	990 (0.00)	$\nu_{54}$ (a <sub>1</sub> )	1525 (1.38) [463.34]	1492 (1.00)
$\nu_{33}$ (a <sub>1</sub> )	1030 (0.03) [2.44]	1013 (0.04)	$\nu_{54}$ (b <sub>2</sub> )	1543 (1.55) [0.55]	1507 (0.41)
$\nu_{34}$ (b <sub>2</sub> )	1051 (2.31) [2.52]	1028 (2.34)	$\nu_{55}$ (a <sub>1</sub> )	1573 (1.79) [34.07]	1537 (1.10)
$\nu_{35}$ (b <sub>2</sub> )	1060 (0.11) [45.51]	1040 (0.10)	$\nu_{56}$ (b <sub>2</sub> )	1668 (4.82) [2.85]	1629 (3.32)
$\nu_{36}$ (a <sub>1</sub> )	1087 (8.72) [0.65]	1063 (7.07)	$\nu_{57}$ (b <sub>2</sub> )	1670 (0.11) [11.69]	1636 (0.05)
$\nu_{37}$ (b <sub>2</sub> )	1169 (0.68) [3.31]	1151 (0.51)	$\nu_{58}$ (a <sub>1</sub> )	1704 (0.01) [105.82]	1698 (0.32)
$\nu_{38}$ (a <sub>1</sub> )	1171 (0.00) [1.76]	1151 (0.02)	$\nu_{59}$ (b <sub>2</sub> )	1717 (0.27) [47.34]	1678 (0.08)

<sup>a</sup> (km/mole) [ $A^4/AMU$ ]      <sup>b</sup> Unscaled DFT frequencies calculated using  $\omega$ B97X-D/6-311++G\*\*.<sup>c</sup> Anharmonics PT2 level using  $\omega$ B97X-D/6-311++G\*\*

Table C.27 – Low wavenumber Raman ad PA infrared spectra of 1-methylfluorene (Cs).

Predicted <sup>a</sup>			Predicted <sup>a</sup>		
Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>
$\nu_1$ (A'')	89 (0.76) [0.20]	90 (0.07)	$\nu_{10}$ (A')	439 (0.47) [4.10]	431 (0.47)
$\nu_2$ (A'')	98 (0.10) [0.06]	111 (2.43)	$\nu_{11}$ (A'')	482 (0.65) [0.75]	477 (0.43)
$\nu_3$ (A'')	119 (0.15) [0.59]	114 (0.16)	$\nu_{12}$ (A')	508 (0.07) [7.13]	501 (0.06)
$\nu_4$ (A')	186 (0.15) [0.41]	182 (0.21)	$\nu_{13}$ (A'')	514 (0.66) [0.45]	509 (0.58)
$\nu_5$ (A'')	217 (0.05) [0.34]	208 (0.02)	$\nu_{14}$ (A')	550 (1.87) [3.23]	544 (1.86)
$\nu_6$ (A'')	253 (9.91) [0.94]	248 (10.02)	$\nu_{15}$ (A')	589 (2.24) [5.60]	581 (2.21)
$\nu_7$ (A'')	278 (0.06) [4.24]	274 (0.04)	$\nu_{16}$ (A'')	593 (1.28) [0.34]	586 (1.14)
$\nu_8$ (A')	304 (0.04) [2.63]	298 (0.13)	$\nu_{17}$ (A')	645 (3.26) [0.20]	636 (3.02)
$\nu_9$ (A'')	432 (3.35) [0.24]	421 (4.15)			

<sup>a</sup> (km/mole) [A<sup>4</sup>/AMU]    <sup>b</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*.<sup>c</sup> Anharmonics PT2 level using ωB97X-D/6-311++G\*\*Table C.28 – Raman ad PA infrared spectra of 1-methylfluorene (Cs), 700 - 2000 cm<sup>-1</sup>

Predicted <sup>a</sup>			Predicted <sup>a</sup>		
Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>	Mode (Symmetry)	Harmonic <sup>b</sup>	Anharmonic <sup>c</sup>
$\nu_{18}$ (A')	717 (1.54) [23.31]	707 (1.36)	$\nu_{40}$ (A')	1195 (0.66) [3.65]	1183 (1.45)
$\nu_{19}$ (A'')	718 (7.22) [0.05]	707 (8.61)	$\nu_{41}$ (A')	1203 (2.06) [1.55]	1184 (2.35)
$\nu_{20}$ (A'')	749 (5.64) [0.32]	739 (10.95)	$\nu_{42}$ (A')	1226 (2.96) [15.34]	1204 (1.26)
$\nu_{21}$ (A'')	775 (106.31) [0.25]	764 (89.80)	$\nu_{43}$ (A')	1244 (2.85) [21.79]	1214 (1.62)
$\nu_{22}$ (A'')	815 (6.47) [1.47]	810 (6.49)	$\nu_{44}$ (A')	1281 (1.09) [38.17]	1257 (1.18)
$\nu_{23}$ (A')	821 (0.41) [1.99]	805 (0.42)	$\nu_{45}$ (A')	1317 (3.21) [79.21]	1290 (1.67)
$\nu_{24}$ (A')	854 (0.33) [23.71]	843 (0.15)	$\nu_{46}$ (A')	1333 (0.78) [113.24]	1309 (0.36)
$\nu_{25}$ (A'')	888 (0.36) [0.08]	865 (0.26)	$\nu_{47}$ (A')	1355 (5.24) [2.14]	1320 (3.76)
$\nu_{26}$ (A')	907 (1.60) [11.85]	892 (0.53)	$\nu_{48}$ (A')	1380 (0.38) [54.68]	1356 (0.25)
$\nu_{27}$ (A'')	923 (0.11) [0.12]	912 (0.00)	$\nu_{49}$ (A')	1421 (2.78) [10.58]	1391 (3.16)
$\nu_{28}$ (A'')	944 (0.46) [0.11]	926 (0.41)	$\nu_{50}$ (A')	1453 (9.46) [18.96]	1405 (6.71)
$\nu_{29}$ (A'')	979 (1.98) [0.03]	958 (1.42)	$\nu_{51}$ (A')	1468 (2.57) [18.44]	1438 (2.41)
$\nu_{30}$ (A'')	999 (0.82) [0.15]	990 (0.40)	$\nu_{52}$ (A'')	1489 (10.27) [9.27]	1442 (5.39)
$\nu_{31}$ (A'')	1007 (0.01) [0.18]	997 (0.05)	$\nu_{53}$ (A')	1501 (9.22) [4.23]	1468 (19.27)
$\nu_{32}$ (A')	1015 (1.66) [14.79]	993 (1.00)	$\nu_{54}$ (A')	1506 (34.50) [11.70]	1474 (24.87)
$\nu_{33}$ (A')	1057 (6.75) [32.6]	1038 (0.76)	$\nu_{55}$ (A')	1524 (1.89) [47.75]	1492 (1.00)
$\nu_{34}$ (A'')	1063 (1.38) [0.09]	1043 (0.98)	$\nu_{56}$ (A')	1546 (1.73) [35.66]	1511 (1.60)
$\nu_{35}$ (A')	1096 (5.01) [2.34]	1083 (1.45)	$\nu_{57}$ (A')	1656 (1.63) [31.68]	1618 (0.45)
$\nu_{36}$ (A')	1112 (1.74) [21.12]	1096 (2.87)	$\nu_{58}$ (A')	1670 (14.50) [46.92]	1631 (6.70)
$\nu_{37}$ (A')	1145 (1.89) [0.80]	1122 (1.66)	$\nu_{59}$ (A')	1680 (3.08) [83.45]	1640 (2.89)
$\nu_{38}$ (A'')	1171 (0.40) [3.19]	1139 (0.52)	$\nu_{60}$ (A')	1689 (0.43) [300.56]	1653 (0.29)
$\nu_{39}$ (A')	1184 (0.64) [11.35]	1168 (0.95)			

<sup>a</sup> (km/mole) [A<sup>4</sup>/AMU]    <sup>b</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*.<sup>c</sup> Anharmonics PT2 level using ωB97X-D/6-311++G\*\*

Table C.29 – Low wavenumber Raman ad PA infrared spectra of 1,8-dimethylfluorene ( $C_2V$ ).

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
121 m	138 m	$\nu_1$ (a <sub>2</sub> )	84 (0.00) [0.17]	82 (0.00)
		$\nu_2$ (b <sub>1</sub> )	88 (0.41) [0.27]	90 (0.20)
		$\nu_3$ (b <sub>1</sub> )	100 (1.02) [0.09]	108 (0.82)
		$\nu_4$ (a <sub>2</sub> )	103 (0.00) [0.40]	105 (0.00)
		$\nu_5$ (a <sub>1</sub> )	160 (0.11) [0.11]	145 (0.10)
	166 w			
	219 m	$\nu_6$ (b <sub>1</sub> )	201 (0.06) [0.85]	204 (0.11)
	240m	$\nu_7$ (a <sub>2</sub> )	233 (0.00) [0.43]	234 (0.00)
	264 vw	$\nu_8$ (b <sub>1</sub> )	260 (12.31) [1.03]	263 (10.04)
	285 m	$\nu_9$ (a <sub>2</sub> )	278 (0.00) [4.24]	280 (0.00)
313 m	356 vw	$\nu_{10}$ (b <sub>2</sub> )	282 (0.15) [1.88]	284 (0.20)
		$\nu_{11}$ (a <sub>1</sub> )	315 (0.00) [3.21]	314 (0.00)
	387 vw			
	403 vw			
	430 vw			
	451 vw			
	464 w	$\nu_{12}$ (a <sub>1</sub> )	473 (0.40) [2.06]	468 (0.34)
		$\nu_{13}$ (b <sub>1</sub> )	484 (0.48) [0.14]	472 (0.21)
	478 w	$\nu_{14}$ (b <sub>2</sub> )	490 (0.15) [4.42]	483 (0.11)
500 m	538 m	$\nu_{15}$ (a <sub>2</sub> )	514 (0.00) [0.51]	505 (0.00)
		$\nu_{16}$ (b <sub>1</sub> )	515 (1.14) [0.31]	509 (0.95)
		$\nu_{17}$ (a <sub>1</sub> )	551 (0.54) [12.32]	545 (0.23)
		$\nu_{18}$ (b <sub>1</sub> )	570 (3.52) [3.09]	566 (3.44)
		$\nu_{19}$ (b <sub>2</sub> )	599 (3.98) [2.50]	591 (4.02)
		$\nu_{20}$ (a <sub>2</sub> )	605 (0.00) [0.62]	600 (0.00)
	622 vw			
	645 vw			
	664 w	668 vw		
681 vs	694 vw	$\nu_{21}$ (a <sub>1</sub> )	699 (0.15) [23.45]	693 (0.41)
		$\nu_{22}$ (b <sub>1</sub> )	717 (9.07) [0.11]	705 (10.17)

<sup>a</sup> (km/mole) [ $A^4/AMU$ ]<sup>b</sup> Ref [223], Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder.<sup>c</sup> Unscaled DFT frequencies calculated using  $\omega$ B97X-D/6-311++G\*\*.<sup>d</sup> Anharmonics  $\omega$ B97X-D/6-311++G\*\* calculation using the P\_VMWCI<sub>2</sub> algorithm

Table C.30 – Raman and PA infrared spectra of 1,8-dimethylfluorene ( $C_2V$ ), 700–2000  $\text{cm}^{-1}$ 

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
724 vw		$\nu_{23}$ (a <sub>2</sub> )	749 (0.00) [0.16]	720 (0.00)
749 vw	745 s	$\nu_{25}$ (b <sub>2</sub> )	795 (2.54) [0.003]	777 (1.06)
774 vw	768 vs	$\nu_{24}$ (b <sub>1</sub> )	792 (98.65) [0.29]	780 (70.58)
796 vw		$\nu_{26}$ (a <sub>2</sub> )	821 (0.00) [1.34]	810 (0.00)
	817 vw	$\nu_{27}$ (b <sub>2</sub> )	854 (1.40) [0.76]	840 (0.90)
835 m	835 vw	$\nu_{28}$ (a <sub>1</sub> )	855 (0.01) [23.08]	843 (0.17)
879 vw	885 w	$\nu_{29}$ (b <sub>1</sub> )	921 (0.23) [0.04]	912 (0.20)
	920 w	$\nu_{30}$ (a <sub>2</sub> )	924 (0.00) [0.22]	922 (0.00)
933 m	935 vw	$\nu_{31}$ (b <sub>1</sub> )	945 (1.05) [0.05]	942 (0.47)
	965 vw	$\nu_{32}$ (a <sub>1</sub> )	956 (3.48) [37.32]	958 (2.89)
	996 vw	$\nu_{33}$ (a <sub>2</sub> )	995 (0.00) [0.23]	997 (0.00)
1016 w		$\nu_{34}$ (b <sub>1</sub> )	1000 (0.76) [0.01]	1010 (0.32)
		$\nu_{35}$ (b <sub>2</sub> )	1011 (1.50) [0.28]	1016 (0.74)
	1039 m	$\nu_{36}$ (a <sub>1</sub> )	1046 (0.16) [6.75]	1043 (0.45)
1077 w	1073 m	$\nu_{37}$ (a <sub>2</sub> )	1060 (0.00) [0.03]	1065 (0.00)
		$\nu_{38}$ (b <sub>1</sub> )	1065 (3.19) [0.10]	1070 (3.03)
1113 vw	1111 vw	$\nu_{39}$ (b <sub>2</sub> )	1102 (6.35) [1.32]	1103 (2.96)
		$\nu_{40}$ (a <sub>1</sub> )	1110 (7.70) [27.30]	1113 (3.42)
1160 w	1159 m	$\nu_{41}$ (b <sub>2</sub> )	1140 (0.27) [6.90]	1143 (0.36)
		$\nu_{42}$ (a <sub>2</sub> )	1169 (0.00) [2.78]	1162 (0.00)
	1175 sh	$\nu_{43}$ (a <sub>1</sub> )	1194 (0.07) [0.78]	1190 (0.1)
1194 w	1196 vw	$\nu_{44}$ (b <sub>2</sub> )	1194 (1.48) [6.73]	1202 (1.02)
		$\nu_{45}$ (b <sub>2</sub> )	1216 (1.17) [3.17]	1218 (0.95)
	1224 vw	$\nu_{46}$ (a <sub>1</sub> )	1229 (1.08) [14.89]	1227 (0.74)
1241 w	1244 vw	$\nu_{47}$ (b <sub>2</sub> )	1263 (3.11) [1.21]	1256 (2.96)
1260 sh	1265 vw	$\nu_{48}$ (a <sub>1</sub> )	1280 (0.80) [23.66]	1277 (0.85)
1292 s	1296 sh	$\nu_{49}$ (b <sub>2</sub> )	1303 (1.93) [9.11]	1302 (0.99)
	1312 m	$\nu_{50}$ (a <sub>1</sub> )	1331 (1.80) [221.83]	1320 (1.92)
1344 m	1344 vw	$\nu_{51}$ (b <sub>2</sub> )	1354 (3.56) [1.27]	1350 (1.59)
		$\nu_{52}$ (a <sub>1</sub> )	1379 (0.32) [53.10]	1360 (0.12)
1380 w	1381 m	$\nu_{53}$ (b <sub>2</sub> )	1420 (0.08) [0.95]	1405 (0.20)
	1400 sh	$\nu_{54}$ (a <sub>1</sub> )	1421 (5.22) [19.60]	1408 (3.41)
1418 w	1428 sh	$\nu_{55}$ (a <sub>1</sub> )	1451 (8.68) [21.19]	1440 (5.33)
	1452 s	$\nu_{56}$ (a <sub>1</sub> )	1467 (4.10) [47.00]	1460 (3.62)
		$\nu_{57}$ (b <sub>2</sub> )	1471 (1.37) [3.10]	1466 (1.12)
1484 w	1483 sh	$\nu_{58}$ (a <sub>2</sub> )	1489 (0.00) [9.60]	1486 (0.00)
		$\nu_{59}$ (b <sub>1</sub> )	1489 (20.61) [8.48]	1488 (18.54)
		$\nu_{60}$ (a <sub>1</sub> )	1505 (5.37) [6.73]	1497 (2.87)
		$\nu_{61}$ (b <sub>2</sub> )	1507 (41.17) [0.39]	1500 (29.54)

Observed		Predicted <sup>a</sup>		
Raman <sup>b</sup>	Infrared <sup>b</sup>	Mode (Symmetry)	Harmonic <sup>c</sup>	Anharmonic <sup>d</sup>
1592 m	1602 vs	$\nu_{62}$ (a <sub>1</sub> )	1537 (4.99) [45.30]	1532 (3.23)
		$\nu_{63}$ (b <sub>2</sub> )	1541 (0.08) [7.10]	1533 (0.35)
		$\nu_{64}$ (b <sub>2</sub> )	1669 (17.13) [82.20]	1606 (8.26)
		$\nu_{65}$ (a <sub>1</sub> )	1670 (15.60) [0.03]	1619 (5.30)
		$\nu_{66}$ (b <sub>2</sub> )	1679 (0.05) [10.06]	1615 (0.10)
		$\nu_{67}$ (a <sub>1</sub> )	1681 (0.97) [364.11]	1620 (0.23)

<sup>a</sup> (km/mole) [Å<sup>4</sup>/AMU]

<sup>b</sup> Ref [223], Relative intensities: vw = very weak, w = weak, m = medium, s = strong, vs = very strong, sh = shoulder.    <sup>c</sup> Unscaled DFT frequencies calculated using ωB97X-D/6-311++G\*\*.    <sup>d</sup> Anharmonics ωB97X-D/6-311++G\*\* calculation using the P\_VMWCI<sub>2</sub> algorithm

## C.2 Dimers

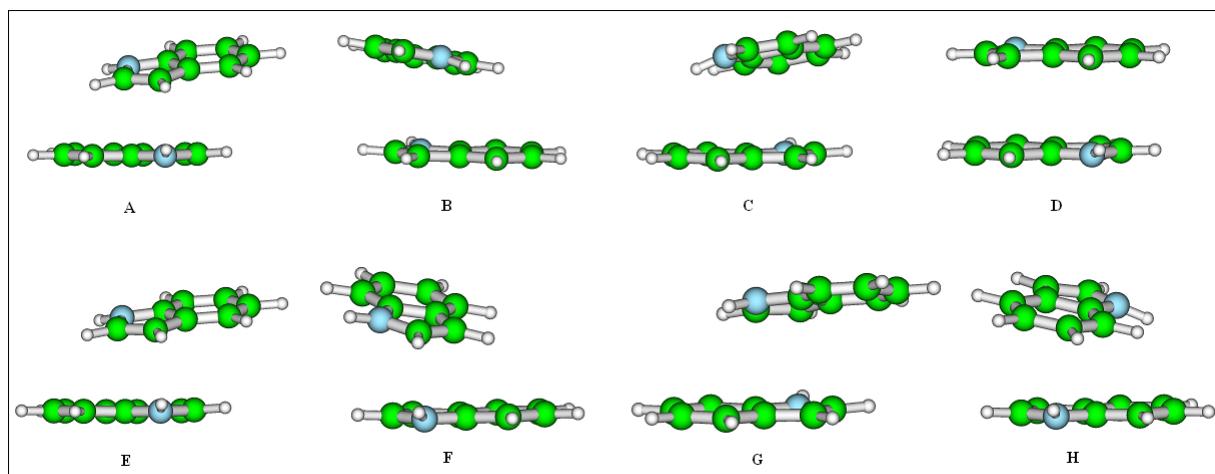


Figure C.1 – Configurations of Indole dimers optimized at the ωB97X-D/6-311++G\*\* level of theory

Table C.31 – Interaction energies of Indole dimers (kcal/mol) in 8 different configurations using ωB97X-D/6-311++G\*\*, and compared with SAPT-DFT.

	A	B	C	D	E	F	G	H
ωB97X-D/ 6-311++G**	-9.22	-9.49	-10.06	-9.34	-9.22	-6.76	-9.22	-10.06
SAPT- DFT(PBE0) aug-cc pvtz	-7.54	-7.75	-8.76	-8.19	-7.54	-5.50	-7.54	-8.76
LMP2/ aug-cc-pVTZ <sup>a</sup>					-7.77			
CAM-B3LYP/ 6-31+G*				-0.91				

<sup>a</sup> Ref [284]

Table C.32 – Intermolecular distance between pyrrole rings center of Indole dimers ( $\text{\AA}$ ), selon figure 4.7

	A=E=G	B	C = H	D	F
$\omega\text{B97x-D/ 6-311++G}^{**}$	3.71	3.55	3.44	4.12	4.01
DFT-SAPT(PBE0)/ aug-cc pvtz	3.72	3.57	3.46	4.13	4.04
CAM-B3LYP/ 6-31+G*			3.57		

Table C.33 – Angle between ring planes of Indole dimers ( $^\circ$ ) at the  $\omega\text{B97x-D/6-311++G}^{**}$  level of theory, according to figure 4.7

	A = E = G	B	C = H	C (CAM-B3LYP/ 6-31+G*)	D	F
$\alpha$	21.80	16.82	18.19	22.91	0.59	15.73
$\beta$	26.05	6.06	15.06	23.94	2.11	22.04
$\gamma$	45.40	76.65	61.23	59.43	0.003	87.25

Table C.34 – Calculated low wavenumber Raman ad PA infrared spectra of Indole Dimer.

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_1$	29	(0.44) [1.84]	$\nu_{13}$	437	(35.80) [0.23]
$\nu_2$	40	(0.57) [7.68]	$\nu_{14}$	444	(0.67) [0.70]
$\nu_3$	47	(0.03) [5.44]	$\nu_{15}$	474	(140.28) [2.59]
$\nu_4$	73	(2.63) [0.58]	$\nu_{16}$	513	(134.92) [2.75]
$\nu_5$	89	(0.09) [2.10]	$\nu_{17}$	555	(0.30) [5.40]
$\nu_6$	129	(0.43) [4.17]	$\nu_{18}$	556	(0.88) [6.59]
$\nu_7$	230	(18.65) [1.38]	$\nu_{19}$	594	(0.33) [0.08]
$\nu_8$	234	(0.37) [1.05]	$\nu_{20}$	596	(0.49) [0.22]
$\nu_9$	259	(1.03) [0.13]	$\nu_{21}$	618	(4.31) [1.08]
$\nu_{10}$	262	(3.19) [0.06]	$\nu_{22}$	623	(8.48) [0.88]
$\nu_{11}$	406	(1.38) [0.11]	$\nu_{23}$	626	(4.37) [8.28]
$\nu_{12}$	408	(5.90) [0.43]	$\nu_{24}$	627	(19.85) [1.87]

<sup>a</sup> km/mole

<sup>b</sup>  $\text{A}^4/\text{AMU}$

Table C.35 – Calculated Raman and PA infrared spectra of Indole Dimer, 700–2000 cm<sup>-1</sup>

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{25}$	734	(63.11) [0.08]	$\nu_{51}$	1150	(0.25) [0.22]
$\nu_{26}$	743	(0.57) [1.09]	$\nu_{52}$	1152	(1.13) [9.76]
$\nu_{27}$	769	(231.13) [0.35]	$\nu_{53}$	1179	(0.75) [1.87]
$\nu_{28}$	771	(4.75) [0.66]	$\nu_{54}$	1179	(0.82) [0.40]
$\nu_{29}$	784	(9.54) [0.99]	$\nu_{55}$	1237	(4.46) [2.12]
$\nu_{30}$	785	(0.65) [40.14]	$\nu_{56}$	1237	(4.58) [1.47]
$\nu_{31}$	794	(65.87) [0.48]	$\nu_{57}$	1278	(10.76) [8.48]
$\nu_{32}$	797	(0.36) [0.79]	$\nu_{58}$	1279	(3.50) [3.26]
$\nu_{33}$	872	(0.75) [0.22]	$\nu_{59}$	1318	(2.92) [25.37]
$\nu_{34}$	877	(1.59) [0.16]	$\nu_{60}$	1321	(16.15) [7.90]
$\nu_{35}$	893	(0.08) [0.58]	$\nu_{61}$	1381	(25.17) [1.48]
$\nu_{36}$	894	(0.22) [2.32]	$\nu_{62}$	1382	(30.25) [38.1]
$\nu_{37}$	898	(1.63) [1.53]	$\nu_{63}$	1391	(6.47) [13.89]
$\nu_{38}$	899	(0.73) [2.74]	$\nu_{64}$	1391	(4.12) [5.82]
$\nu_{39}$	921	(2.94) [9.77]	$\nu_{65}$	1469	(19.27) [21.23]
$\nu_{40}$	921	(7.25) [0.57]	$\nu_{66}$	1506	(4.04) [3.07]
$\nu_{41}$	960	(5.44) [0.22]	$\nu_{67}$	1508	(55.53) [15.45]
$\nu_{42}$	964	(1.32) [0.45]	$\nu_{68}$	1540	(8.42) [9.06]
$\nu_{43}$	995	(0.20) [0.03]	$\nu_{69}$	1543	(0.06) [3.52]
$\nu_{44}$	996	(0.35) [0.24]	$\nu_{70}$	1576	(10.62) [41.44]
$\nu_{45}$	1042	(4.32) [45.88]	$\nu_{71}$	1578	(6.87) [103.97]
$\nu_{46}$	1043	(9.13) [1.20]	$\nu_{72}$	1649	(0.09) [7.24]
$\nu_{47}$	1097	(2.00) [21.10]	$\nu_{73}$	1650	(0.58) [7.24]
$\nu_{48}$	1097	(2.57) [3.72]	$\nu_{74}$	1691	(0.08) [7.09]
$\nu_{49}$	1128	(12.77) [5.03]	$\nu_{75}$	1692	(3.47) [21.19]

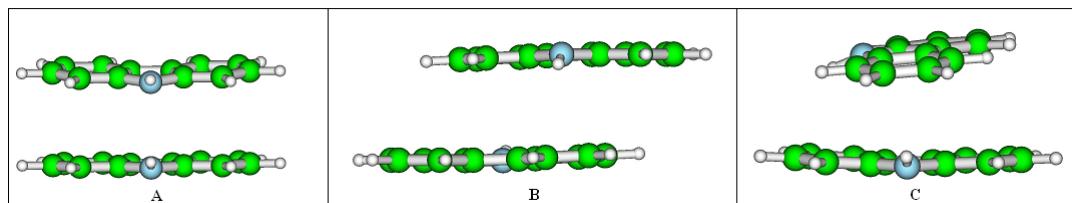
<sup>a</sup> km/mole    <sup>b</sup> Å<sup>4</sup>/AMU

Figure C.2 – Configurations of Carbazole dimers optimized at the ωB97x-D/6-311++G\*\* level of theory.

Table C.36 – Interaction energies of Carbazole dimers (kcal/mol) in 3 different configurations using  $\omega$ B97x-D/6-311++G\*\*.

	<b>A</b>	<b>B</b>	<b>C</b>
$\omega$ B97x-D/ 6-311++G**	-10.09	-12.08	-14.57

Table C.37 – Intermolecular distance between pyrrole ring center ( $\text{\AA}$ ) and angles ( $^\circ$ ) according figure of Carbazole dimers using  $\omega$ B97x-D/6-311++G\*\*.

	<b>A</b>	<b>B</b>	<b>C</b>
<b>d</b>	3.68	4.02	3.30
$\alpha$	5.01	1.59	4.30
$\beta$	12.60	4.17	6.01
$\gamma$			76.83

Table C.38 – Calculated low wavenumber Raman ad PA infrared spectra of Carbazole Dimer.

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_1$	26	(0.05) [4.28]	$\nu_{19}$	437	(77.32) [9.71]
$\nu_2$	31	(0.03) [8.66]	$\nu_{20}$	438	(11.67) [2.22]
$\nu_3$	43	(0.02) [8.30]	$\nu_{21}$	446	(203.96) [0.35]
$\nu_4$	65	(1.32) [0.40]	$\nu_{22}$	447	(30.19) [0.63]
$\nu_5$	70	(0.008) [1.17]	$\nu_{23}$	457	(13.23) [0.08]
$\nu_6$	95	(0.51) [2.21]	$\nu_{24}$	458	(0.01) [0.37]
$\nu_7$	115	(2.85) [0.89]	$\nu_{25}$	520	(1.52) [0.18]
$\nu_8$	153	(0.12) [0.08]	$\nu_{26}$	521	(3.61) [1.57]
$\nu_9$	158	(0.11) [0.54]	$\nu_{27}$	568	(0.37) [0.68]
$\nu_{10}$	168	(0.04) [0.85]	$\nu_{28}$	568	(0.61) [12.89]
$\nu_{11}$	221	(0.47) [1.28]	$\nu_{29}$	584	(0.02) [0.31]
$\nu_{12}$	221	(0.35) [1.46]	$\nu_{30}$	585	(22.52) [0.20]
$\nu_{13}$	301	(0.74) [4.64]	$\nu_{31}$	594	(1.87) [0.07]
$\nu_{14}$	305	0.005 [3.61]	$\nu_{32}$	596	(0.01) [0.03]
$\nu_{15}$	321	(0.01) [0.05]	$\nu_{33}$	635	(2.63) [0.01]
$\nu_{16}$	325	(2.83) [0.06]	$\nu_{34}$	635	(4.06) [0.02]
$\nu_{17}$	431	(7.50) [0.09]	$\nu_{35}$	678	(4.03) [0.48]
$\nu_{18}$	435	(5.63) [17.54]	$\nu_{36}$	679	(1.77) [1.81]

<sup>a</sup> km/mole<sup>b</sup>  $\text{A}^4/\text{AMU}$

Table C.39 – Calculated Raman and PA infrared spectra of Carbazole Dimer, 700–2000 cm<sup>-1</sup>

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{37}$	750	(241.17) [0.24]	$\nu_{73}$	1179	(6.29) [0.40]
$\nu_{38}$	754	(0.02) [0.26]	$\nu_{74}$	1179	(3.39) [7.54]
$\nu_{39}$	762	(4.39) [1.11]	$\nu_{75}$	1187	(0.41) [3.90]
$\nu_{40}$	765	(4.98) [2.93]	$\nu_{76}$	1187	(1.02) [0.43]
$\nu_{41}$	767	(15.18) [0.45]	$\nu_{77}$	1244	(9.48) [17.02]
$\nu_{42}$	767	(0.32) [44.78]	$\nu_{78}$	1245	(2.29) [31.39]
$\nu_{43}$	781	(0.07) [1.06]	$\nu_{79}$	1254	(0.38) [3.53]
$\nu_{44}$	782	(108.04) [0.10]	$\nu_{80}$	1254	(1.17) [1.38]
$\nu_{45}$	796	(1.02) [0.73]	$\nu_{81}$	1277	(44.62) [0.76]
$\nu_{46}$	802	(0.62) [1.65]	$\nu_{82}$	1278	(73.55) [13.45]
$\nu_{47}$	869	(3.99) [0.41]	$\nu_{83}$	1328	(0.17) [89.36]
$\nu_{48}$	870	(3.84) [0.19]	$\nu_{84}$	1329	(2.37) [30.66]
$\nu_{49}$	873	(1.59) [0.04]	$\nu_{85}$	1347	(1.42) [44.81]
$\nu_{50}$	873	(1.99) [1.14]	$\nu_{86}$	1348	(0.02) [169.47]
$\nu_{51}$	877	(4.42) [0.36]	$\nu_{87}$	1365	(4.44) [0.08]
$\nu_{52}$	881	(2.57) [0.67]	$\nu_{88}$	1365	(57.98) [2.21]
$\nu_{53}$	902	(2.28) [3.33]	$\nu_{89}$	1375	(9.18) [23.33]
$\nu_{54}$	902	(0.009) [30.50]	$\nu_{90}$	1377	(11.37) [7.82]
$\nu_{55}$	958	(0.14) [0.30]	$\nu_{91}$	1438	(9.27) [5.06]
$\nu_{56}$	958	(6.99) [0.08]	$\nu_{92}$	1439	(5.14) [0.35]
$\nu_{57}$	962	(0.83) [0.08]	$\nu_{93}$	1501	(12.98) [10.05]
$\nu_{58}$	963	(1.56) [0.28]	$\nu_{94}$	1503	(31.33) [30.01]
$\nu_{59}$	994	(0.03) [0.15]	$\nu_{95}$	1509	(34.71) [0.57]
$\nu_{60}$	996	(1.08) [0.01]	$\nu_{96}$	1509	(43.83) [0.85]
$\nu_{61}$	996	(0.59) [0.26]	$\nu_{97}$	1538	(0.64) [27.23]
$\nu_{62}$	997	(0.00) [0.09]	$\nu_{98}$	1539	(0.95) [39.33]
$\nu_{63}$	1028	(3.64) [0.93]	$\nu_{99}$	1547	(33.20) [2.87]
$\nu_{64}$	1028	(5.18) [2.21]	$\nu_{100}$	1549	(43.02) [4.01]
$\nu_{65}$	1048	(0.00) [88.19]	$\nu_{101}$	1649	(1.39) [57.62]
$\nu_{66}$	1049	(2.21) [5.40]	$\nu_{102}$	1649	(0.02) [19.29]
$\nu_{67}$	1053	(4.46) [0.42]	$\nu_{103}$	1657	(2.86) [2.36]
$\nu_{68}$	1054	(7.89) [3.51]	$\nu_{104}$	1658	(5.23) [20.96]
$\nu_{69}$	1140	(5.75) [1.19]	$\nu_{105}$	1681	(29.56) [17.35]
$\nu_{70}$	1140	(3.15) [17.48]	$\nu_{106}$	1682	(26.68) [0.31]
$\nu_{71}$	1151	(2.91) [0.10]	$\nu_{107}$	1707	(9.47) [121.78]
$\nu_{72}$	1151	(5.66) [2.23]	$\nu_{108}$	1707	(1.58) [124.42]

<sup>a</sup> km/mole<sup>b</sup> Å<sup>4</sup>/AMU

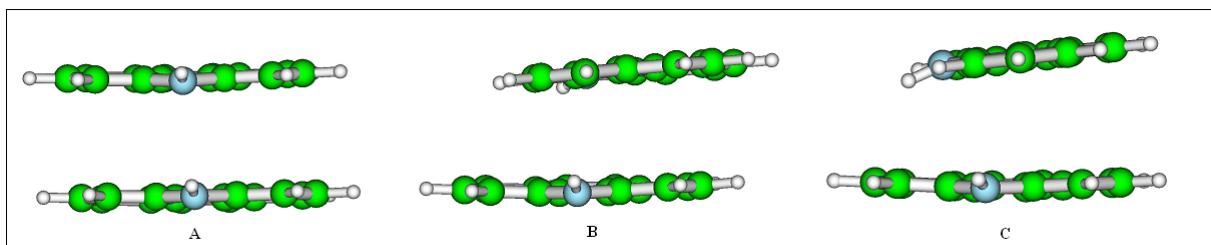


Figure C.3 – Configurations of 4,5-iminophenanthrene dimers optimized at the  $\omega$ B97X-D/6-311++G\*\* level of theory.

Table C.40 – Interaction energies of 4,5-iminophenanthrene dimers (kcal/mol) using  $\omega$ B97X-D/6-311++G\*\*.

	<b>A</b>	<b>B</b>	<b>C</b>
$\omega$ <b>B97x-D/ 6-311++G**</b>	-12.78	-15.13	-14.78

Table C.41 – Intermolecular distance between pyrrole ring center ( $\text{\AA}$ ) and angles ( $^\circ$ ) according figure of 4,5-iminophenanthrene dimers using  $\omega$ B97x-D/6-311++G\*\*.

	<b>A</b>	<b>B</b>	<b>C</b>
<b>d</b>	3.66	3.59	3.66
$\alpha$	2.60	10.25	5.76
$\beta$	4.84	8.42	8.65
$\gamma$			73.64

Table C.42 – Calculated low wavenumber Raman ad PA infrared spectra of 4,5-iminophenanthrene Dimer.

<b>Mode</b>	$\omega \text{ cm}^{-1}$	<b>Intensities</b> ( <sup>a</sup> ) [ <sup>b</sup> ]	<b>Mode</b>	$\omega \text{ cm}^{-1}$	<b>Intensities</b> ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_1$	19	(0.05) [1.55]	$\nu_9$	202	(0.71) [0.62]
$\nu_2$	31	(0.86) [10.03]	$\nu_{10}$	205	(0.10) [1.07]
$\nu_3$	38	(0.06) [8.52]	$\nu_{11}$	220	(0.19) [0.07]
$\nu_4$	61	(0.60) [0.89]	$\nu_{12}$	224	(17.06) [0.10]
$\nu_5$	75	(1.88) (0.84)	$\nu_{13}$	313	(0.15) [0.93]
$\nu_6$	92	(1.31) [5.95]	$\nu_{14}$	314	(0.22) [0.06]
$\nu_7$	114	(6.79) [0.33]	$\nu_{15}$	322	(2.12) [2.57]
$\nu_8$	128	(0.52) [3.59]	$\nu_{16}$	326	(0.97) [1.26]

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{17}$	352	(3.60) [0.45]	$\nu_{30}$	550	(0.07) [1.54]
$\nu_{18}$	353	(3.52) [0.28]	$\nu_{31}$	581	(0.02) [1.66]
$\nu_{19}$	376	(192.54) [11.49]	$\nu_{32}$	581	(0.03) [1.82]
$\nu_{20}$	416	(144.55) [7.09]	$\nu_{33}$	586	(0.03) [0.26]
$\nu_{21}$	446	(0.18) [2.38]	$\nu_{34}$	586	(0.02) [0.17]
$\nu_{22}$	447	(0.12) [1.83]	$\nu_{35}$	613	(1.60) [0.18]
$\nu_{23}$	458	(0.52) [20.91]	$\nu_{36}$	616	(0.48) [0.49]
$\nu_{24}$	458	(5.85) [10.36]	$\nu_{37}$	621	(6.47) [5.71]
$\nu_{25}$	492	(2.23) [2.10]	$\nu_{38}$	622	(1.86) [44.52]
$\nu_{26}$	492	(2.63) [2.11]	$\nu_{39}$	664	(0.69) [0.15]
$\nu_{27}$	508	(5.40) [0.10]	$\nu_{40}$	666	(0.17) [0.11]
$\nu_{28}$	509	(0.38) [0.53]	$\nu_{41}$	698	(2.18) [8.38]
$\nu_{29}$	549	(0.10) [1.74]	$\nu_{42}$	699	(1.51) [22.84]

<sup>a</sup> km/mole    <sup>b</sup>  $\text{Å}^4/\text{AMU}$

Table C.43 – Calculated Raman and PA infrared spectra of 4,5-iminophenanthrene Dimer, 700–2000  $\text{cm}^{-1}$

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{43}$	711	(0.27) [0.70]	$\nu_{63}$	983	(0.70) [0.20]
$\nu_{44}$	711	(0.04) [0.28]	$\nu_{64}$	985	(0.12) [0.11] [10.15]
$\nu_{45}$	739	(128.75) [0.53]	$\nu_{65}$	986	(0.88) [0.09]
$\nu_{46}$	740	(8.29) [0.91]	$\nu_{66}$	987	(2.61) [0.07]
$\nu_{47}$	772	(28.65) [0.63]	$\nu_{67}$	1001	(0.11) [0.79]
$\nu_{48}$	775	(28.81) [0.23]	$\nu_{68}$	1001	(0.45) [2.85]
$\nu_{49}$	776	(3.79) [1.89]	$\nu_{69}$	1030	(0.26) [0.81]
$\nu_{50}$	778	(8.60) [0.79]	$\nu_{70}$	1030	(0.45) [1.22]
$\nu_{51}$	822	(0.06) [0.43]	$\nu_{71}$	1050	(0.05) [5.17]
$\nu_{52}$	823	(0.01) [0.68]	$\nu_{72}$	1051	(0.59) [0.12]
$\nu_{53}$	831	(1.96) [1.33]	$\nu_{73}$	1060	(0.83) [35.07]
$\nu_{54}$	834	(1.06) [2.16]	$\nu_{74}$	1060	(0.56) [24.98]
$\nu_{55}$	843	(43.96) [2.87]	$\nu_{75}$	1076	(0.27) [2.12]
$\nu_{56}$	843	(8.20) [7.55]	$\nu_{76}$	1077	(0.64) [2.08]
$\nu_{57}$	846	(137.09) [0.30]	$\nu_{77}$	1104	(15.85) [5.11]
$\nu_{58}$	847	(2.55) [1.88]	$\nu_{78}$	1105	(57.86) [6.45]
$\nu_{59}$	894	(1.29) [0.51]	$\nu_{79}$	1165	(0.75) [2.80]
$\nu_{60}$	896	(1.13) [0.46]	$\nu_{80}$	1165	(0.90) [6.46]
$\nu_{61}$	903	(6.39) [0.17]	$\nu_{81}$	1202	(0.60) [0.40]
$\nu_{62}$	904	(1.81) [0.36]	$\nu_{82}$	1204	(3.46) [1.00]

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{83}$	1214	(5.39) [10.15]	$\nu_{102}$	1433	(28.00) [0.77]
$\nu_{84}$	1214	(7.53) [7.26]	$\nu_{103}$	1434	(46.35) [0.22]
$\nu_{85}$	1246	(6.75) [15.67]	$\nu_{104}$	1447	(4.55) [36.99]
$\nu_{86}$	1246	(0.32) [64.88]	$\nu_{105}$	1491	(8.97) [45.65]
$\nu_{87}$	1250	(0.52) [0.41]	$\nu_{106}$	1492	(26.35) [11.69]
$\nu_{88}$	1251	(0.15) [2.34]	$\nu_{107}$	1532	(1.09) [327.53]
$\nu_{89}$	1264	(4.78) [2.77]	$\nu_{108}$	1533	(0.87) [141.8]
$\nu_{90}$	1265	(11.08) [2.30]	$\nu_{109}$	1555	(0.62) [0.79]
$\nu_{91}$	1283	(1.07) [31.80]	$\nu_{110}$	1556	(1.31) [0.22]
$\nu_{92}$	1284	(2.77) [8.74]	$\nu_{111}$	1567	(6.88) [8.72]
$\nu_{93}$	1363	(7.13) [3.08]	$\nu_{112}$	1568	(27.68) [3.99]
$\nu_{94}$	1364	(52.26) [1.89]	$\nu_{113}$	1670	(37.28) [0.89]
$\nu_{95}$	1366	(2.83) [73.43]	$\nu_{114}$	1672	(110.36) [2.52]
$\nu_{96}$	1367	(29.20) [34.08]	$\nu_{115}$	1674	(0.85) [11.45]
$\nu_{97}$	1395	(0.34) [14.67]	$\nu_{116}$	1675	(14.04) [13.97]
$\nu_{98}$	1395	(0.29) [16.55]	$\nu_{117}$	1708	(3.04) [66.33]
$\nu_{99}$	1367	(29.20) [34.08]	$\nu_{118}$	1709	(7.00) [81.57]
$\nu_{100}$	1395	(0.34) [14.67]	$\nu_{119}$	1738	(0.23) [4.52]
$\nu_{101}$	1395	(0.29) [16.55]	$\nu_{120}$	1740	(0.60) [5.40]

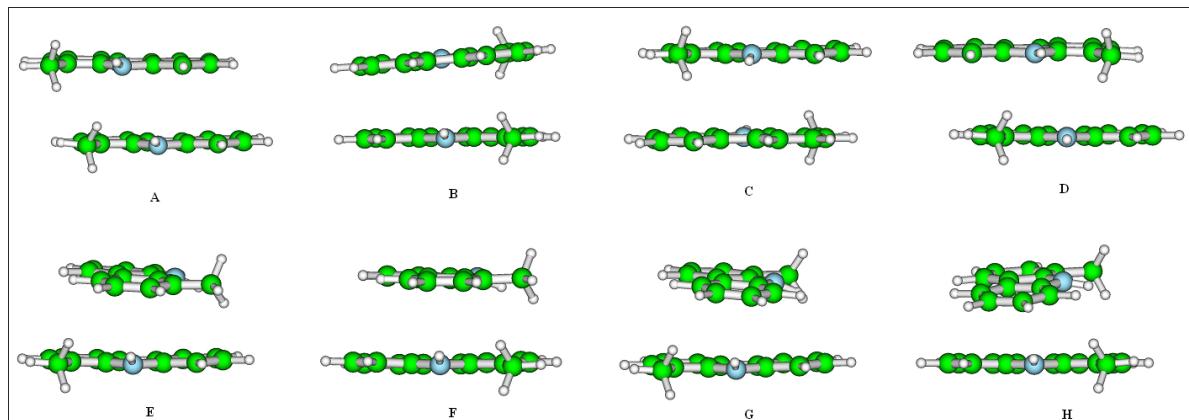
<sup>a</sup> km/mole<sup>b</sup> Å<sup>4</sup>/AMU

Figure C.4 – Configurations of 1-methylcarbazole dimers optimized at the  $\omega$ B97X-D/6-311++G\*\* level of theory.

Table C.44 – Interaction energies of 1-methylcarbazole dimers (kcal/mol) using  $\omega$ B97X-D/6-311++G\*\*.

	A	B	C	D	E	F	G	H
$\omega$ B97X-D/ 6-311++G**	-11.81	-15.04	-15.03	-12.27	-15.62	-16.10	-15.91	-14.81

Table C.45 – Intermolecular distance between pyrrole ring center (Å) and plane angles (°) of 1-methylcarbazole dimers with the  $\omega$ B97x-D/6-311++G\*\*.

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>
<b>d</b>	3.83	3.78	3.76	3.86	3.35	3.36	3.34	3.46
$\alpha$	6.21	5.65	3.65	1.58	5.79	1.62	18.04	8.37
$\beta$	6.28	4.10	5.94	5.98	10.34	10.32	6.35	16.92
$\gamma$					136.11	139.65	135.39	121.90

Table C.46 – Calculated low wavenumber Raman ad PA infrared spectra of 1-methylcarbazole Dimer.

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> )	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> )
$\nu_1$	28	(0.13)	$\nu_{22}$	332	(3.67)
$\nu_2$	33	(0.02)	$\nu_{23}$	440	(9.09)
$\nu_3$	51	(0.00)	$\nu_{24}$	442	(0.40)
$\nu_4$	74	(0.23)	$\nu_{25}$	448	(0.39)
$\nu_5$	76	(0.01)	$\nu_{26}$	449	(1.55)
$\nu_6$	103	(3.40)	$\nu_{27}$	510	(42.79)
$\nu_7$	108	(0.00)	$\nu_{28}$	513	(3.08)
$\nu_8$	125	(0.08)	$\nu_{29}$	522	(23.58)
$\nu_9$	136	(0.21)	$\nu_{30}$	523	(11.52)
$\nu_{10}$	144	(0.27)	$\nu_{31}$	532	(30.72)
$\nu_{11}$	149	(0.09)	$\nu_{32}$	535	(156.32)
$\nu_{12}$	160	(1.74)	$\nu_{33}$	571	(30.52)
$\nu_{13}$	190	(0.49)	$\nu_{34}$	572	(0.27)
$\nu_{14}$	191	(0.19)	$\nu_{35}$	591	(21.30)
$\nu_{15}$	234	(0.01)	$\nu_{36}$	591	(0.92)
$\nu_{16}$	235	(1.82)	$\nu_{37}$	595	(4.02)
$\nu_{17}$	298	(0.29)	$\nu_{38}$	597	(0.22)
$\nu_{18}$	303	(0.07)	$\nu_{39}$	598	(4.52)
$\nu_{19}$	306	(0.16)	$\nu_{40}$	598	(1.61)
$\nu_{20}$	311	(0.00)	$\nu_{41}$	660	(4.41)
$\nu_{21}$	328	(0.34)	$\nu_{42}$	660	(5.78)

<sup>a</sup> km/mole

Table C.47 – Raman ad PA infrared spectra of 1-methylcarbazole Dimer, 700- 200 cm<sup>-1</sup>.

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> )	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> )
$\nu_{43}$	725	(1.53)	$\nu_{83}$	1194	(0.30)
$\nu_{44}$	725	(0.02)	$\nu_{84}$	1194	(0.82)
$\nu_{45}$	755	(119.18)	$\nu_{85}$	1251	(5.99)
$\nu_{46}$	759	(3.30)	$\nu_{86}$	1251	(0.03)
$\nu_{47}$	771	(72.00)	$\nu_{87}$	1273	(33.85)
$\nu_{48}$	774	(0.21)	$\nu_{88}$	1275	(57.37)
$\nu_{49}$	778	(111.76)	$\nu_{89}$	1284	(0.07)
$\nu_{50}$	778	(0.25)	$\nu_{90}$	1284	(0.17)
$\nu_{51}$	814	(0.01)	$\nu_{91}$	1321	(3.08)
$\nu_{52}$	815	(8.52)	$\nu_{92}$	1322	(5.15)
$\nu_{53}$	863	(0.19)	$\nu_{93}$	1350	(0.99)
$\nu_{54}$	864	(4.19)	$\nu_{94}$	1351	(11.68)
$\nu_{55}$	873	(0.57)	$\nu_{95}$	1361	(38.49)
$\nu_{56}$	873	(4.46)	$\nu_{96}$	1361	(56.34)
$\nu_{57}$	883	(1.44)	$\nu_{97}$	1376	(2.20)
$\nu_{58}$	884	(4.35)	$\nu_{98}$	1377	(17.18)
$\nu_{59}$	920	(0.63)	$\nu_{99}$	1420	(14.61)
$\nu_{60}$	921	(0.00)	$\nu_{100}$	1421	(4.72)
$\nu_{61}$	928	(1.70)	$\nu_{101}$	1441	(0.39)
$\nu_{62}$	930	(1.94)	$\nu_{102}$	1442	(16.31)
$\nu_{63}$	959	(0.52)	$\nu_{103}$	1466	(3.15)
$\nu_{64}$	960	(3.76)	$\nu_{104}$	1470	(37.22)
$\nu_{65}$	984	(1.72)	$\nu_{105}$	1485	(18.25)
$\nu_{66}$	984	(0.00)	$\nu_{106}$	1493	(1.58)
$\nu_{67}$	994	(0.66)	$\nu_{107}$	1505	(8.23)
$\nu_{68}$	996	(0.10)	$\nu_{108}$	1505	(37.77)
$\nu_{69}$	1014	(1.61)	$\nu_{109}$	1512	(19.90)
$\nu_{70}$	1015	(2.10)	$\nu_{110}$	1513	(30.90)
$\nu_{71}$	1051	(1.31)	$\nu_{111}$	1542	(10.13)
$\nu_{72}$	1051	(8.90)	$\nu_{112}$	1543	(4.43)
$\nu_{73}$	1066	(0.64)	$\nu_{113}$	1556	(25.32)
$\nu_{74}$	1066	(6.92)	$\nu_{114}$	1557	(30.94)
$\nu_{75}$	1094	(16.80)	$\nu_{115}$	1654	(0.41)
$\nu_{76}$	1095	(1.56)	$\nu_{116}$	1655	(1.90)
$\nu_{77}$	1121	(0.00)	$\nu_{117}$	1667	(2.17)
$\nu_{78}$	1121	(3.83)	$\nu_{118}$	1668	(27.38)
$\nu_{79}$	1153	(1.39)	$\nu_{119}$	1681	(6.50)
$\nu_{80}$	1154	(13.29)	$\nu_{120}$	1681	(16.08)
$\nu_{81}$	1184	(0.01)	$\nu_{121}$	1704	(15.67)
$\nu_{82}$	1184	(3.93)	$\nu_{122}$	1704	(0.36)

<sup>a</sup> km/mole

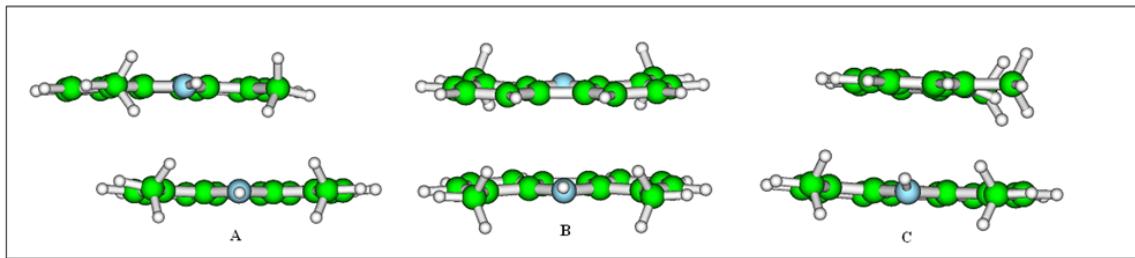


Figure C.5 – Configurations of 1,8-dimethylcarbazole dimers optimized at the  $\omega$ B97X-D/6-311++G\*\* level of theory.

Table C.48 – Interaction energies of 1,8-dimethylcarbazole dimers (kcal/mol) using  $\omega$ B97X-D/6-311++G\*\*.

	<b>A</b>	<b>B</b>	<b>C</b>
$\omega$ <b>B97x-D / 6-311++G**</b>	-12.83	-18.72	-17.72

Table C.49 – Intermolecular distance between pyrrole ring center ( $\text{\AA}$ ) and angles ( $^\circ$ ) according figure of 1,8-dimethylcarbazole dimers using  $\omega$ B97x-D/6-311++G\*\*.

	<b>A</b>	<b>B</b>	<b>C</b>
<b>d</b>	3.90	3.78	3.42
$\alpha$	6.22	3.04	2.10
$\beta$	7.58	2.30	1.64
$\gamma$			96.85

Table C.50 – Calculated low wavenumber Raman ad PA infrared spectra of 1,8-dimethylcarbazole Dimer.

<b>Mode</b>	$\omega \text{ cm}^{-1}$	<b>Intensities</b> ( <sup>a</sup> )	<b>Mode</b>	$\omega \text{ cm}^{-1}$	<b>Intensities</b> ( <sup>a</sup> )
$\nu_1$	31	(0.12)	$\nu_{11}$	160	(0.00)
$\nu_2$	34	(0.00)	$\nu_{12}$	160	(1.48)
$\nu_3$	54	(0.00)	$\nu_{13}$	180	(6.56)
$\nu_4$	73	(0.00)	$\nu_{14}$	198	(0.00)
$\nu_5$	84	(1.31)	$\nu_{15}$	202	(0.00)
$\nu_6$	103	(0.00)	$\nu_{16}$	214	(0.90)
$\nu_7$	105	(4.93)	$\nu_{17}$	229	(0.00)
$\nu_8$	117	(0.00)	$\nu_{18}$	245	(4.36)
$\nu_9$	123	(0.00)	$\nu_{19}$	250	(0.02)
$\nu_{10}$	125	(0.00)	$\nu_{20}$	268	(0.00)

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> )	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> )
$\nu_{20}$	280	(0.00)	$\nu_{34}$	514	(0.00)
$\nu_{21}$	282	(0.13)	$\nu_{35}$	515	(9.44)
$\nu_{22}$	301	(0.19)	$\nu_{36}$	538	(0.12)
$\nu_{23}$	306	(0.00)	$\nu_{37}$	541	(0.00)
$\nu_{24}$	314	(10.50)	$\nu_{38}$	567	(0.00)
$\nu_{25}$	317	(0.00)	$\nu_{39}$	567	(0.98)
$\nu_{26}$	329	(5.77)	$\nu_{40}$	588	(0.00)
$\nu_{27}$	338	(0.00)	$\nu_{41}$	589	(15.27)
$\nu_{28}$	385	(212.09)	$\nu_{42}$	591	(0.00)
$\nu_{29}$	390	(0.00)	$\nu_{43}$	591	(9.91)
$\nu_{30}$	476	(3.47)	$\nu_{44}$	597	(0.00)
$\nu_{31}$	476	(0.00)	$\nu_{45}$	597	(1.32)
$\nu_{32}$	491	(0.00)	$\nu_{46}$	603	(0.00)
$\nu_{33}$	492	(0.00)	$\nu_{47}$	603	(1.69)

<sup>a</sup> km/moleTable C.51 – Calculated Raman and PA infrared spectra of 1,8-dimethylcarbazole Dimer, 700–2000  $\text{cm}^{-1}$ 

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{49}$	712	[0.23]	$\nu_{64}$	883	[0.00]
$\nu_{50}$	712	[0.00]	$\nu_{65}$	921	[0.11]
$\nu_{51}$	767	[1.03]	$\nu_{66}$	922	[0.00]
$\nu_{52}$	768	[0.00]	$\nu_{67}$	924	[0.00]
$\nu_{53}$	771	[0.00]	$\nu_{68}$	924	[1.52]
$\nu_{54}$	774	[71.06]	$\nu_{69}$	974	[8.84]
$\nu_{55}$	786	[168.42]	$\nu_{70}$	975	[0.00]
$\nu_{56}$	787	[0.00]	$\nu_{71}$	985	[0.00]
$\nu_{57}$	817	[0.00]	$\nu_{72}$	986	[0.03]
$\nu_{58}$	818	[0.02]	$\nu_{73}$	987	[1.73]
$\nu_{59}$	821	[0.00]	$\nu_{74}$	988	[0.00]
$\nu_{60}$	822	[0.00]	$\nu_{75}$	1011	[1.27]
$\nu_{61}$	882	[1.22]	$\nu_{76}$	1012	[0.00]
$\nu_{62}$	882	[0.00]	$\nu_{77}$	1054	[0.00]
$\nu_{63}$	882	[8.78]	$\nu_{78}$	1054	[0.25]

Table C.52 – Raman ad PA infrared spectra of 1,8-methylcarbazole Dimer, 700- 200 cm<sup>-1</sup>.

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> )	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> )
$\nu_{78}$	1069	[0.00]	$\nu_{107}$	1427	[14.30]
$\nu_{79}$	1070	[8.53]	$\nu_{108}$	1427	[0.00]
$\nu_{80}$	1071	[1.42]	$\nu_{109}$	1427	[0.00]
$\nu_{81}$	1071	[0.00]	$\nu_{110}$	1440	[12.74]
$\nu_{82}$	1101	[0.00]	$\nu_{111}$	1441	[0.00]
$\nu_{83}$	1101	[26.06]	$\nu_{112}$	1466	[53.13]
$\nu_{84}$	1107	[26.52]	$\nu_{113}$	1466	[0.00]
$\nu_{85}$	1108	[0.00]	$\nu_{114}$	1469	[68.57]
$\nu_{86}$	1155	[0.00]	$\nu_{115}$	1470	[0.00]
$\nu_{87}$	1157	[9.90]	$\nu_{116}$	1486	[44.27]
$\nu_{88}$	1192	[0.00]	$\nu_{117}$	1487	[0.00]
$\nu_{89}$	1192	[0.32]	$\nu_{118}$	1488	[20.51]
$\nu_{90}$	1196	[0.00]	$\nu_{119}$	1488	[0.00]
$\nu_{91}$	1197	[1.52]	$\nu_{120}$	1510	[12.21]
$\nu_{92}$	1267	[35.79]	$\nu_{121}$	1511	[32.25]
$\nu_{93}$	1269	[0.00]	$\nu_{122}$	1512	[0.00]
$\nu_{94}$	1279	[0.62]	$\nu_{123}$	1514	[0.00]
$\nu_{95}$	1279	[0.00]	$\nu_{124}$	1548	[6.60]
$\nu_{96}$	1285	[17.08]	$\nu_{125}$	1550	[0.00]
$\nu_{97}$	1286	[0.00]	$\nu_{126}$	1561	[79.70]
$\nu_{98}$	1313	[0.37]	$\nu_{127}$	1562	[0.00]
$\nu_{99}$	1316	[0.00]	$\nu_{128}$	1663	[33.53]
$\nu_{100}$	1354	[0.00]	$\nu_{129}$	1664	[0.00]
$\nu_{101}$	1355	[15.93]	$\nu_{130}$	1667	[13.77]
$\nu_{102}$	1357	[0.00]	$\nu_{131}$	1668	[20.03]
$\nu_{103}$	1359	[141.23]	$\nu_{132}$	1682	[0.00]
$\nu_{104}$	1376	[10.92]	$\nu_{133}$	1683	[2.44]
$\nu_{105}$	1377	[0.00]	$\nu_{134}$	1694	[0.00]

<sup>a</sup> km/mole

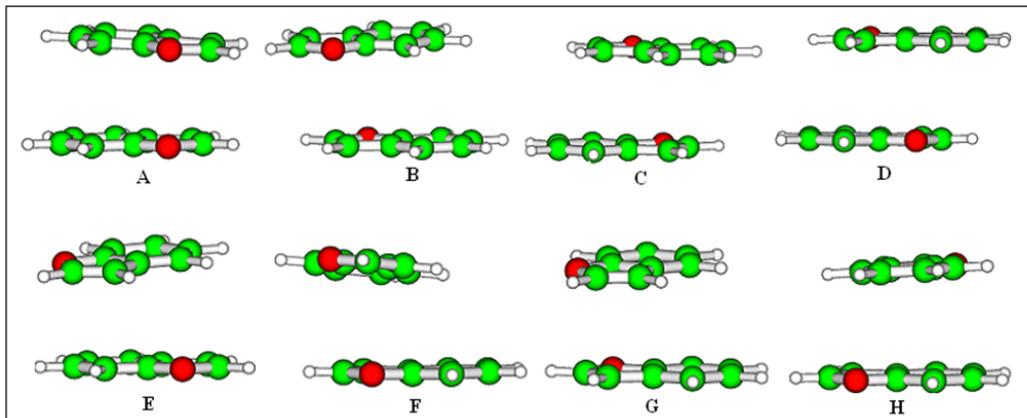


Figure C.6

Configurations of benzofuran dimers optimized at the  $\omega$ B97X-D/6-311++G\*\* level of theory.

Table C.53 – Interaction energies of benzofuran dimers (kcal/mol) in 8 different configurations using  $\omega$ B97X-D/6-311++G\*\*. and compared with other methods.

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>
$\omega$ B97X-D/ 6-311++G**	-5.79	-6.21	-6.40	-6.72	-6.28	-6.50	-6.40	-6.52
DFT-SAPT(PBE0) aug-ccpvtz	-4.47	-4.81	-5.00	-5.29	-4.84	-5.24	-4.91	-4.99
LMP2/ aug-ccpvtz <sup>a</sup>	-4.60							

Table C.54 – Intermolecular distance between furan rings center of benzofuran dimers ( $\text{\AA}$ )

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>
$\omega$ B97x-D/ 6-311++G**	3.84	3.61	3.52	3.65	4.37	3.46	3.68	4.08
DFT-SAPT(PBE0)/ aug-ccpvtz	3.86	3.63	3.53	3.65	4.39	3.47	3.71	4.08

Table C.55 – Angle between ring planes of benzofuran dimers ( $^\circ$ ) at the  $\omega$ B97x-D/6-311++G\*\* level of theory. according to figure

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>
$\alpha$	18.10	4.12	9.26	5.80	9.82	6.47	5.54	9.89
$\beta$	14.08	5.53	6.98	4.08	9.17	4.87	9.02	9.58
$\gamma$					63.58	100.73	61.49	81.51

Table C.56 – Calculated low wavenumber Raman ad PA infrared spectra of Benzofuran Dimer.

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_1$	16	(0.13) [0.00]	$\nu_{12}$	413	(8.20) [0.00]
$\nu_2$	33	(0.00) [7.09]	$\nu_{13}$	433	(20.00) [0.00]
$\nu_3$	60	(0.00) [7.03]	$\nu_{14}$	435	(0.00) [1.23]
$\nu_4$	69	(0.80) [0.00]	$\nu_{15}$	553	(1.18) [0.01]
$\nu_5$	73	(0.34) [0.00]	$\nu_{16}$	553	(0.00) [10.7]
$\nu_6$	93	(0.00) [4.56]	$\nu_{17}$	587	(0.95) [0.00]
$\nu_7$	229	(0.002) [1.03]	$\nu_{18}$	588	(0.00) [0.47]
$\nu_8$	232	(15.24) [0.00]	$\nu_{19}$	602	(0.00) [3.35]
$\nu_9$	258	(2.59) [0.00]	$\nu_{20}$	605	(1.26) [0.00]
$\nu_{10}$	263	(0.00) [0.93]	$\nu_{21}$	628	(0.00) [10.86]
$\nu_{11}$	413	(0.01) [0.27]	$\nu_{22}$	628	(2.02) [0.00]

<sup>a</sup> km/mole    <sup>b</sup> Å<sup>4</sup>/AMUTable C.57 – Raman ad PA infrared spectra of Benzofuran Dimer, 700- 2000 cm<sup>-1</sup>

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{23}$	751	(0.003) [1.11]	$\nu_{43}$	1042	(10.47) [0.05]
$\nu_{24}$	755	(23.17) [0.00]	$\nu_{44}$	1042	(0.01) [42.05]
$\nu_{25}$	771	(243.20) [0.00]	$\nu_{45}$	1073	(0.00) [17.51]
$\nu_{26}$	774	(0.006) [0.44]	$\nu_{46}$	1074	(20.96) [0.00]
$\nu_{27}$	789	(11.41) [0.00]	$\nu_{47}$	1136	(0.003) [6.05]
$\nu_{28}$	789	(0.00) [36.07]	$\nu_{48}$	1137	(21.91) [0.00]
$\nu_{29}$	792	(42.03) [0.00]	$\nu_{49}$	1165	(31.38) [0.00]
$\nu_{30}$	793	(0.06) [0.47]	$\nu_{50}$	1166	(0.00) [10.5]
$\nu_{31}$	878	(0.00) [0.49]	$\nu_{51}$	1182	(0.00) [3.39]
$\nu_{32}$	878	(2.03) [0.00]	$\nu_{52}$	1183	(5.03) [0.00]
$\nu_{33}$	882	(0.00) [4.95]	$\nu_{53}$	1212	(0.02) [2.15]
$\nu_{34}$	882	(13.55) [0.00]	$\nu_{54}$	1212	(34.44) [0.001]
$\nu_{35}$	897	(0.04) [6.85]	$\nu_{55}$	1299	(0.002) [28.04]
$\nu_{36}$	897	(5.37) [0.04]	$\nu_{56}$	1300	(9.36) [0.01]
$\nu_{37}$	922	(0.005) [0.01]	$\nu_{57}$	1301	(0.00) [48.82]
$\nu_{38}$	922	(5.15) [0.01]	$\nu_{58}$	1305	(58.62) [0.00]
$\nu_{39}$	964	(6.58) [0.00]	$\nu_{59}$	1371	(9.76) [0.001]
$\nu_{40}$	965	(0.00) [0.21]	$\nu_{60}$	1373	(0.00) [28.56]
$\nu_{41}$	997	(0.00) [0.49]	$\nu_{61}$	1396	(3.53) [0.001]
$\nu_{42}$	998	(0.26) [0.00]	$\nu_{62}$	1396	0.00) [20.70]

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{63}$	1502	(76.23) [0.003]	$\nu_{68}$	1609	(24.08) [0.00]
$\nu_{64}$	1503	(0.02) [10.81]	$\nu_{69}$	1664	(0.18) [0.01]
$\nu_{65}$	1524	(7.06) [0.002]	$\nu_{70}$	1664	(0.00) [15.60]
$\nu_{66}$	1525	(0.002) [8.11]	$\nu_{71}$	1691	0.50) [0.96]
$\nu_{67}$	1608	(0.00) [143.49]	$\nu_{72}$	1691	(0.008) [65.93]

<sup>a</sup> km/mole    <sup>b</sup> Å<sup>4</sup>/AMU

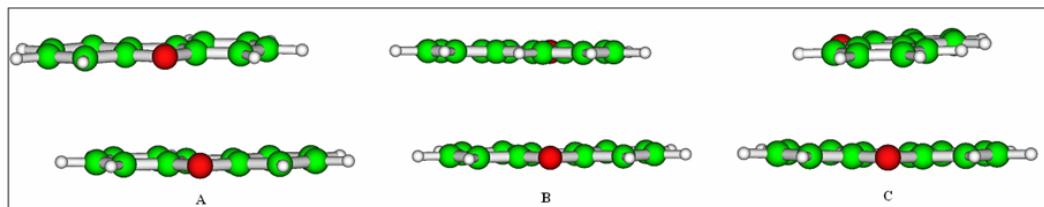


Figure C.7 – Configurations of dibenzofuran dimers optimized at the  $\omega\text{B97x-D}/6-311++\text{G}^{**}$  level of theory.

Table C.58 – Interaction energies of dibenzofuran dimers (kcal/mol) in 3 different configurations using  $\omega\text{B97x-D}/6-311++\text{G}^{**}$ .

	A	B	C
$\omega\text{B97x-D} / 6-311++\text{G}^{**}$	-10.03	-10.38	-9.58

Table C.59 – ntermolecular distance between furan ring center (Å) according figure of dibenzofuran dimers using  $\omega\text{B97x-D}/6-311++\text{G}^{**}$ .

	A	B	C
d	3.67	3.82	3.39

Table C.60 – Calculated low wavenumber Raman ad PA infrared spectra of Dibenzofuran Dimer.

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_1$	12	(0.03) [2.29]	$\nu_{18}$	436	(1.47) [5.71]
$\nu_2$	27	(0.04) [5.45]	$\nu_{19}$	436	(1.41) [7.34]
$\nu_3$	37	(0.09) [9.99]	$\nu_{20}$	437	(2.71) [1.85]
$\nu_4$	63	(0.5) [0.69]	$\nu_{21}$	447	(0.12) [0.38]
$\nu_5$	78	(0.26) [1.31]	$\nu_{22}$	449	(0.19) [0.41]
$\nu_6$	88	(0.08) [1.92]	$\nu_{23}$	531	(0.01) [2.63]
$\nu_7$	121	(4.03) [0.50]	$\nu_{24}$	532	(0.20) [1.01]
$\nu_8$	126	(0.04) [1.14]	$\nu_{25}$	573	(3.35) [4.98]
$\nu_9$	163	(0.04) [0.36]	$\nu_{26}$	573	(1.76) [6.88]
$\nu_{10}$	167	(0.06) [0.32]	$\nu_{27}$	580	(1.47) [0.31]
$\nu_{11}$	222	(0.34) [1.09]	$\nu_{28}$	580	(7.51) [0.38]
$\nu_{12}$	222	(1.52) [0.31]	$\nu_{29}$	586	(0.15) [0.16]
$\nu_{13}$	299	(0.05) [6.95]	$\nu_{30}$	587	(0.01) [0.02]
$\nu_{14}$	302	(0.10) [1.37]	$\nu_{31}$	633	(0.86) [0.05]
$\nu_{15}$	321	(0.28) [0.09]	$\nu_{32}$	634	(7.41) [0.01]
$\nu_{16}$	323	(0.02) [1.90]	$\nu_{33}$	678	(0.08) [1.58]
$\nu_{17}$	434	(10.51) [0.60]	$\nu_{31}$	679	(1.07) [0.12]

<sup>a</sup> km/mole    <sup>b</sup> A<sup>4</sup>/AMUTable C.61 – Raman ad PA infrared spectra of Dibenzofuran Dimer, 700- 2000 cm<sup>-1</sup>

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{32}$	746	(140.16) [0.05]	$\nu_{46}$	874	(0.48) [0.36]
$\nu_{33}$	748	(2.57) [0.45]	$\nu_{47}$	876	(0.37) [5.56]
$\nu_{32}$	763	(0.39) [0.64]	$\nu_{48}$	876	(12.21) [5.69]
$\nu_{33}$	766	(4.67) [0.11]	$\nu_{49}$	877	(10.55) [1.63]
$\nu_{32}$	767	(0.56) [2.67]	$\nu_{50}$	877	(23.66) [8.54]
$\nu_{33}$	768	0.02) [39.98]	$\nu_{51}$	878	(3.55) [6.61]
$\nu_{32}$	774	(0.20) [2.90]	$\nu_{52}$	879	(4.99) [0.55]
$\nu_{33}$	776	(27.11) [0.30]	$\nu_{53}$	962	(4.95) [0.25]
$\nu_{43}$	777	(180.51) [180.51]	$\nu_{54}$	964	(1.78) [0.13]
$\nu_{44}$	780	(8.58) [8.58]	$\nu_{55}$	965	(3.33) [0.18]
$\nu_{45}$	873	(4.45) [4.45]	$\nu_{56}$	966	(0.43) [0.21]

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{56}$	966	(0.43) [0.21]	$\nu_{81}$	1316	(6.02) [1.48]
$\nu_{57}$	996	(0.23) [0.31]	$\nu_{82}$	1316	(26.74) [1.35]
$\nu_{58}$	997	(0.18) [0.11]	$\nu_{83}$	1343	(1.77) [66.22]
$\nu_{59}$	999	(0.05) [0.24]	$\nu_{84}$	1344	(1.88) [92.12]
$\nu_{60}$	1000	(0.18) [0.14]	$\nu_{85}$	1368	(17.78) [0.90]
$\nu_{61}$	1031	(0.08) [1.49]	$\nu_{86}$	1368	(3.08) [2.31]
$\nu_{62}$	1032	(1.08) [0.16]	$\nu_{87}$	1389	(4.41) (26.90)
$\nu_{63}$	1045	(0.81) [2.21]	$\nu_{88}$	1390	(3.81) [24.86]
$\nu_{64}$	1045	(0.09) [81.70]	$\nu_{89}$	1493	(47.13) [8.24]
$\nu_{65}$	1052	(3.86) [0.18]	$\nu_{90}$	1494	(18.13) [19.34]
$\nu_{66}$	1052	(17.66) [0.05]	$\nu_{91}$	1500	(45.74) [0.63]
$\nu_{67}$	1135	(1.19) [10.14]	$\nu_{92}$	1500	(11.05) [1.09]
$\nu_{68}$	1135	(14.60) [0.54]	$\nu_{93}$	1522	(11.12) [8.31]
$\nu_{69}$	1146	(1.53) [4.05]	$\nu_{94}$	1522	(67.33) [4.38]
$\nu_{70}$	1146	(9.83) [1.44]	$\nu_{95}$	1542	(0.44) [54.73]
$\nu_{71}$	1178	(1.34) [12.15]	$\nu_{96}$	1543	(1.04) [6.90]
$\nu_{72}$	1179	(2.83) [8.51]	$\nu_{97}$	1662	(1.41) [6.10]
$\nu_{73}$	1187	(0.18) [1.54]	$\nu_{98}$	1662	(11.95) [1.93]
$\nu_{74}$	1187	(0.14) [1.41]	$\nu_{99}$	1668	(0.41) [17.86]
$\nu_{75}$	1232	(1.15) [13.45]	$\nu_{100}$	1668	(1.61) [9.85]
$\nu_{76}$	1233	(2.12) [5.32]	$\nu_{101}$	1675	(1.42) [3.71]
$\nu_{77}$	1244	(27.73) [6.88]	$\nu_{102}$	1675	(15.67) [1.05]
$\nu_{78}$	1249	(205.63) [2.36]	$\nu_{103}$	1713	(0.23) [269.45]
$\nu_{79}$	1290	(4.37) [123.94]	$\nu_{104}$	1713	(0.50) [199.85]

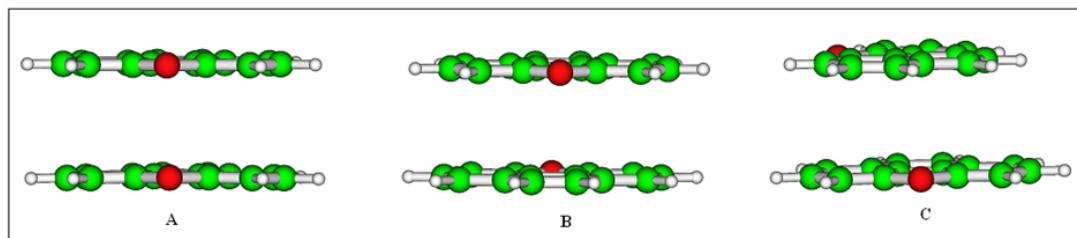
<sup>a</sup> km/mole<sup>b</sup>  $\text{A}^4/\text{AMU}$ 

Figure C.8 – Configurations of Tribenzofuran dimers optimized at the  $\omega\text{B97x-D}/6-311++\text{G}^{**}$  level of theory.

Table C.62 – Interaction energies of Tribenzofuran dimers (kcal/mol) in 3 different configurations using  $\omega$ B97x-D/6-311++G\*\*.

	A	B	C
$\omega$ B97x-D/ 6-311++G**	-12.31	-13.30	-13.58

Table C.63 – ntermolecular distance between furan ring center ( $\text{\AA}$ ) according figure of Tribenzofuran dimers using  $\omega$ B97x-D/6-311++G\*\*.

	A	B	C
d	3.68	4.71	4.25

Table C.64 – Calculated low wavenumber Raman ad PA infrared spectra of Tribenzofuran Dimer.

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_1$	21	(0.05) [2.13]	$\nu_{21}$	457	(0.04) [16.62]
$\nu_2$	23	(0.16) [11.11]	$\nu_{22}$	458	(0.01) [17.46]
$\nu_3$	49	(0.00) [9.65]	$\nu_{23}$	491	(1.36) [1.74]
$\nu_4$	77	(0.40) [0.19]	$\nu_{24}$	491	(1.93) [3.62]
$\nu_5$	79	(0.01) [2.70]	$\nu_{25}$	504	(0.29) [0.34]
$\nu_6$	92	(3.12) [1.14]	$\nu_{26}$	506	(8.12) [0.03]
$\nu_7$	98	(4.42) [1.12]	$\nu_{27}$	557	(0.40) [2.18]
$\nu_8$	119	(0.00) [3.63]	$\nu_{28}$	558	(0.45) [2.84]
$\nu_9$	200	(0.03) [0.18]	$\nu_{29}$	576	(0.004) [0.09]
$\nu_{10}$	212	(0.01) [1.01]	$\nu_{30}$	578	(0.003) [0.28]
$\nu_{11}$	222	(0.08) [2.53]	$\nu_{31}$	585	(2.60) [0.57]
$\nu_{12}$	228	(5.98) [0.46]	$\nu_{32}$	586	(2.93) [0.77]
$\nu_{13}$	307	(0.21) [1.72]	$\nu_{33}$	611	(8.03) [0.17]
$\nu_{14}$	309	(0.09) [1.22]	$\nu_{34}$	612	(0.40) [0.11]
$\nu_{15}$	326	(0.62) [1.06]	$\nu_{35}$	623	(1.39) [48.21]
$\nu_{16}$	329	(0.11) [1.82]	$\nu_{36}$	624	(2.54) [0.54]
$\nu_{17}$	348	(3.43) [0.30]	$\nu_{37}$	663	(0.09) [0.11]
$\nu_{18}$	348	(4.57) [0.26]	$\nu_{38}$	663	(0.19) [0.58]
$\nu_{19}$	437	(0.09) [1.52]	$\nu_{39}$	690	(0.63) [0.15]
$\nu_{20}$	439	(0.02) [2.40]	$\nu_{40}$	691	(0.30) [13.37]

<sup>a</sup> km/mole<sup>b</sup>  $\text{A}^4/\text{AMU}$

Table C.65 – Raman ad PA infrared spectra of Dibenzofuran Dimer, 700- 2000 cm<sup>-1</sup>

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{41}$	707	(1.30) [0.10]	$\nu_{79}$	1196	(2.85) [0.37]
$\nu_{42}$	707	(1.01) [0.08]	$\nu_{80}$	1196	(4.37) [0.89]
$\nu_{43}$	734	(91.56) [0.03]	$\nu_{81}$	1205	(0.00) [0.55]
$\nu_{44}$	735	(0.70) [0.38]	$\nu_{82}$	1206	(0.003) [0.51]
$\nu_{45}$	767	(1.42) [0.20]	$\nu_{83}$	1243	(6.01) [998.56]
$\nu_{46}$	768	(2.08) [0.16]	$\nu_{84}$	1243	(10.43) [24.23]
$\nu_{47}$	775	(12.66) [0.13]	$\nu_{85}$	1253	(2.95) [0.48]
$\nu_{48}$	776	(64.56) [0.12]	$\nu_{86}$	1254	(0.06) [8.25]
$\nu_{49}$	808	(0.58) [1.90]	$\nu_{87}$	1257	(9.46) [2.56]
$\nu_{50}$	208	(2.49) [1.20]	$\nu_{88}$	1258	(20.06) [7.40]
$\nu_{51}$	815	(3.45) [0.92]	$\nu_{89}$	1274	(16.27) [0.49]
$\nu_{52}$	815	(5.20) [1.93]	$\nu_{90}$	1275	(35.04) [3.28]
$\nu_{53}$	838	(3.35) [7.73]	$\nu_{91}$	1352	(16.11) [42.19]
$\nu_{54}$	839	(3.04) [10.57]	$\nu_{92}$	1352	(10.15) [62.41]
$\nu_{55}$	842	(173.42) [0.30]	$\nu_{93}$	1403	(1.25) [5.26]
$\nu_{56}$	845	(29.05) [0.72]	$\nu_{94}$	1404	(3.42) [14.86]
$\nu_{57}$	906	(0.12) [0.48]	$\nu_{95}$	1405	(8.16) [0.72]
$\nu_{58}$	907	(0.03) [0.56]	$\nu_{96}$	1407	(19.96) [0.53]
$\nu_{59}$	913	(3.04) [0.23]	$\nu_{97}$	1441	(0.28) [15.89]
$\nu_{60}$	914	(3.85) [0.27]	$\nu_{98}$	1442	(0.33) [25.67]
$\nu_{61}$	987	(0.61) [0.13]	$\nu_{99}$	1467	(34.56) [0.39]
$\nu_{62}$	987	(0.15) [0.04]	$\nu_{100}$	1467	(20.10) [0.61]
$\nu_{63}$	989	(0.35) [0.06]	$\nu_{101}$	1491	(20.89) [5.82]
$\nu_{64}$	989	(3.52) [0.02]	$\nu_{102}$	1492	(11.19) [34.98]
$\nu_{65}$	998	(1.54) [1.93]	$\nu_{103}$	1535	(0.003) [361.12]
$\nu_{66}$	1000	(1.47) [2.13]	$\nu_{104}$	1536	(0.24) [50.34]
$\nu_{67}$	1025	(5.30) [1.67]	$\nu_{105}$	1543	(0.11) [0.83]
$\nu_{68}$	1026	(3.01) [1.85]	$\nu_{106}$	1543	(0.25) [2.20]
$\nu_{69}$	1027	(7.67) [2.94]	$\nu_{107}$	1563	(12.63) [2.49]
$\nu_{70}$	1028	(0.23) [0.37]	$\nu_{108}$	1564	(21.52) [0.59]
$\nu_{71}$	1045	(0.14) [48.95]	$\nu_{109}$	1657	(34.71) [0.18]
$\nu_{72}$	1046	(0.47) [12.77]	$\nu_{110}$	1658	(67.22) [3.42]
$\nu_{73}$	1071	(19.53) [0.67]	$\nu_{111}$	1699	(0.82) [29.36]
$\nu_{74}$	1072	(24.13) [2.47]	$\nu_{112}$	1699	(1.19) [16.13]
$\nu_{75}$	1075	(32.97) [0.31]	$\nu_{113}$	1722	(1.37) [70.71]
$\nu_{76}$	1075	(72.14) [1.23]	$\nu_{114}$	1722	(0.74) [38.60]
$\nu_{77}$	1165	(0.21) [3.04]	$\nu_{115}$	1752	(0.06) [23.64]
$\nu_{78}$	1166	(0.89) [0.35]	$\nu_{116}$	1753	(0.58) [26.38]

<sup>a</sup> km/mole<sup>b</sup> A<sup>4</sup>/AMU

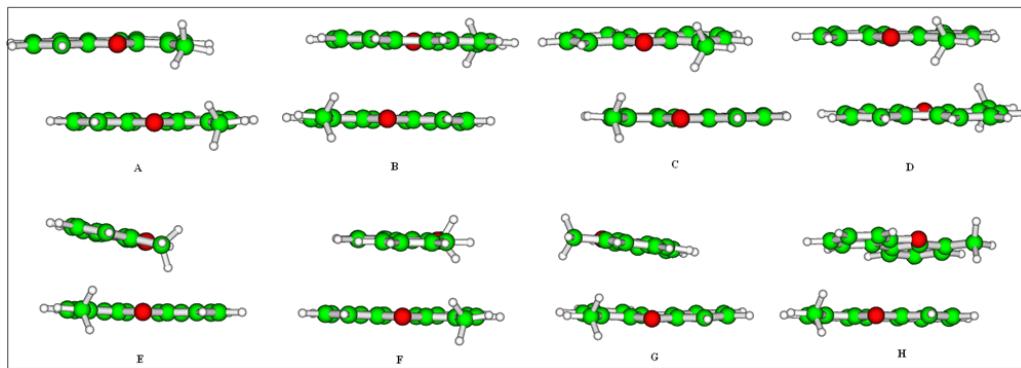


Figure C.9 – Configurations of 4-methyldibenzofuran dimers optimized at the  $\omega$ B97x-D/6-311++G\*\* level of theory.

Table C.66 – Interaction energies of 4-methyldibenzofuran dimers (kcal/mol) using  $\omega$ B97X-D/6-311++G\*\*.

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>
<b><math>\omega</math>B97X-D / 6-311++G**</b>	-11.54	-11.74	-11.59	-11.29	-12.67	-12.30	-12.79	-11.87

Table C.67 – Calculated low wavenumber Raman ad PA infrared spectra of 4-methyldibenzofuran Dimer.

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_1$	15	(0.01) [0.00]	$\nu_{16}$	163	(1.17) [0.00]
$\nu_2$	24	(0.00) [6.54]	$\nu_{17}$	209	(6.26) [0.00]
$\nu_3$	39	(0.00) [7.95]	$\nu_{18}$	221	(0.00) [1.98]
$\nu_4$	78	(0.05) [0.00]	$\nu_{19}$	247	(0.00) [2.5]
$\nu_5$	79	(0.40) [0.00]	$\nu_{20}$	259	(0.35) [0.00]
$\nu_6$	90	(0.00) [1.58]	$\nu_{21}$	268	(0.00) [3.52]
$\nu_7$	99	(0.12) [0.00]	$\nu_{22}$	268	(1.88) [0.00]
$\nu_8$	102	(0.00) [0.27]	$\nu_{23}$	305	(0.00) [4.76]
$\nu_9$	111	(0.001) [1.7]	$\nu_{24}$	309	(0.40) [0.00]
$\nu_{10}$	111	(5.17) [0.00]	$\nu_{25}$	319	(0.00) [6.38]
$\nu_{11}$	120	(0.18) [0.00]	$\nu_{26}$	319	(0.22) [0.00]
$\nu_{12}$	141	(0.00) [1.24]	$\nu_{27}$	341	(3.68) [0.00]
$\nu_{13}$	149	(0.00) [0.75]	$\nu_{28}$	345	(0.00) [3.36]
$\nu_{14}$	150	(0.70) [0.00]	$\nu_{29}$	472	(4.21) [0.00]
$\nu_{15}$	162	(0.00) [0.46]	$\nu_{30}$	472	(0.00) [3.31]

<sup>a</sup> km/mole

<sup>b</sup> A<sup>4</sup>/AMU

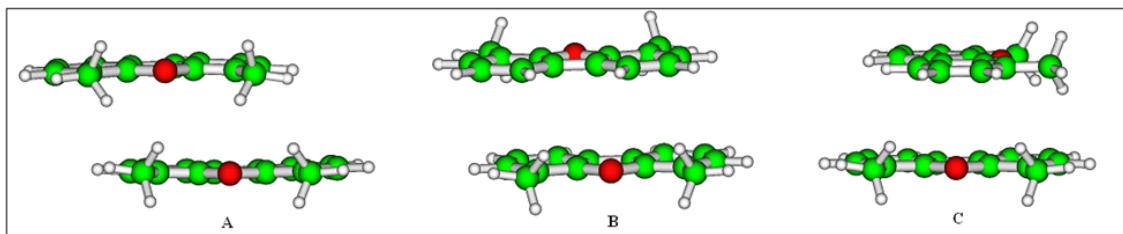


Figure C.10 – Configurations of 4,6-dimethyldibenzofuran dimers optimized at the  $\omega$ B97x-D/6-311++G\*\* level of theory.

Table C.68 – Interaction energies of 4,6-dimethyldibenzofuran dimers (kcal/mol) in 3 different configurations using  $\omega$ B97x-D/6-311++G\*\*.

	A	B	C
$\omega$ B97x-D/ 6-311++G**	-13.32	-13.74	-13.63

Table C.69 – Calculated low wavenumber Raman ad PA infrared spectra of 4,6-dimethyldibenzofuran Dimer.

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_1$	15	(0.01) [0.00]	$\nu_{19}$	247	(0.00) [2.5]
$\nu_2$	24	(0.00) [6.54]	$\nu_{20}$	259	(0.35) [0.00]
$\nu_3$	39	(0.00) [7.95]	$\nu_{21}$	268	(0.00) [3.52]
$\nu_4$	78	(0.05) [0.00]	$\nu_{22}$	268	(1.88) [0.00]
$\nu_5$	79	(0.40) [0.00]	$\nu_{23}$	305	(0.00) [4.76]
$\nu_6$	90	(0.00) [1.58]	$\nu_{24}$	309	(0.40) [0.00]
$\nu_7$	99	(0.12) [0.00]	$\nu_{25}$	319	(0.00) [6.38]
$\nu_8$	102	(0.00) [0.27]	$\nu_{26}$	319	(0.22) [0.00]
$\nu_9$	111	(0.001) [1.7]	$\nu_{27}$	341	(3.68) [0.00]
$\nu_{10}$	111	(5.17) [0.00]	$\nu_{28}$	345	(0.00) [3.36]
$\nu_{11}$	120	(0.18) [0.00]	$\nu_{29}$	472	(4.21) [0.00]
$\nu_{12}$	141	(0.00) [1.24]	$\nu_{30}$	472	(0.00) [3.31]
$\nu_{13}$	149	(0.00) [0.75]	$\nu_{31}$	493	(0.00) [5.85]
$\nu_{14}$	150	(0.70) [0.00]	$\nu_{32}$	493	(0.03) [0.00]
$\nu_{15}$	162	(0.00) [0.46]	$\nu_{33}$	507	(0.00) [2.00]
$\nu_{16}$	163	(1.17) [0.00]	$\nu_{34}$	510	(4.30) [0.00]
$\nu_{17}$	209	(6.26) [0.00]	$\nu_{35}$	534	(0.43) [0.00]
$\nu_{18}$	221	(0.00) [1.98]	$\nu_{36}$	536	(0.00) [0.43]

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{37}$	571	(2.60) [0.00]	$\nu_{42}$	594	(0.00) [0.99]
$\nu_{38}$	571	(0.00) [36.63]	$\nu_{43}$	597	(0.00) [2.52]
$\nu_{39}$	589	(0.62) [0.00]	$\nu_{44}$	597	(12.52) [0.00]
$\nu_{40}$	589	(0.00) [0.25]	$\nu_{45}$	606	(1.37) [0.00]
$\nu_{41}$	593	(0.21) [0.00]	$\nu_{46}$	606	(0.00) [6.12]

<sup>a</sup> km/mole    <sup>b</sup>  $\text{A}^4/\text{AMU}$ Table C.70 – Raman ad PA infrared spectra of 4,6-dimethylbibenzofuran Dimer, 700- 2000  $\text{cm}^{-1}$ 

Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega \text{ cm}^{-1}$	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{47}$	710	(1.11) [0.00]	$\nu_{79}$	1067	(0.03) [1.87]
$\nu_{48}$	710	(0.00) [19.60]	$\nu_{80}$	1067	(11.96) [0.003]
$\nu_{49}$	766	(0.00) [0.38]	$\nu_{81}$	1102	(29.32) [0.00]
$\nu_{50}$	767	(9.31) [0.00]	$\nu_{82}$	1103	(0.00) [2.77]
$\nu_{51}$	773	(74.58) [0.00]	$\nu_{83}$	1110	(0.00) [27.92]
$\nu_{52}$	773	(0.00) [0.79]	$\nu_{84}$	1110	(17.19) [0.00]
$\nu_{53}$	786	(0.004) [1.25]	$\nu_{85}$	1175	(9.69) [0.00]
$\nu_{54}$	786	(187.62) [0.00]	$\nu_{86}$	1175	(0.00) [1.37]
$\nu_{55}$	810	(0.00) [2.44]	$\nu_{87}$	1192	(0.23) [0.00]
$\nu_{56}$	810	(2.31) [0.00]	$\nu_{88}$	1192	(0.00) [2.28]
$\nu_{57}$	818	(0.00) [1.11]	$\nu_{89}$	1211	(0.00) [2.26]
$\nu_{58}$	818	(4.17) [0.00]	$\nu_{90}$	1211	(11.49) [0.00]
$\nu_{60}$	866	(0.00) [22.18]	$\nu_{91}$	1238	(0.00) [3.85]
$\nu_{60}$	866	(0.00) [22.18]	$\nu_{92}$	1243	(222.22) [0.00]
$\nu_{61}$	887	(0.00) [2.73]	$\nu_{93}$	1262	(0.00) [5.19]
$\nu_{62}$	888	(36.91) [0.00]	$\nu_{94}$	1263	(3.02) [0.00]
$\nu_{63}$	916	(2.51) [0.00]	$\nu_{95}$	1299	(0.00) [54.18]
$\nu_{64}$	917	(0.00) [0.61]	$\nu_{96}$	1299	(2.87) [0.00]
$\nu_{65}$	919	(0.00) [0.21]	$\nu_{97}$	1305	(0.00) 4.15]
$\nu_{66}$	920	(1.53) [0.00]	$\nu_{98}$	1306	(31.80) [0.00]
$\nu_{67}$	969	(0.00) [45.72]	$\nu_{99}$	1348	(0.00) 346.8
$\nu_{68}$	970	(11.86) [0.00]	$\nu_{100}$	1349	(12.02) [0.00]
$\nu_{69}$	987	(0.13) [0.00]	$\nu_{101}$	1366	(17.21) [0.00]
$\nu_{70}$	987	(0.00) [0.09]	$\nu_{102}$	1367	(0.00) [4.93]
$\nu_{71}$	989	(0.00) [0.37]	$\nu_{103}$	1388	(0.00) [31.6]
$\nu_{72}$	989	(1.63) [0.00]	$\nu_{104}$	1389	(2.37) [0.00]

<sup>a</sup> km/mole    <sup>b</sup>  $\text{A}^4/\text{AMU}$

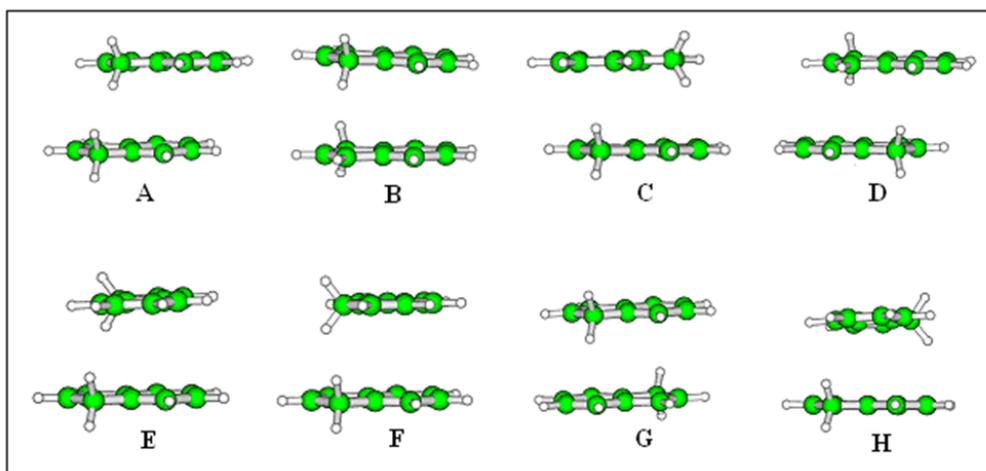


Figure C.11 – Configurations of Indene dimers optimized at the  $\omega$ B97X-D/6-311++G\*\* level of theory.

Table C.71 – Interaction energies of Indene dimers (kcal/mol) in 8 different configurations using  $\omega$ B97X-D/6-311++G\*\*, and compared with SAPT-DFT.

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>
<b><math>\omega</math>97X-D / 6-311++G**</b>	-7.06	-7.53	-8.20	-7.19	-7.47	-7.33	-8.97	-7.98
<b>SAPT- DFT(PBE0) aug-ccpvtz</b>	-5.11	-5.29	-6.05	-5.15	-5.27	-5.22	-6.76	-5.81

Table C.72 – Intermolecular distance between ciclopentadiene rings center of indene dimers (Å).

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>
<b><math>\omega</math>97X-D / 6-311++G**</b>	3.98	3.93	3.83	3.67	4.16	3.74	4.20	3.96
<b>SAPT- DFT(PBE0) aug-ccpvtz</b>	4.01	3.96	3.86	3.72	4.21	3.78	4.21	3.99

Table C.73 – Calculated low wavenumber Raman ad PA infrared spectra of Indene Dimer.

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_1$	28	(0.04) [0.00]	$\nu_7$	205	(4.5) [0.00]
$\nu_2$	44	(0.00) [8.00]	$\nu_8$	212	(0.00) [1.15]
$\nu_3$	62	(0.00) [4.39]	$\nu_9$	230	(0.00) [0.56]
$\nu_4$	71	(0.34) [0.00]	$\nu_{10}$	238	(7.39) [0.00]
$\nu_5$	86	(0.60) [0.00]	$\nu_{11}$	390	(0.00) [1.09]
$\nu_6$	116	(0.00) [12.40]	$\nu_{12}$	390	(2.78) [0.00]

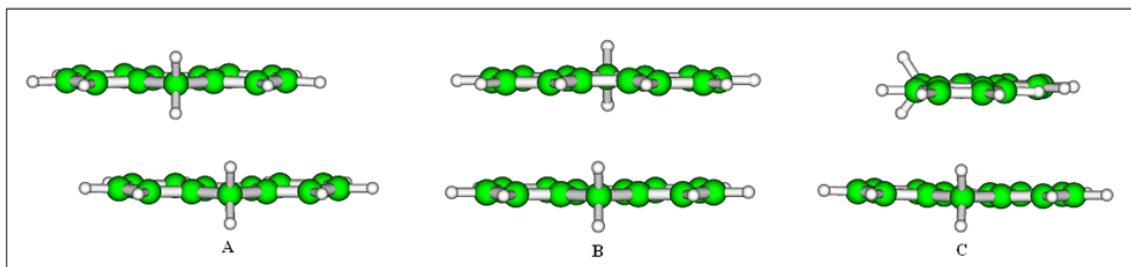
Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{13}$	400	(0.00) [0.95]	$\nu_{18}$	547	(0.00) [12.10]
$\nu_{14}$	402	(9.13) [0.00]	$\nu_{19}$	570	(11.82) [0.00]
$\nu_{15}$	432	(26.07) [0.00]	$\nu_{20}$	570	(0.00) [0.50]
$\nu_{16}$	434	(0.00) [0.07]	$\nu_{21}$	610	(2.32) [0.00]
$\nu_{17}$	546	(0.08) [0.00]	$\nu_{22}$	610	(0.00) [10.13]

<sup>a</sup> km/mole    <sup>b</sup> Å<sup>4</sup>/AMUTable C.74 – Calculated Raman and PA infrared spectra of Indene Dimer, 700–2000 cm<sup>-1</sup>

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{23}$	715	(0.00) [0.74]	$\nu_{47}$	1053	(0.00) [47.16]
$\nu_{24}$	716	(34.29) [0.00]	$\nu_{48}$	1053	(8.74) [0.00]
$\nu_{25}$	742	(62.20) [0.00]	$\nu_{49}$	1096	(0.00) [13.49]
$\nu_{26}$	743	(0.00) [0.84]	$\nu_{50}$	1096	(3.09) [0.00]
$\nu_{27}$	751	(0.00) [21.71]	$\nu_{51}$	1140	(1.96) [0.00]
$\nu_{28}$	752	(4.11) [0.00]	$\nu_{52}$	1141	(0.00) [23.12]
$\nu_{29}$	794	(146.91) [0.00]	$\nu_{53}$	1163	(4.90) [0.00]
$\nu_{30}$	795	(0.00) [1.79]	$\nu_{54}$	1164	(0.00) [11.78]
$\nu_{31}$	853	(0.00) [17.94]	$\nu_{55}$	1183	(0.16) [0.00]
$\nu_{32}$	853	(2.18) [0.002]	$\nu_{56}$	1184	(0.00) [5.36]
$\nu_{33}$	880	(0.00) [9.36]	$\nu_{57}$	1196	(0.00) [1.84]
$\nu_{34}$	880	(3.55) [0.00]	$\nu_{58}$	1197	(2.81) [0.00]
$\nu_{35}$	884	(0.00) [1.33]	$\nu_{59}$	1241	(0.00) [45.29]
$\nu_{36}$	885	(1.51) [0.00]	$\nu_{60}$	1242	(3.21) [0.00]
$\nu_{37}$	947	(16.20) [0.00]	$\nu_{61}$	1265	(5.49) [0.00]
$\nu_{38}$	947	(0.00) [1.05]	$\nu_{62}$	1266	(0.00) [24.77]
$\nu_{39}$	967	(0.00) [1.24]	$\nu_{63}$	1327	(0.00) [11.63]
$\nu_{40}$	969	(2.21) [0.00]	$\nu_{64}$	1327	(1.37) [0.00]
$\nu_{41}$	971	(0.00) [7.77]	$\nu_{65}$	1352	(0.00) [11.06]
$\nu_{42}$	971	(13.75) [0.00]	$\nu_{66}$	1353	(7.05) [0.00]
$\nu_{43}$	985	(3.54) [0.00]	$\nu_{67}$	1401	(4.87) [0.00]
$\nu_{44}$	986	(0.00) [5.74]	$\nu_{68}$	1401	(0.00) [42.12]
$\nu_{45}$	1007	(0.00) [0.20]	$\nu_{69}$	1439	(0.01) [24.51]
$\nu_{46}$	1008	(0.21) [0.00]	$\nu_{70}$	1439	(18.54) [0.01]

Table C.75 – Calculated Raman and PA infrared spectra of Indene Dimer, 700–2000 cm<sup>-1</sup>

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{71}$	1505	(10.12) [0.00]	$\nu_{76}$	1634	(0.00) [122.78]
$\nu_{72}$	1505	(0.00) [8.24]	$\nu_{77}$	1667	(0.00) [40.40]
$\nu_{73}$	1510	(0.00) [21.80]	$\nu_{78}$	1668	(0.86) [0.00]
$\nu_{74}$	1511	(23.19) [0.00]	$\nu_{79}$	1689	(5.18) [0.00]
$\nu_{75}$	1634	(0.66) [0.00]	$\nu_{80}$	1689	(0.00) [102.66]

<sup>a</sup> km/mole<sup>b</sup> A<sup>4</sup>/AMUFigure C.12 – Configurations of Fluorene dimers optimized at the  $\omega$ B97x-D/6-311++G\*\* level of theory.Table C.76 – Interaction energies of Fluorene dimers (kcal/mol) in 3 different configurations using  $\omega$ B97x-D/6-311++G\*\*.

	A	B	C
$\omega$ B97x-D / 6-311++G**	-10.82	-11.74	-13.61

Table C.77 – Calculated low wavenumber Raman ad PA infrared spectra of Fluorene Dimer.

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_1$	27	(0.01) [2.45]	$\nu_{18}$	423	(0.15) [10.09]
$\nu_2$	40	(0.04) [7.86]	$\nu_{19}$	426	(8.62) [2.33]
$\nu_3$	45	(0.01) [7.26]	$\nu_{20}$	427	(0.02) [2.13]
$\nu_4$	76	(0.002) [1.94]	$\nu_{21}$	443	(0.42) [1.17]
$\nu_5$	79	(0.05) [1.55]	$\nu_{22}$	444	(0.37) [0.84]
$\nu_6$	93	(0.02) [3.82]	$\nu_{23}$	484	(0.001) [1.08]
$\nu_7$	111	(0.74) [0.36]	$\nu_{24}$	486	(3.56) [0.10]
$\nu_8$	144	(0.07) [0.75]	$\nu_{25}$	502	(0.21) [0.05]
$\nu_9$	144	(0.03) [1.60]	$\nu_{26}$	503	(0.32) [0.35]
$\nu_{10}$	158	(0.01) [0.48]	$\nu_{27}$	558	(0.16) [0.75]
$\nu_{11}$	218	(0.20) [0.49]	$\nu_{28}$	558	(0.27) [12.06]
$\nu_{12}$	219	(0.13) [0.58]	$\nu_{29}$	582	(0.01) [0.10]
$\nu_{13}$	266	(0.02) [1.06]	$\nu_{30}$	584	(0.08) [0.31]
$\nu_{14}$	272	(15.68) [1.55]	$\nu_{31}$	640	(3.47) [0.06]
$\nu_{15}$	292	(0.14) [4.88]	$\nu_{32}$	640	(5.65) [0.26]
$\nu_{16}$	292	(4.08) [3.92]	$\nu_{33}$	649	(0.24) [0.01]
$\nu_{17}$	422	(9.91) [1.50]	$\nu_{34}$	650	(0.12) [0.10]

<sup>a</sup> km/mole    <sup>b</sup> Å<sup>4</sup>/AMUTable C.78 – Raman ad PA infrared spectra of Fluorene Dimer, 700- 2000 cm<sup>-1</sup>

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{35}$	718	(0.03) [0.58]	$\nu_{47}$	863	(0.10) [5.61]
$\nu_{36}$	718	(18.08) [0.06]	$\nu_{48}$	863	(0.10) [30.70]
$\nu_{37}$	747	(15.67) [0.41]	$\nu_{49}$	880	(2.27) [0.52]
$\nu_{38}$	750	(1.89) [0.50]	$\nu_{50}$	881	(5.50) [0.30]
$\nu_{39}$	761	(275.51) [0.08]	$\nu_{51}$	891	(0.26) [0.002]
$\nu_{40}$	765	(1.77) [1.61]	$\nu_{53}$	943	(0.01) [1.26]
$\nu_{41}$	765	(0.01) [25.21]	$\nu_{54}$	943	(0.41) [0.11]
$\nu_{42}$	765	(0.16) [13.68]	$\nu_{55}$	970	(0.26) [0.02]
$\nu_{43}$	806	(1.33) [2.69]	$\nu_{56}$	971	(0.15) [0.08]
$\nu_{44}$	807	(2.27) [2.36]	$\nu_{57}$	978	(13.33) [0.21]
$\nu_{45}$	826	(0.04) [0.10]	$\nu_{58}$	989	(0.01) [0.98]
$\nu_{46}$	826	(0.02) [0.35]	$\nu_{59}$	1006	(0.21) [0.32]

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_{60}$	1008	(0.48) [0.24]	$\nu_{87}$	1331	(0.52) [31.59]
$\nu_{61}$	1008	(0.28) [0.12]	$\nu_{88}$	1331	(0.03) [117.03]
$\nu_{62}$	1009	(0.02) [0.18]	$\nu_{89}$	1337	(1.89) [0.05]
$\nu_{63}$	1034	(1.63) [0.91]	$\nu_{90}$	1337	(2.26) [1.26]
$\nu_{64}$	1034	(2.57) [1.62]	$\nu_{91}$	1357	(7.71) [1.06]
$\nu_{65}$	1057	(0.02) [94.68]	$\nu_{92}$	1358	(5.36) [0.002]
$\nu_{66}$	1058	(1.58) [4.53]	$\nu_{93}$	1381	(0.41) [29.60]
$\nu_{67}$	1061	(4.39) [0.36]	$\nu_{94}$	1381	(0.09) [51.06]
$\nu_{68}$	1061	(7.88) [2.55]	$\nu_{95}$	1452	(5.88) [29.20]
$\nu_{69}$	1128	(4.32) [1.05]	$\nu_{96}$	1452	(12.93) [1.19]
$\nu_{70}$	1129	(2.71) [6.44]	$\nu_{97}$	1495	(7.54) [1.19]
$\nu_{71}$	1143	(0.01) [0.61]	$\nu_{98}$	1495	(10.03) [4.70]
$\nu_{72}$	1143	(0.02) [0.03]	$\nu_{99}$	1498	(20.07) [0.61]
$\nu_{73}$	1177	(0.44) [3.33]	$\nu_{100}$	1499	(25.63) [1.26]
$\nu_{74}$	1178	(0.69) [4.58]	$\nu_{101}$	1528	(5.12) [1.65]
$\nu_{75}$	1182	(0.20) [9.65]	$\nu_{102}$	1529	(8.30) [9.70]
$\nu_{76}$	1182	(0.47) [0.88]	$\nu_{103}$	1534	(0.02) [66.91]
$\nu_{77}$	1186	(0.002) [16.57]	$\nu_{104}$	1534	(1.26) [43.91]
$\nu_{78}$	1186	(0.10) [0.03]	$\nu_{105}$	1652	(0.33) [37.41]
$\nu_{79}$	1202	(1.44) [1.07]	$\nu_{106}$	1652	(0.99) [14.28]
$\nu_{80}$	1202	(0.95) [0.40]	$\nu_{107}$	1660	(0.98) [0.80]
$\nu_{81}$	1223	(2.59) [8.74]	$\nu_{108}$	1660	(1.38) [1.11]
$\nu_{82}$	1223	(5.15) [0.89]	$\nu_{109}$	1686	(1.26) [11.36]
$\nu_{83}$	1238	(2.90) [1.33]	$\nu_{110}$	1687	(2.21) [39.23]
$\nu_{84}$	1238	(4.86) [18.36]	$\nu_{111}$	1692	(0.01) [209.97]
$\nu_{85}$	1269	(2.36) [36.28]	$\nu_{112}$	1693	(0.07) [243.77]
$\nu_{86}$	1270	(0.57) [108.92]			

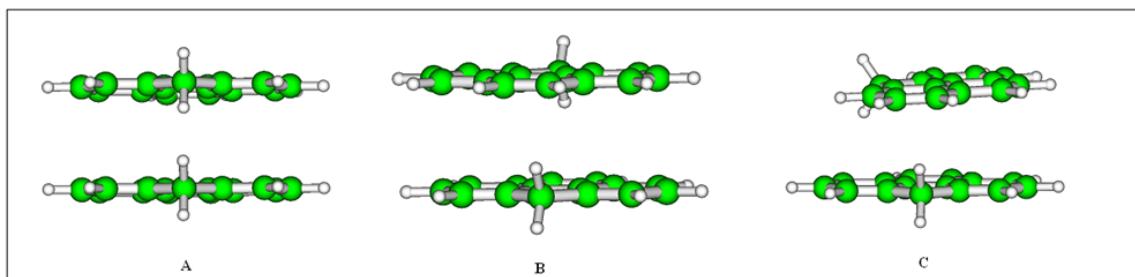
<sup>a</sup> km/mole<sup>b</sup> Å<sup>4</sup>/AMU

Figure C.13 – Configurations of Cyclopentaphenanthrene dimers optimized at the  $\omega$ B97x-D/6-311++G\*\* level of theory.

Table C.79 – Interaction energies of Cyclopentaphenanthrene dimers (kcal/mol) in 3 different configurations using  $\omega$ B97x-D/6-311++G\*\*.

	A	B	C
$\omega$ B97x-D/ 6-311++G**	-13.20	-14.61	-15.21

Table C.80 – Calculated low wavenumber Raman ad PA infrared spectra of Cyclopentaphenanthrene Dimer.

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_1$	29	(0.00) [1.77]	$\nu_{21}$	447	(0.02) [16.59]
$\nu_2$	34	(0.07) [10.73]	$\nu_{22}$	447	(0.03) [18.27]
$\nu_3$	45	(0.00) [11.47]	$\nu_{23}$	478	(0.04) [0.66]
$\nu_4$	77	(0.01) [0.48]	$\nu_{24}$	478	(1.20) [0.15]
$\nu_5$	88	(0.00) [2.91]	$\nu_{25}$	495	(0.88) [0.51]
$\nu_6$	90	(0.32) [0.13]	$\nu_{26}$	495	(1.02) [3.95]
$\nu_7$	110	(4.19) [0.90]	$\nu_{27}$	508	(6.63) [0.01]
$\nu_8$	116	(0.02) [0.92]	$\nu_{28}$	510	(0.24) [0.36]
$\nu_9$	194	(0.08) [1.62]	$\nu_{29}$	534	(0.28) [0.14]
$\nu_{10}$	195	(0.01) [0.95]	$\nu_{30}$	534	(0.77) [0.41]
$\nu_{11}$	206	(25.22) [0.75]	$\nu_{31}$	560	(0.01) [0.44]
$\nu_{12}$	210	(0.00) [3.23]	$\nu_{32}$	561	(0.01) [3.40]
$\nu_{13}$	257	(4.41) [0.57]	$\nu_{33}$	579	(0.02) [0.03]
$\nu_{14}$	264	(0.00) [1.78]	$\nu_{34}$	579	(0.03) [0.11]
$\nu_{15}$	294	(0.34) [1.66]	$\nu_{35}$	604	(2.33) [0.09]
$\nu_{16}$	295	(0.18) [1.36]	$\nu_{36}$	605	(0.80) [32.26]
$\nu_{17}$	344	(1.74) [0.50]	$\nu_{37}$	632	(0.02) [0.05]
$\nu_{18}$	344	(0.82) [0.11]	$\nu_{38}$	635	(0.10) [0.03]
$\nu_{19}$	437	(0.04) [3.63]	$\nu_{39}$	672	(0.47) [0.97]
$\nu_{20}$	438	(0.12) [3.09]	$\nu_{40}$	672	(0.28) [40.10]

<sup>a</sup> km/mole<sup>b</sup> A<sup>4</sup>/AMU

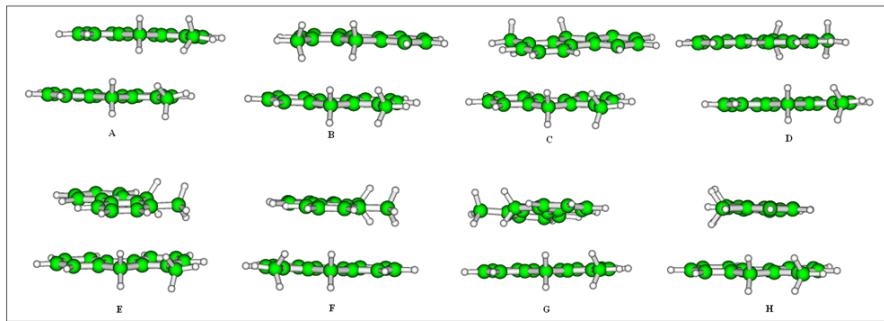


Figure C.14 – Configurations of 1-methylfluorene dimers optimized at the  $\omega$ B97X-D/6-311++G\*\* level of theory.

Table C.81 – Interaction energies of 1-methylfluorene dimers (kcal/mol) in 8 different configurations using  $\omega$ B97X-D/6-311++G\*\*.

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>
$\omega$ <b>B97X-D / 6-311++G**</b>	-12.13	-12.92	-15.78	-14.95	-15.94	-15.45	-14.80	-15.01

Table C.82 – Calculated low wavenumber Raman ad PA infrared spectra of 1-methylfluorene Dimer.

<b>Mode</b>	$\omega$ $\text{cm}^{-1}$	<b>Intensities</b> <sup>(a)</sup> [ <sup>b</sup> ]	<b>Mode</b>	$\omega$ $\text{cm}^{-1}$	<b>Intensities</b> <sup>(a)</sup> [ <sup>b</sup> ]
$\nu_1$	22	(0.02) [1.74]	$\nu_{21}$	305	(0.02) [4.24]
$\nu_2$	40	(0.00) [5.00]	$\nu_{22}$	307	(0.06) [0.87]
$\nu_3$	50	(0.00) [5.37]	$\nu_{23}$	433	(1.65) [0.48]
$\nu_4$	52	(0.02) [3.48]	$\nu_{24}$	435	(6.42) [0.21]
$\nu_5$	69	(0.06) [1.53]	$\nu_{25}$	439	(0.66) [2.04]
$\nu_6$	87	(0.71) [1.80]	$\nu_{26}$	440	(0.18) [5.68]
$\nu_7$	93	(0.02) [0.73]	$\nu_{27}$	482	(3.08) [0.39]
$\nu_8$	114	(1.56) [0.55]	$\nu_{28}$	486	(0.27) [0.34]
$\nu_9$	128	(0.03) [3.80]	$\nu_{29}$	508	(0.04) [6.34]
$\nu_{10}$	131	(0.12) [1.77]	$\nu_{30}$	508	(0.12) [4.90]
$\nu_{11}$	138	(0.05) [0.45]	$\nu_{31}$	518	(1.16) [0.46]
$\nu_{12}$	187	(0.25) [0.31]	$\nu_{32}$	522	(0.48) [0.76]
$\nu_{13}$	189	(0.07) [0.45]	$\nu_{33}$	550	(1.49) [1.74]
$\nu_{14}$	213	(0.09) [1.67]	$\nu_{34}$	551	(0.99) [2.09]
$\nu_{15}$	225	(0.06) [0.87]	$\nu_{35}$	589	(2.08) [3.48]
$\nu_{16}$	245	(0.06) [0.61]	$\nu_{36}$	590	(1.30) [4.81]
$\nu_{17}$	260	(6.30) [1.37]	$\nu_{37}$	593	(2.94) [0.49]
$\nu_{18}$	277	(14.65) [1.88]	$\nu_{38}$	598	(1.25) [0.61]

<sup>a</sup> km/mole

<sup>b</sup>  $\text{A}^4/\text{AMU}$

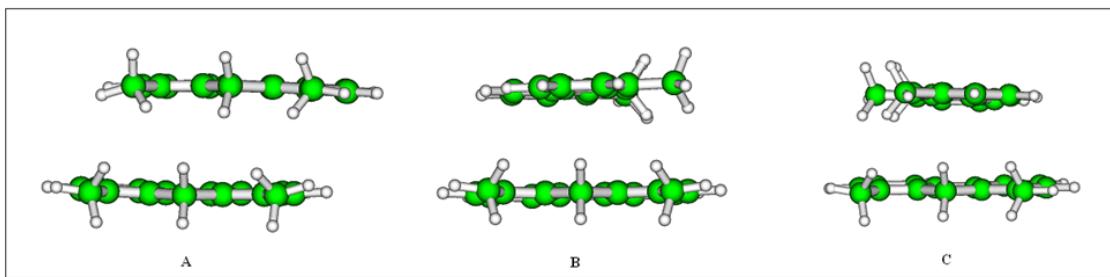


Figure C.15 – Configurations of 1,8-dimethylfluorene dimers optimized at the  $\omega$ B97x-D/6-311++G\*\* level of theory.

Table C.83 – Interaction energies of 18-dimethylfluorene dimers (kcal/mol) in 3 different configurations using  $\omega$ B97x-D/6-311++G\*\*.

	A	B	C
$\omega$ B97x-D/ 6-311++G**	-14.12	-15.90	-16.29

Table C.84 – Calculated low wavenumber Raman ad PA infrared spectra of 18-dimethylfluorene Dimer.

Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]	Mode	$\omega$ cm <sup>-1</sup>	Intensities ( <sup>a</sup> ) [ <sup>b</sup> ]
$\nu_1$	22	(0.02) [1.74]	$\nu_{21}$	305	(0.02) [4.24]
$\nu_2$	40	(0.00) [5.00]	$\nu_{22}$	307	(0.06) [0.87]
$\nu_3$	50	(0.00) [5.37]	$\nu_{23}$	433	(1.65) [0.48]
$\nu_4$	52	(0.02) [3.48]	$\nu_{24}$	435	(6.42) [0.21]
$\nu_5$	69	(0.06) [1.53]	$\nu_{25}$	439	(0.66) [2.04]
$\nu_6$	87	(0.71) [1.80]	$\nu_{26}$	440	(0.18) [5.68]
$\nu_7$	93	(0.02) [0.73]	$\nu_{27}$	482	(3.08) [0.39]
$\nu_8$	114	(1.56) [0.55]	$\nu_{28}$	486	(0.27) [0.34]
$\nu_9$	128	(0.03) [3.80]	$\nu_{29}$	508	(0.04) [6.34]
$\nu_{10}$	131	(0.12) [1.77]	$\nu_{30}$	508	(0.12) [4.90]
$\nu_{11}$	138	(0.05) [0.45]	$\nu_{31}$	518	(1.16) [0.46]
$\nu_{12}$	187	(0.25) [0.31]	$\nu_{32}$	522	(0.48) [0.76]
$\nu_{13}$	189	(0.07) [0.45]	$\nu_{33}$	550	(1.49) [1.74]
$\nu_{14}$	213	(0.09) [1.67]	$\nu_{34}$	551	(0.99) [2.09]
$\nu_{15}$	225	(0.06) [0.87]	$\nu_{35}$	589	(2.08) [3.48]

<sup>a</sup> km/mole

<sup>b</sup> A<sup>4</sup>/AMU

# Références bibliographiques

- [1] J G Speight and S. 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.
- [2] J. A Calles, J Dufour, J Marugán, J. L Peña, R Giménez-Aguirre, and D 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.
- [3] L Goual. *Petroleum asphaltenes*. INTECH Open Access Publisher, 2012.
- [4] J.G Speight. Petroleum asphaltenes-part 1: Asphaltenes, resins and the structure of petroleum. *Oil & gas science and technology*, 59(5):467–477, 2004.
- [5] I Merdignac 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.
- [6] J Eyssautier, P Levitz, D Espinat, J Jestin, J Gummel, I Grillo, and L 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.
- [7] O. C Mullins, B Martínez-Haya, and A. 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.
- [8] H Groenzin and O. C Mullins. Molecular size and structure of asphaltenes from various sources. *Energy & Fuels*, 14(3):677–684, 2000.
- [9] S Zhao, L.S Kotlyar, J.R Woods, B.D Sparks, K Hardacre, and K.H Chung. Molecular transformation of athabasca bitumen end-cuts during coking and hydrocracking. *Fuel*, 80(8):1155–1163, 2001.
- [10] E Rogel and L Carbognani. Density estimation of asphaltenes using molecular dynamics simulations. *Energy & fuels*, 17(2):378–386, 2003.
- [11] K. L Gawrys, P Matthew Spiecker, and P. K Kilpatrick. The role of asphaltene solubility and chemical composition on asphaltene aggregation. *Petroleum science and technology*, 21(3-4):461–489, 2003.
- [12] J. M Sheremata, M. R Gray, H. D Dettman, and W. C McCaffrey. Quantitative molecular representation and sequential optimization of Athabasca asphaltenes. *Energy & Fuels*, 18(5):1377–1384, 2004.

- [13] P. D Murphy, B. C Gerstein, V. L Weinberg, and T. F 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.
- [14] T. F Yen, W. H Wu, and G. V Chilingar. A study of the structure of petroleum asphaltenes and related substances by infrared spectroscopy. *Energy Sources*, 7(3):203–235, 1984.
- [15] E Durand, M Clemancey, J.M Lancelin, J Verstraete, D Espinat, and A Quoineaud. Effect of chemical composition on asphaltenes aggregation. *Energy Fuels*, 24(2):1051–1062, 2010.
- [16] T. F Yen, J. G Erdman, and S. S Pollack. Investigation of the structure of petroleum asphaltenes by X-ray diffraction. *Analytical Chemistry*, 33(11):1587–1594, 1961.
- [17] J. W Shirokoff, M. N Siddiqui, and M. F Ali. Characterization of the structure of saudi crude asphaltenes by X-ray diffraction. *Energy & Fuels*, 11(3):561–565, 1997.
- [18] K.H Altgelt and M.M Boduszynski. Composition and analysis of heavy petroleum fractions. 1994.
- [19] H Groenzin and O. C Mullins. Asphaltene molecular size and structure. *The Journal of Physical Chemistry A*, 103(50):11237–11245, 1999.
- [20] R Souza, D. E Nicodem, S. J Garden, and R. J Corrêa. Study of the asphaltene aggregation structure by time-resolved fluorescence spectroscopy. *Energy & Fuels*, 24(2):1135–1138, 2009.
- [21] Y Hongfu, C Xiaoli, L Haoran, and X Yupeng. Determination of multi-properties of residual oils using mid-infrared attenuated total reflection spectroscopy. *Fuel*, 85(12):1720–1728, 2006.
- [22] A.A El-Bassoussi, M.H Ahmed, S. M Sayed, J. S Basta, and E. S Attia. Characterization of some local petroleum residues by spectroscopic techniques. *Petroleum Science and Technology*, 28(5):430–444, 2010.
- [23] R. R Coelho, I Hovell, E. L Moreno, A 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.
- [24] J Laxalde, N Caillol, F Wahl, C Ruckebusch, and L Duponchel. Combining near and mid infrared spectroscopy for heavy oil characterisation. *Fuel*, 133:310–316, 2014.
- [25] S Acevedo, O Escobar, L Echevarria, L. B Gutiérrez, and B 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.
- [26] L. B Gutiérrez, M. A Ranaudo, B Méndez, and S 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.
- [27] J Murgich, J Rodríguez, and Y Aray. Molecular recognition and molecular mechanics of micelles of some model asphaltenes and resins. *Energy & Fuels*, 10(1):68–76, 1996.

- [28] F Alvarez-Ramírez and Y Ruiz-Morales. Island versus archipelago architecture for asphaltenes: polycyclic aromatic hydrocarbon dimer theoretical studies. *Energy & Fuels*, 27(4):1791–1808, 2013.
- [29] T Takanohashi, S Sato, and R Tanaka. Structural relaxation behaviors of three different asphaltenes using MD calculations. *Petroleum science and technology*, 22(7-8):901–914, 2004.
- [30] M.E Elyashberg, A. J Williams, and G. E Martin. Computer-assisted structure verification and elucidation tools in NMR-based structure elucidation. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 53(1):1–104, 2008.
- [31] 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.
- [32] J. L 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.
- [33] H. K Al Halwachi, D. S Yakovlev, and E. S Boek. Systematic optimization of asphaltene molecular structure and molecular weight using the quantitative molecular representation approach. *Energy & Fuels*, 26(10):6177–6185, 2012.
- [34] L. P de Oliveira, J Verstraete, and M Kolb. A monte carlo modeling methodology for the simulation of hydrotreating processes. *Chemical engineering journal*, 207:94–102, 2012.
- [35] P. J Krusic, E Wasserman, B. A Parkinson, B Malone, Jr E. R Holler, P. N Keizer, J. R Morton, and K. F Preston. Electron spin resonance study of the radical reactivity of C<sub>60</sub>. *Journal of the American Chemical Society*, 113(16):6274–6275, 1991.
- [36] P Barchewitz. *Spectroscopie atomique et moléculaire: Spectroscopie moléculaire*, volume 2. Masson, 1971.
- [37] J. G Wilson Jr, E. Band Decius, P. G Cross, and R. T Lagemann. Molecular vibrations. *American Journal of Physics*, 23(8):550–550, 1955.
- [38] C 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.
- [39] K Zaki. *Etude des modes de vibration au-delà de l'approximation harmonique: Application à la résolution des problèmes vibrationnels de molécules de type X<sub>2</sub>CNX(X= F, H)*. PhD thesis, 1996.
- [40] C Eckart. Some studies concerning rotating axes and polyatomic molecules. *Physical Review*, 47(7):552, 1935.
- [41] H Casimir. Rotation of a rigid body in quantum mechanics. 1931.
- [42] P Pulay, G Fogarasi, F Pang, and J. 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.

- [43] G Simons, R. G Parr, and J. M Finlan. New alternative to the dunham potential for diatomic molecules. *The Journal of Chemical Physics*, 59(6):3229–3234, 1973.
- [44] W Meyer, P Botschwina, and P Burton. Ab initio calculation of near-equilibrium potential and multipole moment surfaces and vibrational frequencies of  $\text{H}^{+3}$  and its isotopomers. *The Journal of chemical physics*, 84(2):891–900, 1986.
- [45] P. E Maslen, D Jayatilaka, S. M Colwell, R. D Amos, and N. C Handy. Higher analytic derivatives. II. the fourth derivative of self-consistent-field energy. *The Journal of chemical physics*, 95(10):7409–7417, 1991.
- [46] A. R Hoy, I. M Mills, and G Strey. Anharmonic force constant calculations. *Molecular Physics*, 24(6):1265–1290, 1972.
- [47] H. H Nielsen. The vibration-rotation energies of molecules. *Reviews of Modern Physics*, 23(2):90, 1951.
- [48] P Pulay. Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules: I. theory. *Molecular Physics*, 17(2):197–204, 1969.
- [49] F Spillebout, D Begue, I Baraille, and J. M Shaw. On discerning intermolecular and intramolecular vibrations in experimental acene spectra. *Energy & Fuels*, 28(5):2933–2947, 2014.
- [50] G Fradet. Master 1, 2016.
- [51] P Barchewitz. *Spectroscopie infrarouge: Vibrations moléculaires*, volume 1. Gauthier-Villars, 1961.
- [52] P Carbonnière. *Calcul ab-initio de spectres vibrationnels anharmoniques.: Développement méthodologique pour le traitement de systèmes de grande taille*. PhD thesis, Pau, 2002.
- [53] D Papousek and M. R Aliev. Molecular vibrational-rotational spectra. 1982.
- [54] J. H Van Vleck. On  $\sigma$ -type doubling and electron spin in the spectra of diatomic molecules. *Physical Review*, 33(4):467, 1929.
- [55] T 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.
- [56] M. J Frisch, G. W Trucks, H. B Schlegel, G. E Scuseria, M. A Robb, J. R Cheeseman, G Scalmani, V Barone, B Mennucci, G. A Petersson, et al. Gaussian 09, revision a. 02; gaussian, inc: Wallingford, CT, 2009. *There is no corresponding record for this reference*, 2015.
- [57] E Fermi. Über den ramaneffekt des kohlendioxyds. *Zeitschrift für Physik*, 71(3-4):250–259, 1931.
- [58] B. T Darling and D. M Dennison. The water vapor molecule. *Physical Review*, 57(2):128, 1940.

- [59] J. M Martin and P. 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.
- [60] J. O Jung and R. B Gerber. Vibrational wave functions and spectroscopy of  $(\text{H}_2\text{O})_n$ , n= 2, 3, 4, 5: Vibrational self-consistent field with correlation corrections. *The Journal of chemical physics*, 105(23):10332–10348, 1996.
- [61] L. S Norris, M. A Ratner, A. E Roitberg, and R. B Gerber. Moller–plesset perturbation theory applied to vibrational problems. *The Journal of chemical physics*, 105(24):11261–11267, 1996.
- [62] C Pouchan, K Zaki, et al. Ab-initio configuration-interaction determination of the overtone vibrations of methyleneimine in the region 2800–3200  $\text{cm}^{-1}$ . *Journal of Chemical Physics*, 107(2):342–345, 1997.
- [63] S Carter, S. J Culik, and J. 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.
- [64] J Koput, S Carter, N. 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.
- [65] P Cassam-Chenaï and J Liévin. Alternative perturbation method for the molecular vibration–rotation problem. *International journal of quantum chemistry*, 93(3):245–264, 2003.
- [66] N Gohaud, D Begue, C Darrigan, and C Pouchan. New parallel software (P\_Anhar) for anharmonic vibrational calculations: Application to  $(\text{CH}_3\text{Li})_2$ . *Journal of computational chemistry*, 26(7):743–754, 2005.
- [67] 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.
- [68] O Coulaud, M Odulami, V LeBris, D Bégué, and I Baraille. Notes PyRAVCI : Python impélementation of the A-VCI algorithm. Technical report, INRIA and IPREM, 2016.
- [69] R. E Wyatt, C Iung, and C Leforestier. Toward ab initio intramolecular dynamics. *Accounts of chemical research*, 28(10):423–429, 1995.
- [70] D Bégué, N Gohaud, C Pouchan, P Cassam-Chenaï, and J 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.
- [71] R Garnier, M Odunlami, V Le Bris, D Bégué, I Baraille, and O 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.
- [72] J. M Bowman. Self-consistent field energies and wavefunctions for coupled oscillators. *The Journal of Chemical Physics*, 68(2):608–610, 1978.

- [73] J. M Bowman. The self-consistent-field approach to polyatomic vibrations. *Accounts of Chemical Research*, 19(7):202–208, 1986.
- [74] R. B Gerber and M. 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.
- [75] R. B Gerber and M. A Ratner. A semiclassical self-consistent field (SC SCF) approximation for eigenvalues of coupled-vibration systems. *Chemical Physics Letters*, 68(1):195–198, 1979.
- [76] S Carter, N. C Handy, and J. M Bowman. High torsional vibrational energies of H<sub>2</sub>O<sub>2</sub> and CH<sub>3</sub>OH studied by multimode with a large amplitude motion coupled to two effective contraction schemes. *Molecular Physics*, 107(8-12):727–737, 2009.
- [77] Ml Sparta, M.l B Hansen, E Matito, D Toffoli, and O 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.
- [78] K Yagi, S Hirata, and K Hirao. Vibrational quasi-degenerate perturbation theory: applications to fermi resonance in CO<sub>2</sub>, H<sub>2</sub>CO, and C<sub>6</sub>H<sub>6</sub>. *Physical Chemistry Chemical Physics*, 10(13):1781–1788, 2008.
- [79] B Huron, J. P 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.
- [80] I Gerber. *Description des forces de van der Waals dans le cadre de la théorie de la fonctionnelle de la densité par un traitement explicite des interactions de longue portée*. PhD thesis, Université Henri Poincaré-Nancy I, 2005.
- [81] R. G Gordon and Y. S Kim. Theory for the forces between closed-shell atoms and molecules. *The Journal of Chemical Physics*, 56(6):3122–3133, 1972.
- [82] R Valero and D. G Truhlar. Nonadiabatic effects in c–br bond scission in the photo-dissociation of bromoacetyl chloride. *The Journal of chemical physics*, 125(19):194305, 2006.
- [83] S Grimme. Semiempirical gga-type density functional constructed with a long-range dispersion correction. *Journal of computational chemistry*, 27(15):1787–1799, 2006.
- [84] S Grimme, J Antony, S Ehrlich, and H Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of chemical physics*, 132(15):154104, 2010.
- [85] J Klimeš and A Michaelides. Perspective: Advances and challenges in treating van der waals dispersion forces in density functional theory. *The Journal of chemical physics*, 137(12):120901, 2012.
- [86] F Tran and J Hutter. Nonlocal van der waals functionals: The case of rare-gas dimers and solids. *The Journal of chemical physics*, 138(20):204103, 2013.

- [87] M Dion, H Rydberg, E Schröder, D C Langreth, and B. I Lundqvist. Van der waals density functional for general geometries. *Physical review letters*, 92(24):246401, 2004.
- [88] J Wellendorff, K. T Lundgaard, A Møgelhøj, V Petzold, D. D Landis, J. K Nørskov, T Bligaard, and K. W Jacobsen. Density functionals for surface science: Exchange-correlation model development with bayesian error estimation. *Physical Review B*, 85(23):235149, 2012.
- [89] S Lebegue, J Harl, T Gould, J. G Ángyán, G Kresse, and J. F Dobson. Cohesive properties and asymptotics of the dispersion interaction in graphite by the random phase approximation. *Physical review letters*, 105(19):196401, 2010.
- [90] B Mussard. *Modélisations quantochimiques des forces de dispersion de London par la méthode des phases aléatoires (RPA): développements méthodologiques*. PhD thesis, Université de Lorraine-Collegium Sciences et Technologies, 2013.
- [91] S Grimme. Do special noncovalent  $\pi$ - $\pi$  stacking interactions really exist? *Angewandte Chemie International Edition*, 47(18):3430–3434, 2008.
- [92] C. R Martinez and B. L Iverson. Rethinking the term “pi-stacking”. *Chemical Science*, 3(7):2191–2201, 2012.
- [93] G. B McGaughey, M Gagné, and A. K Rappé.  $\pi$ -stacking interactions alive and well in proteins. *Journal of Biological Chemistry*, 273(25):15458–15463, 1998.
- [94] A. D Buckingham. Permanent and induced molecular moments and long-range intermolecular forces. *Advances in Chemical Physics: Intermolecular Forces, Volume 12*, pages 107–142, 1967.
- [95] C. D Sherrill. Counterpoise correction and basis set superposition error. *School of Chemistry and Biochemistry, Georgia Institute of Technology*, 2010.
- [96] S. F Boys and F Bernardi. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics*, 19(4):553–566, 1970.
- [97] J Garza, J Ramírez, and R Vargas. Role of hartree-fock and kohn-sham orbitals in the basis set superposition error for systems linked by hydrogen bonds. *The Journal of Physical Chemistry A*, 109(4):643–651, 2005.
- [98] S. C Wang. The mutual influence between hydrogen atoms. *Phys Z*, 28:663–666, 1927.
- [99] F London. Z phys chem b 1930, 11, 222; eisenschitz, r.; london, f. *Z Phys*, 60:491, 1930.
- [100] K 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.
- [101] W. J Stevens and W. 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.

- [102] D Begue, M Mérawa, M Rerat, and C Pouchan. Dynamic dipole polarizabilities and  $c_6$  dispersion coefficients for small clusters of beryllium  $\text{Be}_n$  ( $n = 2, 3, 4$ ). *Chemical physics letters*, 301(1):43–52, 1999.
- [103] P. R. Fontana. Theory of long-range interatomic forces. I. dispersion energies between unexcited atoms. *Physical Review*, 123(5):1865, 1961.
- [104] P. R. Fontana. Theory of long-range interatomic forces. III. First-Order Interaction Energies in the Coupled Representation. *Physical Review*, 125(5):1597, 1962.
- [105] A. D. Buckingham, H. F. Knaap, L. J. Hermans, R. M. Jonkman, J. J. Beenakker, W. B. Brown, N. R. Kestner, W. D. Davison, P. Claverie, and L. Salem. General discussion. *Discussions of the Faraday Society*, 40:171–176, 1965.
- [106] B. C. Hall. *Lie groups, Lie algebras, and representations: an elementary introduction*, volume 222. Springer, 2015.
- [107] R. J. Buehler and J. O. Hirschfelder. Bipolar expansion of coulombic potentials. *Physical Review*, 83(3):628, 1951.
- [108] E. A. Hylleraas. über die elektronenterme des wasserstoffmoleküls. *Zeitschrift für Physik*, 71(11-12):739–763, 1931.
- [109] T. R. Proctor and W. C. Stwalley. The long-range interactions of S-state alkali atoms with rare gas and hydrogen atoms. *The Journal of Chemical Physics*, 66(5):2063–2073, 1977.
- [110] W. D. Davison. Atomic dipole-quadrupole dispersion forces. *Journal of Physics B: Atomic and Molecular Physics*, 1(2):139, 1968.
- [111] M. Saute and M. Aubert-Frécon. Calculated long-range potential-energy curves for the 23 molecular states of  $\text{I}_2$ . *The Journal of Chemical Physics*, 77(11):5639–5646, 1982.
- [112] D. M. Chipman, J. D. Bowman, and J. O. Hirschfelder. Perturbation theories for the calculation of molecular interaction energies. I. general formalism. *The Journal of Chemical Physics*, 59(6):2830–2837, 1973.
- [113] K. Patkowski, B. Jeziorski, and K. Szalewicz. Unified treatment of chemical and van der waals forces via symmetry-adapted perturbation expansion. *The Journal of chemical physics*, 120(15):6849–6862, 2004.
- [114] B. Jeziorski, R. Moszynski, and K. Szalewicz. Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chemical Reviews*, 94(7):1887–1930, 1994.
- [115] A. J. Misquitta, R. Podeszwa, B. Jeziorski, and K. Szalewicz. Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional calculations. *The Journal of chemical physics*, 123(21):214103, 2005.
- [116] A. Heßelmann and G. Jansen. First-order intermolecular interaction energies from kohn–sham orbitals. *Chemical physics letters*, 357(5):464–470, 2002.

- [117] H L Williams and C. 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.
- [118] H Krause, B Ernstberger, and H. J Neusser. Binding energies of small benzene clusters. *Chemical physics letters*, 184(5-6):411–417, 1991.
- [119] J. R Grover, E. A Walters, and E. T Hui. Dissociation energies of the benzene dimer and dimer cation. *Journal of Physical Chemistry*, 91(12):3233–3237, 1987.
- [120] C A Hunter, J Singh, and J. M Thornton.  $\pi$ - $\pi$  interactions: The geometry and energetics of phenylalanine-phenylalanine interactions in proteins. *Journal of molecular biology*, 218(4):837–846, 1991.
- [121] M. C Fyfe and J. F Stoddart. Synthetic supramolecular chemistry. *Accounts of chemical research*, 30(10):393–401, 1997.
- [122] J Rebek. Assembly and encapsulation with self-complementary molecules. *Chemical Society Reviews*, 25(4):255–264, 1996.
- [123] S. K Burley and G. A Petsko. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science*, 229(4708):23–28, 1985.
- [124] K. C Janda, J. C Hemminger, J. S Winn, S. E Novick, S. J Harris, and W Klemperer. Benzene dimer: a polar molecule. *The Journal of Chemical Physics*, 63(4):1419–1421, 1975.
- [125] E Arunan and H. S Gutowsky. The rotational spectrum, structure and dynamics of a benzene dimer. *The Journal of chemical physics*, 98(5):4294–4296, 1993.
- [126] B. F Henson, G. V Hartland, V. A Venturo, and P. M Felker. Raman-vibronic double-resonance spectroscopy of benzene dimer isotopomers. *The Journal of chemical physics*, 97(4):2189–2208, 1992.
- [127] Y Park and J Lee. Accurate ab initio binding energies of the benzene dimer. *The Journal of Physical Chemistry A*, 110(15):5091–5095, 2006.
- [128] S Tsuzuki, K Honda, T Uchimaru, M Mikami, and K Tanabe. Origin of attraction and directionality of the  $\pi/\pi$  interaction: model chemistry calculations of benzene dimer interaction. *Journal of the American Chemical Society*, 124(1):104–112, 2002.
- [129] M Sinnokrot, E Valeev, and C. D Sherrill. Estimates of the ab initio limit for  $\pi$ - $\pi$  interactions: The benzene dimer. *Journal of the American Chemical Society*, 124(36):10887–10893, 2002.
- [130] P Hobza, H. L Selzle, and E. 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.

- [131] C. D Sherrill, T Takatani, and E. G Hohenstein. An assessment of theoretical methods for nonbonded interactions: Comparison to complete basis set limit coupled-cluster potential energy curves for the benzene dimer, the methane dimer, benzene- methane, and benzene- h<sub>2</sub>s†. *The Journal of Physical Chemistry A*, 113(38):10146–10159, 2009.
- [132] D Rocca. Random-phase approximation correlation energies from lanczos chains and an optimal basis set: Theory and applications to the benzene dimer. *The Journal of chemical physics*, 140(18):18A501, 2014.
- [133] J. P Perdew, K Burke, and M Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.
- [134] C Adamo and V Barone. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *The Journal of chemical physics*, 110(13):6158–6170, 1999.
- [135] C Adamo and V Barone. Toward reliable adiabatic connection models free from adjustable parameters. *Chemical Physics Letters*, 274(1):242–250, 1997.
- [136] A. D Becke. Density-functional thermochemistry. III. the role of exact exchange. *The Journal of chemical physics*, 98(7):5648–5652, 1993.
- [137] A. D Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical review A*, 38(6):3098, 1988.
- [138] L Chengteh, Y Weitao, and R. G 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.
- [139] A. D Boese and J. M Martin. Development of density functionals for thermochemical kinetics. *The Journal of chemical physics*, 121(8):3405–3416, 2004.
- [140] C Møller and M. S Plesset. Note on an approximation treatment for many-electron systems. *Physical Review*, 46(7):618, 1934.
- [141] A. D Becke. Density-functional thermochemistry. V. systematic optimization of exchange-correlation functionals. *The Journal of chemical physics*, 107(20):8554–8560, 1997.
- [142] F. A Hamprecht, A. J Cohen, D. J Tozer, and N. C Handy. Development and assessment of new exchange-correlation functionals. *The Journal of chemical physics*, 109(15):6264–6271, 1998.
- [143] P. J Wilson, T. J Bradley, and D. J Tozer. Hybrid exchange-correlation functional determined from thermochemical data and ab initio potentials. *The Journal of Chemical Physics*, 115(20):9233–9242, 2001.
- [144] T. W Keal and D. J Tozer. Semiempirical hybrid functional with improved performance in an extensive chemical assessment. *The Journal of chemical physics*, 123(12):121103, 2005.
- [145] H Iikura, T Tsuneda, T Yanai, and K Hirao. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *The Journal of Chemical Physics*, 115(8):3540–3544, 2001.

- [146] J Chai and M Head-Gordon. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Physical Chemistry Chemical Physics*, 10(44):6615–6620, 2008.
- [147] J. G Angyán, I. C Gerber, A Savin, and J Toulouse. van der waals forces in density functional theory: Perturbational long-range electron-interaction corrections. *Physical Review A*, 72(1):012510, 2005.
- [148] S Grimme. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *Journal of computational chemistry*, 25(12):1463–1473, 2004.
- [149] J Park, B Deok Yu, and S Hong. Ab initio calculations with van der waals corrections: Benzene-benzene intermolecular case and graphite. *J. Korean Phys. Soc.*, 59:196–19, 2011.
- [150] C. M Lee, N. M Mohamed, H. D Watts, J. D Kubicki, and S. H Kim. Sum-frequency-generation vibration spectroscopy and density functional theory calculations with dispersion corrections (DFT-D2) for cellulose i $\alpha$  and i $\beta$ . *The Journal of Physical Chemistry B*, 117(22):6681–6692, 2013.
- [151] G Starkschall and R. G Gordon. Error bounds to long-range three-body, and relativistic interactions between atoms. *The Journal of Chemical Physics*, 57(8):3213–3220, 1972.
- [152] A Tkatchenko and M Scheffler. Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Physical review letters*, 102(7):073005, 2009.
- [153] 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.
- [154] T Bučko, S Lebègue, J. G Ángyán, and J Hafner. Extending the applicability of the Tkatchenko-Scheffler dispersion correction via iterative Hirshfeld partitioning. *The Journal of chemical physics*, 141(3):034114, 2014.
- [155] A Tkatchenko, R A DiStasio Jr, R Car, and M Scheffler. Accurate and efficient method for many-body van der Waals interactions. *Physical review letters*, 108(23):236402, 2012.
- [156] T Bučko, S Lebègue, J Hafner, and J. G Ángyán. Tkatchenko-scheffler van der Waals correction method with and without self-consistent screening applied to solids. *Physical Review B*, 87(6):064110, 2013.
- [157] G. S Waldo, R Carlson, J Moldowan, K. E Peters, and J. E Penner-hahn. Sulfur speciation in heavy petroleums: Information from X-ray absorption near-edge structure. *Geochimica et Cosmochimica Acta*, 55(3):801–814, 1991.
- [158] S Mitra-Kirtley, O. C Mullins, C Y Ralston, D Sellis, and C Pareis. Determination of sulfur species in asphaltene, resin, and oil fractions of crude oils. *Applied spectroscopy*, 52(12):1522–1525, 1998.

- [159] V. V Lobodin, W. K Robbins, J Lu, and R. P Rodgers. Separation and characterization of reactive and non-reactive sulfur in petroleum and its fractions. *Energy & Fuels*, 29(10):6177–6186, 2015.
- [160] H. E Katz. Recent advances in semiconductor performance and printing processes for organic transistor-based electronics. *Chemistry of Materials*, 16(23):4748–4756, 2004.
- [161] B. S Ong, Y Wu, Y Li, P Liu, and H Pan. Thiophene polymer semiconductors for organic thin-film transistors. *Chemistry—A European Journal*, 14(16):4766–4778, 2008.
- [162] M G Vivas, S. L Nogueira, H. S Silva, N. M Barbosa Neto, A Marletta, F Serein-Spirau, S Lois, T Jarrosson, L De Boni, R. A Silva, et al. Linear and nonlinear optical properties of the thiophene/phenylene-based oligomer and polymer. *The Journal of Physical Chemistry B*, 115(44):12687–12693, 2011.
- [163] H. S Silva, S. L Nogueira, J. E Manzoli, N. M Barbosa Neto, A Marletta, F Serein-Spirau, J. P Lere-Porte, S Lois, and R. A Silva. Controlling bandgap energy and multivibronic modes of a poly (2, 5-thiophene-1, 4-dialkoxyphenylene) derivative by gamma photons. *The Journal of Physical Chemistry A*, 115(29):8288–8294, 2011.
- [164] F Kim, G Ren, and S. A Jenekhe. One-dimensional nanostructures of  $\pi$ -conjugated molecular systems: assembly, properties, and applications from photovoltaics, sensors, and nanophotonics to nanoelectronics. *Chemistry of Materials*, 23(3):682–732, 2010.
- [165] A Kishita, S Takahashi, H Kamimura, M Miki, T Moriya, and H Enomoto. Upgrading of bitumen by hydrothermal visbreaking in supercritical water with alkali. *Journal of the Japan Petroleum Institute*, 46(4):215–221, 2003.
- [166] A Samokhvalov. Heterogeneous photocatalytic reactions of sulfur aromatic compounds. *ChemPhysChem*, 12(16):2870–2885, 2011.
- [167] X Chen, J Gao, Y Lu, H Meng, and C Li. Acylation desulfurization of heavy cracking oil as a supplementary oil upgrading pathway. *Fuel Processing Technology*, 130:7–11, 2015.
- [168] O. P Strausz, T. W Mojelsky, and E. M Lown. The molecular structure of asphaltene: an unfolding story. *Fuel*, 71(12):1355–1363, 1992.
- [169] P Liu, Q Shi, K. H Chung, Y Zhang, N Pan, S Zhao, and C Xu. Molecular characterization of sulfur compounds in venezuela crude oil and its SARA fractions by electrospray ionization fourier transform ion cyclotron resonance mass spectrometry. *Energy & Fuels*, 24(9):5089–5096, 2010.
- [170] A Karimi, K Qian, W. N Olmstead, H Freund, C Yung, and M. R Gray. Quantitative evidence for bridged structures in asphaltenes by thin film pyrolysis. *Energy & Fuels*, 25(8):3581–3589, 2011.
- [171] R Coelho, I Hovell, and K 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.
- [172] F. S AlHumaidan, A Hauser, M. S Rana, and H. M Lababidi. Impact of thermal treatment on asphaltene functional groups. *Energy & Fuels*, 30(4):2892–2903, 2016.

- [173] J Widany, G Daminelli, A Di Carlo, P Lugli, G Jungnickel, M Elstner, and T Frauenheim. Electronic band structure and intermolecular interaction in substituted thiophene polymorphs. *Physical Review B*, 63(23):233204, 2001.
- [174] H Liu, E Brémond, A Prlj, J. F Gonthier, and C Corminboeuf. Adjusting the local arrangement of  $\pi$ -stacked oligothiophenes through hydrogen bonds: a viable route to promote charge transfer. *The journal of physical chemistry letters*, 5(13):2320–2324, 2014.
- [175] J. G Speight. *The chemistry and technology of petroleum*. CRC press, 2014.
- [176] O. C. Mullins, S Hassan, J Eyssautier, A. E Pomerantz, L Barré, A Ballard, Y Ruiz-Morales, M Mostowfi, R McFarlane, L Goual, R Lepkowicz, T Cooper, J Orbulescu, R. M Leblanc, J Edwards, and R Zare. Advances in asphaltene science and the yen–mullins model. *Energy & Fuels*, 26(7):3986–4003, 2012.
- [177] J. G Speight. *The desulfurization of heavy oils and residua*. CRC Press, 1999.
- [178] I. D Mackie and G. A DiLabio. Importance of the Inclusion of Dispersion in the Modeling of Asphaltene Dimers. *Energy & Fuels*, 24(12):6468–6475, 2010.
- [179] R. S Morgan, C. E Tatsch, R. H Gushard, J. M Mcadon, and P. K Warne. CHAINS OF ALTERNATING SULFUR AND  $\pi$ -BONDED ATOMS IN EIGHT SMALL PROTEINS. *International journal of peptide and protein research*, 11(3):209–217, 1978.
- [180] T. P Tauer, M. E Derrick, and C. D Sherrill. Estimates of the ab initio limit for sulfur- $\pi$  interactions: the H<sub>2</sub>S-benzene dimer. *The Journal of Physical Chemistry A*, 109(1):191–196, 2005.
- [181] S Tsuzuki, K Honda, and R Azumi. Model chemistry calculations of thiophene dimer interactions: origin of  $\pi$ -stacking. *Journal of the American Chemical Society*, 124(41):12200–12209, 2002.
- [182] A Bree and R Zwarich. The vibrations of dibenzothiophene. *Spectrochimica Acta Part A: Molecular Spectroscopy*, 27(4):599–620, 1971.
- [183] K. H Michaelian, B. E Billinghamurst, J. M Shaw, and V Lastovka. Far-infrared photoacoustic spectra of tetracene, pentacene, perylene and pyrene. *Vibrational Spectroscopy*, 49(1):28–31, 2009.
- [184] J Chai and M Head-Gordon. Systematic optimization of long-range corrected hybrid density functionals. *The Journal of chemical physics*, 128(8):084106, 2008.
- [185] U Salzner and A Aydin. Improved prediction of properties of  $\pi$ -conjugated oligomers with range-separated hybrid density functionals. *Journal of chemical theory and computation*, 7(8):2568–2583, 2011.
- [186] K. B Wiberg. Basis set effects on calculated geometries: 6-311++ G\*\* vs. aug-cc-pVDZ. *Journal of computational chemistry*, 25(11):1342–1346, 2004.
- [187] A Hesselmann, G Jansen, and M Schütz. Density-functional theory-symmetry-adapted intermolecular perturbation theory with density fitting: A new efficient method to study intermolecular interaction energies. *The Journal of chemical physics*, 122(1):014103, 2005.

- [188] R Podeszwa, R Bukowski, and K Szalewicz. Potential energy surface for the benzene dimer and perturbational analysis of  $\pi$ - $\pi$  interactions. *The Journal of Physical Chemistry A*, 110(34):10345–10354, 2006.
- [189] H.J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, et al. Molpro, version 2012, a package of ab initio programs, 2012. see.
- [190] V Barone, M Biczysko, and J Bloino. Fully anharmonic IR and Raman spectra of medium-size molecular systems: accuracy and interpretation. *Physical Chemistry Chemical Physics*, 16(5):1759–1787, 2014.
- [191] 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,  $\text{h}_2 \text{cn}_2$ . *Chemical Physics*, 273(2):91–101, 2001.
- [192] Y Scribano and D. M Benoit. Iterative active-space selection for vibrational configuration interaction calculations using a reduced-coupling VSCf basis. *Chemical Physics Letters*, 458(4):384–387, 2008.
- [193] E. R Johnson, S Keinan, P Mori-Sánchez, J Contreras-Garcia, A. J Cohen, and W Yang. Revealing noncovalent interactions. *Journal of the American Chemical Society*, 132(18):6498–6506, 2010.
- [194] H. S Silva, A. C Sodero, B Bouyssiere, H Carrier, J Korb, A Alfarra, G Vallverdu, D Bégué, and I Baraille. Molecular Dynamics Study of Nanoaggregation in Asphaltene Mixtures: Effects of the N, O, and S Heteroatoms. *Energy & Fuels*, 30(7):5656–5664, 2016.
- [195] J Goupy and L Creighton. *Introduction aux plans d'expériences-5e éd.: Toutes les techniques nécessaires à la conduite d'une étude*. Dunod, 2013.
- [196] R. B Campbell, J. M 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.
- [197] C. C Mattheus, A. B Dros, J Baas, A Meetsma, J. L de Boer, and T. T Palstra. Polymorphism in pentacene. *Acta Crystallographica Section C: Crystal Structure Communications*, 57(8):939–941, 2001.
- [198] C. P Brock and J. D Dunitz. Temperature dependence of thermal motion in crystalline naphthalene. *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry*, 38(8):2218–2228, 1982.
- [199] J. C Facelli and D. M Grant. Determination of molecular symmetry in crystalline naphthalene using solid-state NMR. 1993.
- [200] C. P Brock and J. D Dunitz. Temperature dependence of thermal motion in crystalline anthracene. *Acta Crystallographica Section B: Structural Science*, 46(6):795–806, 1990.
- [201] K. H Michaelian, Q Wen, B. E Billinghamurst, J. M Shaw, and V Lastovka. Far-and mid-infrared photoacoustic spectra of tetracene, pentacene, perylene and pyrene. *Vibrational Spectroscopy*, 58:50–56, 2012.

- [202] B Billinghamurst, T May, L Dallin, W Wurtz, M de Jong, and K. H Michaelian. Photoacoustic spectroscopy using coherent synchrotron radiation: application to  $\alpha$ -lactose monohydrate. *Optics letters*, 35(18):3090–3092, 2010.
- [203] 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.
- [204] G Kresse and J Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical review B*, 54(16):11169, 1996.
- [205] S Grimme, S Ehrlich, and L Goerigk. Effect of the damping function in dispersion corrected density functional theory. *Journal of computational chemistry*, 32(7):1456–1465, 2011.
- [206] J. M Bowman, T Carrington, and H Meyer. Variational quantum approaches for computing vibrational energies of polyatomic molecules. *Molecular Physics*, 106(16-18):2145–2182, 2008.
- [207] R Dovesi, R Orlando, A Erba, C M Zicovich-Wilson, B Civalleri, S Casassa, L Maschio, M Ferrabone, M De La Pierre, P 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.
- [208] 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.
- [209] E. J Meijer and M Sprik. A density-functional study of the intermolecular interactions of benzene. *The Journal of chemical physics*, 105(19):8684–8689, 1996.
- [210] A Katrusiak, M Podsiadło, and A Budzianowski. Association ch...  $\pi$  and no van der waals contacts at the lowest limits of crystalline benzene i and ii stability regions. *Crystal Growth & Design*, 10(8):3461–3465, 2010.
- [211] F Fabbiani, D. R Allan, S Parsons, and C. R Pulham. Exploration of the high-pressure behaviour of polycyclic aromatic hydrocarbons: naphthalene, phenanthrene and pyrene. *Acta Crystallographica Section B: Structural Science*, 62(5):826–842, 2006.
- [212] U Sondermann, A Kutoglu, and H Bassler. X-ray diffraction study of the phase transition in crystalline tetracene. *The Journal of Physical Chemistry*, 89(9):1735–1741, 1985.
- [213] R. B Campbell, J. M Robertson, and J Trotter. The crystal and molecular structure of pentacene. *Acta Crystallographica*, 14(7):705–711, 1961.
- [214] E. P Krainov. The vibrational spectra of aromatic compounds. xix. calculation and interpretation of the vibrational spectra of Naphthalene and some Deuterionaphthalenes. *Opt. Spectr.(USSR)(English Transl.)*, 16, 1964.
- [215] R Kalesky, E Kraka, and D Cremer. Description of aromaticity with the help of vibrational spectroscopy: Anthracene and phenanthrene. *The Journal of Physical Chemistry A*, 118(1):223–237, 2013.

- [216] A Bree and RA Kydd. Infrared spectrum of anthracene crystals. *The Journal of Chemical Physics*, 48(12):5319–5325, 1968.
- [217] D. M Burns and J Iball. Refinement of the structure of chrysene. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 257, pages 491–514. The Royal Society, 1960.
- [218] C Näther, H Bock, Z Havlas, and T Hauck. Solvent-shared and solvent-separated ion multiples of perylene radical anions and dianions: an exemplary case of alkali metal cation solvation. *Organometallics*, 17(21):4707–4715, 1998.
- [219] V. K Belsky, V. E Zavodnik, and V. M Vozzhennikov. Fluorene,  $c_{13}h_{10}$ . *Acta Crystallographica Section C: Crystal Structure Communications*, 40(7):1210–1211, 1984.
- [220] J Godec and L Colombo. Interpretation of the vibrational spectrum of crystalline phenanthrene. *The Journal of Chemical Physics*, 65(11):4693–4700, 1976.
- [221] A Bree, F. G t Solven, and V. V Vilkos. A vibrational analysis for phenanthrene. *Journal of Molecular Spectroscopy*, 44(2):298–319, 1972.
- [222] K Zhang, B Guo, P Colarusso, and P. F Bernath. Far-infrared emission spectra of selected gas-phase PAHs: Spectroscopic fingerprints. *Science*, 274(5287):582, 1996.
- [223] K. H Michaelian, S. A Oladepo, J. M Shaw, X Liu, D Bégué, and I Baraille. Raman and photoacoustic infrared spectra of fluorene derivatives: Experiment and calculations. *Vibrational Spectroscopy*, 74:33–46, 2014.
- [224] VK Belskii. Structure of carbazole. *Kristallografiya*, 30:193–194, 1985.
- [225] O Dideberg, L Dupont, and J. M André. The crystal structure of dibenzofuran. *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry*, 28(4):1002–1007, 1972.
- [226] R. M Schaffrin and J Trotter. Structure of dibenzothiophen. *Journal of the Chemical Society A: Inorganic, Physical, Theoretical*, pages 1561–1565, 1970.
- [227] T. D Klots and W. B Collier. Vibrational spectra, structure, assignment, and ideal-gas thermodynamics of a three-ring molecule: dibenzofuran. *Journal of molecular structure*, 380(1):1–14, 1996.
- [228] A. C Sodero, H Santos-Silva, P Guevara Level, B Bouyssiere, J Korb, H Carrier, A Alfarra, D Bégué, and I Baraille. Investigation of the effect of sulfur heteroatom on asphaltene aggregation. *Energy & Fuels*, 2016.
- [229] O. C Mullins. The asphaltenes. *Annual Review of Analytical Chemistry*, 4:393–418, 2011.
- [230] J. J Adams. Asphaltene adsorption, a literature review. *Energy & Fuels*, 28(5):2831–2856, 2014.
- [231] L Benamsili, J Korb, G Hamon, A Louis-Joseph, B Bouyssiere, H Zhou, and R. 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.

- [232] O. C Mullins. The modified yen model. *Energy & Fuels*, 24(4):2179–2207, 2010.
- [233] H Sabbah, A. L Morrow, A. E Pomerantz, and R. N Zare. Evidence for island structures as the dominant architecture of asphaltenes. *Energy & Fuels*, 25(4):1597–1604, 2011.
- [234] 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.
- [235] M Sedghi, L Goual, W Welch, and J Kubelka. Effect of asphaltene structure on association and aggregation using molecular dynamics. *The Journal of Physical Chemistry B*, 117(18):5765–5776, 2013.
- [236] J Liu, Y Zhao, and S 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.
- [237] F Gao, Z Xu, G Liu, and S Yuan. Molecular dynamics simulation: the behavior of asphaltene in crude oil and at the oil/water interface. *Energy & Fuels*, 28(12):7368–7376, 2014.
- [238] R. B Teklebrhan, L Ge, S Bhattacharjee, Z Xu, and J Sjöblom. Initial partition and aggregation of uncharged polyaromatic molecules at the oil–water interface: a molecular dynamics simulation study. *The Journal of Physical Chemistry B*, 118(4):1040–1051, 2014.
- [239] B Schuler, G Meyer, D Peña, O C Mullins, and L Gross. Unraveling the molecular structures of asphaltenes by atomic force microscopy. *Journal of the American Chemical Society*, 137(31):9870–9876, 2015.
- [240] S Kim, R. P Rodgers, G. T Blakney, C. L Hendrickson, and A. 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.
- [241] L. A Stanford, R. P Rodgers, A. G Marshall, J Czarnecki, and X. 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.
- [242] E Kendrick. A mass scale based on  $\text{CH}_2 = 14.0000$  for high resolution mass spectrometry of organic compounds. *Analytical Chemistry*, 35(13):2146–2154, 1963.
- [243] C. A Hughey, C. L Hendrickson, R. P Rodgers, A. G Marshall, and K Qian. Kendrick mass defect spectrum: a compact visual analysis for ultrahigh-resolution broadband mass spectra. *Analytical Chemistry*, 73(19):4676–4681, 2001.
- [244] Q Shi, D Hou, K. H Chung, C Xu, S Zhao, and Y 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.

- [245] F Frigerio and D Molinari. A multiscale approach to the simulation of asphaltenes. *Computational and Theoretical Chemistry*, 975(1):76–82, 2011.
- [246] Y Mikami, Y Liang, T Matsuoka, and E. 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.
- [247] F Yang, P Tchoukov, H Dettman, R. B Teklebrhan, L Liu, T Dabros, J Czarnecki, J Masliyah, and Z 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.
- [248] H Berendsen, D van der Spoel, and R van Drunen. Gromacs: a message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91(1):43–56, 1995.
- [249] B Hess, C Kutzner, D Van Der Spoel, and E Lindahl. Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of chemical theory and computation*, 4(3):435–447, 2008.
- [250] D Van Der Spoel, E Lindahl, B Hess, G Groenhof, A. E Mark, and H Berendsen. Gromacs: fast, flexible, and free. *Journal of computational chemistry*, 26(16):1701–1718, 2005.
- [251] W. F van Gunsteren, X Daura, and A. E Mark. GROMOS force field. *Encyclopedia of computational chemistry*, 1998.
- [252] C Oostenbrink, A Villa, A. E Mark, and W. 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.
- [253] T Kuznicki, J. H Masliyah, and S Bhattacharjee. Molecular dynamics study of model molecules resembling asphaltene-like structures in aqueous organic solvent systems. *Energy & Fuels*, 22(4):2379–2389, 2008.
- [254] T Kuznicki, J. H Masliyah, and S Bhattacharjee. Aggregation and partitioning of model asphaltenes at toluene– water interfaces: Molecular dynamics simulations. *Energy & Fuels*, 23(10):5027–5035, 2009.
- [255] R. B Teklebrhan, L Ge, S Bhattacharjee, Z Xu, and J 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.
- [256] A. W Schüttelkopf and D Van Aalten. Prodrg: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallographica Section D: Biological Crystallography*, 60(8):1355–1363, 2004.
- [257] D. W.; Krekelberg W. P. Shen, V. K.; Siderius. Nist standard reference simulation website, nist standard reference database number 173, 2015.

- [258] W Huang, A. P Eichenberger, and W. F van Gunsteren. Molecular dynamics simulation of thionated hen egg white lysozyme. *Protein Science*, 21(8):1153–1161, 2012.
- [259] C Lee, W Yang, and R. G Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical review B*, 37(2):785, 1988.
- [260] P. J Stephens, F. J Devlin, C. F Chabalowski, and M. J Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *The Journal of Physical Chemistry*, 98(45):11623–11627, 1994.
- [261] A Schäfer, H Horn, and R Ahlrichs. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *The Journal of Chemical Physics*, 97(4):2571–2577, 1992.
- [262] F Weigend and R 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.
- [263] F Neese. The orca program system. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(1):73–78, 2012.
- [264] G Bussi, D Donadio, and M Parrinello. Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 126(1):014101, 2007.
- [265] H Berendsen, J van Postma, W. F van Gunsteren, A DiNola, and J. R Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.
- [266] W. G Hoover. Canonical dynamics: equilibrium phase-space distributions. *Physical Review A*, 31(3):1695, 1985.
- [267] S Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molecular physics*, 52(2):255–268, 1984.
- [268] M Parrinello and A Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.
- [269] L Verlet. Computer" experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Physical review*, 159(1):98, 1967.
- [270] U Essmann, L Perera, M. L Berkowitz, T Darden, H Lee, and L. G Pedersen. A smooth particle mesh ewald method. *The Journal of chemical physics*, 103(19):8577–8593, 1995.
- [271] B Hess, H Bekker, H Berendsen, and J Fraaije. Lincs: a linear constraint solver for molecular simulations. *Journal of computational chemistry*, 18(12):1463–1472, 1997.
- [272] N. E Jackson, B. M Savoie, K. L Kohlstedt, M Olvera De La Cruz, G. C Schatz, L. X Chen, and M 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.
- [273] D Ayma. *Etude des propriétés électroniques de diffusion des systèmes périodiques. Application aux semi-conducteurs à structure zinc-blende*. PhD thesis, 1997.

- [274] A Dalgarno and W. D Davison. The calculation of van der Waals interactions. *Advances in atomic and molecular physics*, 2:1–32, 1966.
- [275] B Linder. Generalized form for dispersion interaction. *The Journal of Chemical Physics*, 37(5):963–966, 1962.
- [276] Y. M Chan and A Dalgarno. The long-range interaction of atoms and molecules. *Molecular Physics*, 9(4):349–356, 1965.
- [277] A Dalgarno and G. A Victor. Long range three-body forces between Helium and Hydrogen atoms. *Molecular Physics*, 10(4):333–337, 1966.
- [278] H. B Casimir and D Polder. The influence of retardation on the London-van der Waals forces. *Physical Review*, 73(4):360, 1948.
- [279] B. J Orr and J. F Ward. Perturbation theory of the non-linear optical polarization of an isolated system. *Molecular Physics*, 20(3):513–526, 1971.
- [280] W. B Collier and T. D Klots. Heteroatom derivatives of indene. part 1. vibrational frequencies and a refined scaled overlay of the am1 force field of indole, benzofuran benzothiophene, benzoxazole and benzothiazole. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 51(8):1255–1272, 1995.
- [281] V. B Singh. Ab initio and DFT studies of the vibrational spectra of benzofuran and some of its derivatives. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 65(5):1125–1130, 2006.
- [282] T. D Klots. Vibrational spectra of indene. part 4. calibration, assignment, and ideal-gas thermodynamics. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 51(13):2307–2324, 1995.
- [283] A. A El-Azhary. A DFT study of the geometries and vibrational spectra of indene and some of its heterocyclic analogues, benzofuran, benzoxazole, bensothiophene, benzothiazole, indole and indazole. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 55(12):2437–2446, 1999.
- [284] S Fomine, M Tlenkopatchev, S Martinez, and L Fomina. Local MP2 study of naphthalene, indole, and 2, 3-benzofuran dimers. *The Journal of Physical Chemistry A*, 106(15):3941–3946, 2002.