

N° D'ORDRE : 7310

UNIVERSITÉ PARIS XI  
UFR SCIENTIFIQUE D'ORSAY

# THÈSE

Présentée pour obtenir

Le GRADE de DOCTEUR en SCIENCES

De l'UNIVERSITÉ PARIS XI - ORSAY

par

**Benjamin BOUVIER**

Étude théorique du transfert de charge et d'excitation dans les  
assemblées de molécules : application à des agrégats de van  
der Waals et à des doubles brins d'oligonucléotides

Soutenue le 29 septembre 2003 devant la commission d'examen :

M. Laurent DUCASSE  
M. Fernand SPIEGELMANN  
M. Alain FUCHS  
M. Richard LAVERY  
M. Philippe MILLIÉ

Rapporteur  
Rapporteur  
Président



N° D'ORDRE : 7310

UNIVERSITÉ PARIS XI  
UFR SCIENTIFIQUE D'ORSAY

# THÈSE

Présentée pour obtenir

Le GRADE de DOCTEUR en SCIENCES

De l'UNIVERSITÉ PARIS XI - ORSAY

par

**Benjamin BOUVIER**

Étude théorique du transfert de charge et d'excitation dans les  
assemblées de molécules : application à des agrégats de van  
der Waals et à des doubles brins d'oligonucléotides

Soutenue le 29 septembre 2003 devant la commission d'examen :

M. Laurent DUCASSE  
M. Fernand SPIEGELMANN  
M. Alain FUCHS  
M. Richard LAVERY  
M. Philippe MILLIÉ

Rapporteur  
Rapporteur  
Président



Ces trois années de thèse ont été effectuées au Commissariat à l'Énergie Atomique, sur le site de Saclay, dans le Service des Photons, Atomes et Molécules (SPAM). Je tiens à remercier Didier Normand de m'y avoir accueilli.

Alain Fuchs a bien voulu présider le jury de cette thèse ; qu'il en soit remercié ici. Ce travail n'eut pas été possible sans la confiance qu'il m'a accordée en m'intégrant dans le DEA de Physico-Chimie Moléculaire, il y a de cela quatre ans ; à nouveau, merci.

J'exprime ma gratitude à Laurent Ducasse et Fernand Spiegelmann, qui ont accepté d'être les rapporteurs de ce travail. Qu'ils soient remerciés de l'attention avec laquelle ils ont lu ce manuscrit, de l'intérêt qu'ils lui ont témoigné, et des critiques constructives qui ont contribué à son amélioration.

L'étude du transfert d'excitation dans l'ADN doit beaucoup à Richard Lavery qui, par ses conseils, sa connaissance des acides nucléiques et ses calculs de dynamique moléculaire, y a contribué depuis le commencement. J'ai beaucoup apprécié sa passion pour la recherche et sa rigueur, et je garde un excellent souvenir de nos discussions, scientifiques et autres. Qu'il soit en outre remercié d'avoir bien voulu faire partie du jury de cette thèse. Merci également à Krystyna Zakrzewska, avec qui j'ai eu grand plaisir à travailler et à discuter.

Au cours de ces trois années de thèse, Philippe Millié a su me communiquer son goût pour la recherche, la discussion scientifique à bâtons rompus et la collaboration entre théoriciens et expérimentateurs. J'ai beaucoup apprécié sa patience, sa pédagogie, son goût pour l'enseignement et sa très grande culture générale scientifique. Je garde un excellent souvenir de nos débats (même houleux !) et lui souhaite une retraite aussi riche et féconde que le fut sa carrière scientifique.

Dimitra Markovitsi est l'initiatrice des présents travaux sur le transfert d'excitation dans l'ADN. J'ai grandement apprécié son optimisme et ses qualités scientifiques. Travailler avec elle fut un réel plaisir. Je la remercie également de m'avoir accueilli au sein du Laboratoire Francis Perrin (LFP) qu'elle dirige.

Michel Mons m'a patiemment expliqué les subtilités de la photoionisation des agrégats aromatiques ; l'article à ce sujet lui doit beaucoup. J'ai beaucoup apprécié tant l'étendue de ses connaissances scientifiques que la finesse de son humour... pour ne pas parler de sa maîtrise de Photoshop.

Je tiens à remercier les membres du Groupe de Chimie Théorique : Valérie Brenner, Jean-Pierre Dognon, Jean-Maïk Soudan et Patrick de Pujo. Outre les innombrables fois où, chacun dans son domaine, ils m'ont sorti d'une ornière, je voudrais leur exprimer ma

gratitude pour l'ambiance de tous les jours au laboratoire, ingrédient essentiel d'une thèse réussie.

J'adresse ma gratitude à l'ensemble des membres du SPAM et du LFP pour leur accueil, et plus particulièrement à Jacqueline Bandura, Pascal Salières, Véronique Gereczy, Iliana Dimicoli, François PiuZZi, Mohamed Elhanine, Benjamin Tardivel (date de péremption: mai 2003), Thomas Gustavsson, Tu-Hoa Tran-Thi, Cécile Reynaud, Dominique Porterat et Daniel Crozat.

Je remercie également le groupe Informatique du DRECAM, pour de multiples dépannages et d'intéressantes discussions techniques: Angel Sanchez, Sylvie Jubéra, Stéphane Delaporte et Laurent Desèze. Merci également à Anne Capdepon pour la partie Unix.

Merci ensuite à Rodolphe Pollet d'avoir partagé avec moi ses goûts pour la science, l'informatique, la musique, la lecture, la politique, la photo... Je lui souhaite toute la réussite qu'il mérite dans chacun de ces domaines. Merci également à Jean-François Fuchs pour sa bonne humeur constante (une denrée rare !), sa camaraderie, son hospitalité, et nos débats passionnés sur la politique militaire française et les jeux vidéo. Je remercie aussi tous les thésards, post-docs et stagiaires que j'ai côtoyés au cours de ces trois années, et plus particulièrement Delphine Onidas, Anne-Laure Derepas, Luca Quaglia, Vincent Dubois, Sébastien Sorgues, Denis Céolin, Wutharath Chin et Léa Trichet.

Un grand merci à mes parents d'avoir toujours cru en moi et de m'avoir encouragé à chercher ce pourquoi je suis fait. Merci également à Aude qui a toujours été là dans les moments difficiles. Et merci à Paul, qui m'a appris bien des choses malgré lui... Ah, et aussi à Jazz.

Je remercie Fred et FX d'être tout simplement mes amis.

Ma gratitue à tous ceux qui m'ont fourni l'inspiration au cours de la rédaction du manuscrit de thèse: Jean-Sébastien Bach, Frédéric Chopin, Ludwig van Beethoven, Domenico Scarlatti, Felix Mendelssohn, Antonio Vivaldi, Wolfgang Amadeus Mozart, mais aussi Joy Division, Megadeth, Iron Maiden, Coldplay, Lemonheads, Ride et les autres...

Et enfin, merci à Christine qui a supporté mes sautes d'humeur, incertitudes, grands discours et autres journées B, avec patience, humour et affection. Cette thèse, je te la dois et te la dédie.

*À Christine.*

*À tous ceux qui y croient.*



*Quiconque réfléchit sur quatre choses, mieux vaudrait qu'il ne soit jamais né : ce qui est dessus, ce qui est dessous, ce qui est avant et ce qui est après.*

*Talmud, Hagigah 2.1.*

*Mais on ne se bat pas dans l'espoir du succès !  
Non ! Non ! C'est bien plus beau lorsque c'est inutile !*

*Edmond Rostand, Cyrano de Bergerac*



# Table des Matières

INTRODUCTION GÉNÉRALE.....	13
<b>Introduction Générale</b> .....	<b>17</b>
OUTILS ET MÉTHODES .....	25
<b>1. Principes de Chimie Quantique</b> .....	<b>29</b>
1.1. Introduction .....	29
1.2. L'approximation de Born – Oppenheimer .....	29
1.3. Les déterminants de Slater .....	30
1.4. Théorème variationnel.....	31
1.5. Théorie de Hartree-Fock .....	32
1.6. Méthodes post Hartree-Fock .....	34
1.7. Méthodes semi-empiriques .....	39
1.8. Le modèle <i>valence bond</i> .....	41
1.9. Fonctions de base .....	42
<b>2. Potentiel Modèle Intermoléculaire</b> .....	<b>45</b>
2.1. Méthode de perturbation d'échange.....	46
<b>3. Description d'un Supersystème : Méthodes à Hamiltonien Effectif</b> .....	<b>53</b>
<b>4. Etude de Surfaces de d'Énergie Potentielle</b> .....	<b>57</b>
4.1. Surfaces d'énergie potentielle .....	57
4.2. Degrés de liberté d'un ensemble de corps rigides.....	58
4.3. Classification des méthodes d'optimisation.....	59
4.4. Méthodes globales.....	62
4.5. Méthodes locales .....	66
4.6. Méthodes hybrides et extensions.....	67
<b>5. Dynamique Moléculaire Classique</b> .....	<b>69</b>
5.1 Généralités.....	69
5.2. Mouvement de translation des molécules .....	69
5.3. Mouvement de rotation des molécules.....	70
5.4. Fonction d'autocorrélation des vitesses .....	72
5.5. Thermostats .....	73
5.6. Principe de Gauss .....	76
<b>6. Dynamique Moléculaire avec Transitions Électroniques</b> .....	<b>79</b>
6.1. Méthode de <i>Trajectory Surface Hopping</i> .....	79
6.2. Mise en œuvre .....	83
<b>7. Développement et Programmation</b> .....	<b>85</b>
7.1. Langages compilés .....	86
7.2. Langages interprétés.....	88
LE TRANSFERT DE CHARGE .....	91
<b>1. Description du Transfert de Charge</b> .....	<b>95</b>
1.1. Introduction .....	95
1.2. Le modèle à hamiltonien effectif .....	98
1.3. Calcul des éléments diagonaux de l'hamiltonien effectif .....	100
1.4. Calcul des éléments non diagonaux de l'hamiltonien effectif .....	111

<b>2. Phénomènes de Résonance de Charge dans les Agrégats Aromatiques Ionisés</b>	<b>119</b>
2.1. Introduction	119
2.2. Article	119
<b>3. Mécanismes de l'Ionisation dans les Agrégats Aromatiques</b>	<b>137</b>
3.1. Introduction	137
3.2. Article	137
<b>4. Conclusion</b>	<b>161</b>
<b>LE TRANSFERT D'EXCITATION</b>	<b>165</b>
<b>1. Les Acides Nucléiques</b>	<b>169</b>
1.1. Présentation des acides nucléiques	169
1.2. Les doubles brins poly-dA et poly-(dAdT)	174
1.3. Tautomérie des bases azotées	176
<b>2. Description des États Électroniques Excités de Fragments d'ADN</b>	<b>183</b>
2.1. Théorie excitonique	183
2.2. États excités des bases de l'ADN	192
<b>3. États Excités Collectifs de Fragments d'ADN : Approche Statique</b>	<b>197</b>
3.1. Introduction	197
3.2. Article	197
<b>4. États Excités Collectifs de Fragments d'ADN : Approche Dynamique</b>	<b>217</b>
4.1. Introduction	217
4.2. Paramétrage du modèle électrostatique	217
4.3. Génération de la matrice excitonique	230
4.4. Article	232
<b>5. Dynamique Semi-Classique du Transfert de l'Excitation</b>	<b>265</b>
5.1. Introduction	265
5.2. Mise en place du potentiel et des couplages	267
5.3. Désexcitation non radiative du dimère de naphthalène excité	272
5.4. Désexcitation non radiative d'agrégats colonnaires de naphthalène	277
5.5. Conclusion	283
<b>6. Conclusion</b>	<b>285</b>
<b>CONCLUSION GÉNÉRALE</b>	<b>287</b>
<b>RÉFÉRENCES BIBLIOGRAPHIQUES</b>	<b>295</b>
<b>Références Bibliographiques</b>	<b>297</b>

# INTRODUCTION GÉNÉRALE



*La carte n'est pas le territoire.*

Alfred Korzybski, *Science and Sanity*



## Introduction Générale

**S**i l'histoire de l'investigation des propriétés stationnaires et dynamiques de systèmes moléculaires s'étend sur l'intégralité du siècle dernier, elle a connu, au cours des deux décennies précédentes, une évolution accélérée caractérisée par deux sous-tendances. En premier lieu, l'étude de molécules à leurs échelles naturelles est devenue envisageable, avec des résolutions spatiales de quelques ångströms et des échelles temporelles de quelques femtosecondes. En second lieu, il est à présent possible de détecter et d'agir sur les propriétés des molécules elles-mêmes<sup>1,2</sup>. Simultanément, plusieurs communautés scientifiques (physiciens, chimistes et biologistes), précédemment disjointes, ont œuvré au franchissement des barrières traditionnelles séparant leurs disciplines respectives, par un nombre toujours plus important de travaux tant théoriques qu'expérimentaux. En particulier, l'étude des processus de transfert au niveau moléculaire, mettant en jeu des entités telles que le proton, l'électron ou l'exciton, ont facilité la compréhension de phénomènes aussi divers que la catalyse hétérogène, les mécanismes photoinduits dans les macromolécules (tels la redistribution vibrationnelle, les transitions conformationnelles ou les réactions de dégradation), et les transferts de charge ou d'énergie élémentaires caractéristiques des biomolécules. Dans cette dernière classe de phénomènes prennent place bon nombre de mécanismes élémentaires du vivant : la photosynthèse, l'activité enzymatique, les mécanismes de dégradation et de réparation de l'ADN n'en sont que quelques exemples. L'électronique moléculaire, c'est-à-dire le stockage et le traitement de l'information à l'échelle nanométrique, procède également de tels mécanismes ; son étude a été aiguillonnée par le fait qu'elle pourrait, à terme, supplanter l'électronique traditionnelle à base de silice dont elle pallierait les manquements.

L'identification des phénomènes de transfert suit de peu l'avènement de la mécanique quantique. Dès les années 1920, les bases théoriques, telles que le principe de Franck-Condon<sup>3,4</sup> et l'approximation de Born-Oppenheimer<sup>5,6</sup>, avaient été posées. Le transfert de particule par effet tunnel fut mis en évidence pour la première fois dans le cadre de la réaction d'inversion de l'ammoniac<sup>7</sup>. En 1932, Landau et Zener<sup>8-10</sup> posèrent les fondements quantiques de la probabilité de transfert lors de l'intersection de surfaces d'énergie potentielle. Les travaux de Gurney<sup>11</sup> prouvèrent alors, bien que de façon qualitative, le couplage entre niveaux électroniques et mouvement des noyaux. La période suivant la deuxième guerre mondiale se révéla particulièrement féconde, notamment en ce qui concerne les processus de

transfert : dans les solides, avec les travaux de Davydov<sup>12</sup> ou Kubo<sup>13</sup> par exemple, et les biomolécules avec ceux de Duysens<sup>14</sup>. La nature des états vibroniques locaux dans les semiconducteurs et les cristaux polaires, de même que celle des mécanismes afférents (radiatifs ou non), fut ainsi conceptualisée. Le rôle du solvant dans les phénomènes de transfert fut également reconnu<sup>15</sup>. Les théories actuelles du transfert d'électron apparurent dans les années 1950 suite aux travaux de Marcus<sup>16,17</sup>, précédant de peu ceux de Hush<sup>18,19</sup>. Trois années plus tard, Levich et Dogonadze<sup>20-22</sup> furent les premiers à considérer le transfert d'électron comme une transition quantique entre états électroniques, induite par des couplages perturbateurs. Leur théorie fut complétée peu après par les notions de couplages forts et faibles, intimement liées aux limites adiabatiques et diabatiques.

Le domaine de l'électronique moléculaire est l'un des premiers à bénéficier des avancées scientifiques et technologiques liées aux phénomènes de transfert. Il s'agit ici de transférer, de traiter et de stocker de l'information sous forme élémentaire, comme en sont capables les semiconducteurs utilisés aujourd'hui. En perpétuelle évolution, la technologie des semiconducteurs se rapproche pourtant inexorablement de ses limites. Les applications actuelles requièrent une puissance et une cadence d'accès telles que la dissipation de la chaleur produite par effet Joule dans les circuits intégrés est devenue un problème insoluble. La miniaturisation toujours croissante des composants élémentaires ne parvient pas à enrayer la tendance et demande des investissements colossaux de la part des industriels. L'utilisation de systèmes moléculaires pallie cet inconvénient et permet d'atteindre un degré de miniaturisation inaccessible à l'électronique à base de silice. Le transfert de charge entre entités moléculaires est un des vecteurs possibles de l'information au sein de ces systèmes ; il a été montré, par exemple, que le transfert de charge intramoléculaire dans un polymère multichromophorique est compatible avec la logique de Spencer-Brown sur laquelle sont fondées les machines de von Neumann<sup>23,24</sup>. De tels systèmes trouvent également des applications pour le stockage de l'information au sein d'un futur 'ordinateur chimique'<sup>25</sup>. Dans l'électronique moléculaire comme dans bien des domaines, les interactions avec la biologie génèrent des concepts novateurs. Des systèmes biomimétiques visant à émuler le fonctionnement des neurones du cerveau humain, ont ainsi reçu une certaine attention. Le transfert d'excitation électronique entre groupes fonctionnels au sein d'une assemblée de vésicules phospholipidiques<sup>26</sup> en constitue un exemple. La mise en pratique de tels concepts passe par la synthèse de systèmes pour l'électronique moléculaire ; la palette de techniques de nanolithographie disponibles, au premier rang desquelles les films de Langmuir-Blodgett<sup>27</sup>, ne cesse parallèlement de s'accroître.

Bon nombre de phénomènes biologiques ont pour origine des phénomènes de transfert de charge ou d'excitation. La complexité des systèmes concernés a longtemps restreint l'étude de ces phénomènes à une approche macroscopique et phénoménologique. Le développement rapide des méthodologies et de la puissance de calcul, couplé aux informations structurales issues de techniques analytiques puissantes (RMN et diffraction des rayons X à haute résolution), permet aujourd'hui une compréhension plus détaillée de ces phénomènes et des mécanismes élémentaires dont ils participent.

La photosynthèse, par exemple, peut être envisagée comme un ensemble de réactions couplées de transfert d'excitation et de proton. Elle est caractérisée par le transfert ultra-rapide de l'énergie photonique depuis une antenne vers un centre de réaction, où elle est convertie en une séparation de charge stable, *via* un système photosynthétique complexe jouant le rôle de pont. L'énergie ainsi récoltée est stabilisée sous forme de potentiel rédox ou d'énergie de liaison dans les couples ATP/ADP,  $\text{NAD}^+/\text{NADH}$  et  $\text{NADP}^+/\text{NADPH}$ . Elle peut alors être employée, lors de la phase dite obscure de la photosynthèse, à la production de glucose. Le mécanisme inverse, c'est-à-dire la libération de l'énergie contenue sous forme de liaisons chimiques dans le glucose, peut alors être mise en œuvre à tout moment *via* le cycle de Krebs. Une chaîne de mécanismes de transport conduit aux produits de dégradation finaux (cf. Figure 1.1). Le mécanisme de transfert de l'excitation photonique est d'une efficacité telle qu'environ 3000 sauts entre sites sont possibles pendant la durée de vie moyenne de l'excitation<sup>14</sup>. La théorie de Förster<sup>28</sup>, invoquée jusqu'à présent comme base conceptuelle du phénomène, est maintenant compromise par la mise en évidence d'états délocalisés et de recouvrements entre densités électroniques, caractéristiques de couplages forts<sup>29</sup>, et dont la retranscription correcte est délicate.

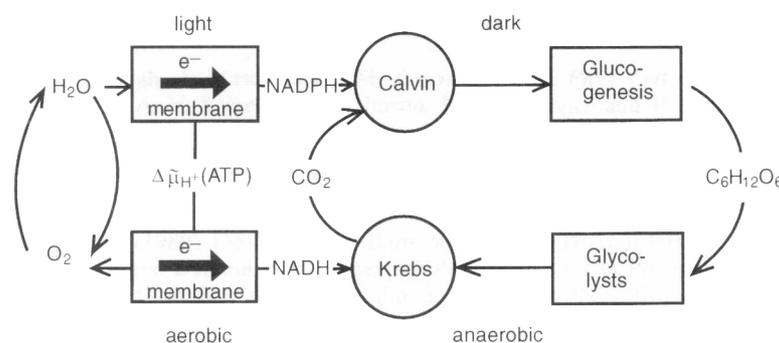


Figure 1.1

La biochimie des acides nucléiques, vecteurs de l'information génétique, fait également intervenir bon nombre de phénomènes de transfert. Depuis la découverte, en 1953,

## Introduction Générale

de la structure de l'ADN par Watson et Crick<sup>30</sup>, la connaissance de ces systèmes n'a cessé de s'accroître à un rythme soutenu, et ce à différentes échelles temporelles et spatiales. Deux consortiums internationaux de généticiens, l'un public<sup>31</sup> et l'autre privé<sup>32</sup>, présentèrent de manière quasi-simultanée, en mars 2001, une première ébauche de cartographie du génôme humain – 30 000 gènes et 3 milliards de paires de bases. En 1996, la première expérience concluante de clonage par transfert nucléaire fut accomplie, avec la brebis Dolly, au Roslin Institute d'Edinburgh<sup>33</sup>. Les mécanismes du vieillissement et de la mort cellulaire programmée (l'apoptose), contenus dans l'information génétique, ont été l'objet de travaux soutenus<sup>34-36</sup>. La biologie et la génétique moléculaires ne sont pas en reste : l'élucidation de la structure de fragments d'ADN a grandement progressé avec l'avènement de résolutions spatiales de l'ordre de un ångström<sup>37,38</sup>. Elles permettent une meilleure compréhension des relations structure-fonction, qui régissent par exemple l'interaction entre l'ADN et les enzymes. De telles interactions, de même que les réactions auxquelles elles donnent naissance, peuvent être simulées par des calculs de dynamique moléculaire, dont la précision et la fidélité par rapport aux systèmes réels ne cessent de croître sous l'effet conjoint de l'amélioration des champs de forces<sup>39-42</sup> et de l'augmentation de la puissance de calcul disponible ; le lecteur est renvoyé aux publications suivantes<sup>43-48</sup> pour une sélection d'exemples récents. L'échelle de taille inférieure, occupée par les bases et les paires de bases, est le domaine de la chimie quantique ; les informations détaillées qu'elle fournit constituent souvent autant de précieuses briques élémentaires extrapolables à des phénomènes plus globaux<sup>49-52</sup>.

Les dégâts causés à l'ADN par l'environnement dans lequel il se trouve (rayonnements<sup>53-58</sup>, conditions chimiques<sup>59-62</sup>, molécules intercalantes<sup>63-65</sup>...) sont maintenant directement reliés à l'apparition de mutations aléatoires qui peuvent se révéler carcinogènes et provoquer la mort cellulaire<sup>66-69</sup>. Or, les réactions induites par irradiation de l'ADN sont favorisées pour peu qu'elles soient précédées par un processus de transfert de charge ou d'énergie d'excitation électronique à travers la double hélice, vers des points névralgiques de la chaîne. L'étude de ces phénomènes a donc connu un regain d'intérêt expérimental comme théorique.

L'existence d'un transfert de charge dans l'ADN, responsable de réactions de dégradation oxydatives, fut mis en évidence en 1962<sup>70</sup>. Ce phénomène focalisa l'attention de plusieurs communautés scientifiques. Outre l'attrait provoqué chez les biologistes par la compréhension des processus de dégradation<sup>71-75</sup> et de mutation<sup>76-78</sup> qu'il est à même d'induire, ce phénomène ne se limite pas à ces applications. L'analogie manifeste entre un

empilement de paires de bases et certains cristaux moléculaires<sup>79,80</sup> et matériaux discotiques<sup>81</sup> laisse présager une compréhension plus générale des phénomènes de transfert dans les agrégats et macromolécules à système  $\pi$ . Les technologies et l'électronique moléculaire à base d'ADN<sup>82-86</sup> peuvent également y trouver leur compte. Un grand nombre de travaux, tant théoriques qu'expérimentaux, ont tenté de caractériser les mécanismes de ce processus. Le lecteur est renvoyé aux articles de revue suivants<sup>87-91</sup> pour un panorama des techniques expérimentales employées. Théoriquement, le transfert de charge est décrit comme une compétition entre un super-échange à longue distance et une série de sauts entre sites voisins, régie par des valeurs de couplages électroniques calculées *ab initio*<sup>92-99</sup>. Ce modèle, bien qu'efficace et reconnu, n'est applicable qu'à des structures de l'ADN figées. Bien qu'il soit possible de contourner le problème à l'aide d'une statistique configurationnelle issue, par exemple, d'un calcul de dynamique moléculaire<sup>100</sup>, le calcul en temps réel de couplages électroniques raisonnables pour une géométrie quelconque de la double hélice apparaît clairement comme la pierre angulaire de toute étude novatrice future sur le sujet.

En comparaison avec celle du transfert de charge, l'étude du transfert d'excitation électronique (états singulets) dans l'ADN a relativement peu évolué au cours de la dernière trentaine d'année. Au rang des pierres d'achoppement expérimentales figure la difficulté d'obtenir des spectres d'absorption correctement résolus étant donnés les rendements quantiques de fluorescence très faibles dont font montre les bases de l'ADN<sup>101-106</sup>. Le très petit nombre d'études théoriques sur le sujet ont recours à la théorie excitonique mais pèchent par des calculs de couplages simplistes<sup>107</sup>. Le transfert d'excitation à l'état triplet n'est pas mieux caractérisé à cause de rendements quantiques de conversion intersystème et de coefficients d'extinction très faibles<sup>108-111</sup> et de termes de couplages à courte distance dont la modélisation est délicate. L'amélioration de la description du phénomène de transfert d'excitation dans l'ADN passe donc par le calcul de couplages réalistes, à géométrie figée tout d'abord, puis en incluant l'indispensable aspect dynamique.

Le but de ce travail est donc de s'efforcer de pallier les limitations des modèles de transfert de charge et d'excitation décrits ci-dessus sur l'exemple de l'ADN. Ceci nécessite la mise au point d'une méthodologie de calcul des couplages, qu'il est bien plus aisé de mettre au point sur de petits systèmes modèle plutôt que sur l'ADN. Les agrégats de molécules aromatiques remplissent le cahier des charges : plus simples à décrire que les acides nucléiques, ils en partagent les caractéristiques essentielles (géométries planes, anisotropie marquée des interactions, promotion des couplages par recouvrements de type  $\pi$ ...).

Le groupe de chimie théorique du CEA/Saclay s'est doté de potentiels modèle de bonne qualité, paramétrés à l'aide de calcul *ab initio*, permettant le calcul de l'énergie d'interaction intermoléculaire au sein d'un système avec un très bon rapport entre précision et coût en temps de calcul<sup>112</sup>. Leur efficacité a été démontrée sur des agrégats formés d'ions et de molécules d'eau<sup>113</sup>. Couplés à des méthodes d'exploration de surfaces d'énergie potentielle<sup>114</sup> ou asservis à une méthode de dynamique moléculaire classique<sup>115</sup>, ils permettent l'étude raisonnée de systèmes de taille moyenne et se positionnent en intermédiaire entre les calculs *ab initio* et les méthodes à champs de forces. Le laboratoire a également posé les bases de la description du transfert de charge au moyen d'un formalisme *valence bond*<sup>113</sup>. Il dispose qui est plus d'un certain savoir-faire quant aux calculs des états excités de grandes molécules<sup>116</sup>, et s'est intéressé, avec la collaboration d'expérimentateurs au sein du Laboratoire Francis Perrin (LFP), au transfert d'excitation dans des agrégats colonnaires<sup>117-119</sup>. Enfin, le LFP participe à un vaste programme de recherche CEA qui a pour objet la toxicologie nucléaire ; il compte y contribuer par des études théoriques et expérimentales sur les radiodommages causés dans l'ADN par le transfert de l'excitation électronique.

La présente étude est divisée en trois parties. La première est dédiée aux méthodes mises en œuvre. Méthodes de calcul de l'énergie d'un système tout d'abord, *ab initio* ou par potentiels modèle. Méthodes d'exploration des surfaces d'énergie potentielle du système ensuite, ou comment rechercher, le plus efficacement possible, les configurations optimales d'un système une fois le niveau de calcul choisi. Méthodes de dynamique moléculaire classique ou semi-classique enfin, pour simuler l'évolution au cours du temps du système sur une ou plusieurs de ces surfaces, compte tenu des conditions 'expérimentales'. La seconde partie est consacrée à l'étude du transfert de charge dans les agrégats aromatiques. L'utilisation conjointe d'un potentiel modèle spécifique et d'une approche *valence bond*, dont l'élaboration est présentée, offrent de nouvelles perspectives sur ces systèmes dont la modélisation est notoirement délicate. La troisième partie traite du transfert d'excitation dans l'ADN et les agrégats aromatiques. Les états excités collectifs de la double hélice d'ADN sont analysés en fonction de sa séquence de paires de bases et des diverses configurations qu'elle est susceptible d'adopter. L'extension de la problématique aux aspects dynamiques du transfert d'excitation est abordée avec l'aide d'une méthode de dynamique moléculaire semiclassique appliquée à des agrégats aromatiques excités.





*Première Partie*

OUTILS ET MÉTHODES



*Je résolu d'être accablant :*

*—Borges est vertigineux à ce sujet. Ne m'en veuillez pas de citer ce passage si connu des Enquêtes : 'Dans les pages lointaines de certaine encyclopédie chinoise intitulée Le Marché céleste des connaissances bénévoles, il est écrit que les animaux se divisent en a) appartenant à l'Empereur, b) embaumés, c) apprivoisés, d) cochons de lait, e) sirènes, f) fabuleux, g) chiens en liberté, h) inclus dans la présente classification, i) qui s'agitent comme des fous, j) innombrables, k) dessinés avec un très fin pinceau de poils de chameau, l) et cætera, m) qui viennent de casser la cruche, n) qui de loin semblent des mouches.' N'est-ce pas une classification qui, pour un scientifique de votre espèce, prête à sourire, sinon à rire franchement ?*

*Je pouffai de la manière la plus civilisée qui fût. Monsieur Bernardin restait de marbre.*

*—Ceci dit, je connais des gens que cela ne déride pas du tout. Et il est vrai qu'au-delà du comique de l'affaire, cet exemple illustre l'épineux problème de la démarche taxinomique. Il n'y a aucune raison de penser que nos catégories mentales soient moins absurdes que celles des Chinois.*

*Amélie Nothomb, Les Catilinaires*



# 1. Principes de Chimie Quantique

## 1.1. Introduction

L'un des plus gros obstacles de la chimie quantique demeure le nombre de particules qui doivent être traitées de manière quantique au sein d'un système chimique. La résolution de l'équation de Schrödinger pour des systèmes de grande taille nécessite en effet le traitement d'équations différentielles du second ordre faisant intervenir des potentiels comptant des centaines, voire des milliers, de variables. Les exemples abondent : les biomolécules, la structure des liquides, les réactions en surface (corrosion, catalyse hétérogène...). L'approche quantique de tels systèmes, si elle était inenvisageable il y a quelques décennies, s'est trouvée remise en question par les progrès rapides et simultanés de la technologie informatique (matérielle et algorithmique) d'une part, et des méthodologies de calcul d'autre part. Il n'en reste pas moins que, même dans le cas de systèmes simples, l'application 'à la lettre' de méthodes numériques est vouée à l'échec. Il convient au contraire d'exploiter la connaissance des aspects physico-chimiques du problème pour formuler des approximations, sans pour autant biaiser la méthodologie dans le but de retrouver un résultat connu *a priori*.

La méthode Hartree-Fock est l'exemple typique d'une telle approximation. Elle introduit le concept d'orbitales moléculaires, très utile pour la description de la structure électronique de systèmes complexes, et qui fournit souvent des résultats quantitatifs d'une grande précision.

## 1.2. L'approximation de Born – Oppenheimer

La séparation des variables est une méthode classique de simplification de l'équation de Schrödinger. Celle-ci fournit parfois des solutions exactes (séparation des coordonnées spatiales et temporelles dans un système stationnaire...), mais ce n'est souvent pas le cas. La procédure doit alors être considérée comme une approximation visant à simplifier le traitement numérique du problème. Il en est ainsi de l'approximation de Born-Oppenheimer<sup>6</sup>, qui postule la séparation des mouvements électroniques et nucléaires.

Invoquant le rapport de masse entre électrons et protons ou neutrons, elle suppose que le mouvement des électrons s'adapte infiniment rapidement à celui des noyaux.

L'hamiltonien non relativiste d'un système de particules chargées (électrons et noyaux) peut être écrit :

$$\hat{H} = \hat{T}_{\text{noy}} + \hat{H}_{\text{el}}$$

$$\hat{T}_{\text{noy}} = -\sum_{\lambda} \frac{1}{2m_{\lambda}} \nabla_{\lambda}^2 \text{ et } \hat{H}_{\text{el}} = -\sum_i \frac{1}{2} \nabla_i^2 - \sum_{\lambda,i} \frac{Z_{\lambda}}{|\vec{r}_{\lambda} - \vec{r}_i|} + \sum_{i<j} \frac{1}{|\vec{r}_i - \vec{r}_j|} + \sum_{\lambda<\mu} \frac{Z_{\lambda}Z_{\mu}}{|\vec{r}_{\lambda} - \vec{r}_{\mu}|}$$

où  $\lambda$  et  $\mu$  sont des noyaux,  $i$  et  $j$  des électrons,  $m_x$  la masse de la particule  $x$ ,  $\vec{r}_x$  son rayon vecteur, et  $Z_x$  la charge du noyau  $x$ . Il apparaît clairement que l'hamiltonien électronique dépend de la position des noyaux de façon paramétrique. En revanche, ce même hamiltonien ne dépend pas des dérivées des coordonnées des noyaux, de sorte que le problème peut être résolu en considérant les noyaux momentanément fixes dans l'espace :

$$\hat{H}_{\text{el}}(\mathbf{r}, \mathbf{R}) \Psi_{\text{el}}(\mathbf{r}, \mathbf{R}) = E_{\text{el}}(\mathbf{R}) \Psi_{\text{el}}(\mathbf{r}, \mathbf{R})$$

où  $\mathbf{r}$  et  $\mathbf{R}$  dénotent l'intégralité des coordonnées électroniques et nucléaires, respectivement. La fonction d'onde totale est alors approximée par :

$$\Psi_{\text{BO}}(\mathbf{r}, \mathbf{R}) = \Psi_{\text{noy}}(\mathbf{R}) \Psi_{\text{el}}(\mathbf{r}, \mathbf{R})$$

où la fonction d'onde nucléaire  $\Psi_{\text{noy}}(\mathbf{R})$  est solution de l'équation :

$$\left( \hat{T}_{\text{noy}}(\mathbf{R}) + E_{\text{el}}(\mathbf{R}) \right) \Psi_{\text{noy}}(\mathbf{R}) = E \Psi_{\text{noy}}(\mathbf{R})$$

$E$  étant l'énergie totale du système. La fonction d'onde électronique dépend toujours des coordonnées des noyaux, et est donc affectée par l'opérateur énergie cinétique de ceux-ci. Cependant, cette dépendance est généralement négligeable en raison du rapport de masse entre noyaux et électrons, qui a pour conséquence que les premiers se meuvent bien plus lentement que les seconds. Il est donc possible d'écrire :

$$\hat{H} \Psi_{\text{BO}} \approx \Psi_{\text{el}}(\mathbf{r}, \mathbf{R}) \left( \hat{T}_{\text{noy}}(\mathbf{R}) + E_{\text{el}}(\mathbf{R}) \right) \Psi_{\text{noy}}(\mathbf{R}) = \Psi_{\text{el}}(\mathbf{r}, \mathbf{R}) E \Psi_{\text{noy}}(\mathbf{R}) = E \Psi_{\text{BO}}$$

L'approximation de Born-Oppenheimer revêt une importance capitale en mécanique quantique, puisqu'elle sous-tend des notions telles que la structure moléculaire ou les surfaces d'énergie potentielle. En revanche, elle ne permet pas à elle seule de résoudre numériquement l'équation de Schrödinger électronique, qui demeure bien trop complexe.

### 1.3. Les déterminants de Slater

## 1. Principes de Chimie Quantique

La méthode de séparation des variables abordée ci-dessus peut également être envisagée pour élaborer une fonction d'onde d'essai pour un système polyélectronique, à partir de fonctions d'ondes monoélectroniques dénommées orbitales.

Les observables du système, et en particulier la probabilité de présence des électrons fournie par le carré de la fonction d'onde, ne doivent pas se modifier dans l'échange de deux électrons, ces particules étant indiscernables. Les électrons étant des fermions, la fonction d'onde doit de plus être antisymétrique dans l'échange de deux particules. Enfin, le principe de Pauli requiert de deux électrons qu'ils diffèrent au moins par un nombre quantique. Toutes ces propriétés se trouvent satisfaites si la fonction d'onde du système est écrite sous la forme d'un déterminant de Slater :

$$\Psi(1,2,\dots,n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_n(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_n(2) \\ \dots & \dots & \dots & \dots \\ \phi_1(n) & \phi_2(n) & \dots & \phi_n(n) \end{vmatrix}$$

qui s'exprime usuellement sous la forme plus compacte:

$$|\Psi(1,2,\dots,n)\rangle = |\phi_1(1)\phi_2(2)\dots\phi_n(n)\rangle$$

où la fonction  $\phi_i(j)$  décrit l'électron  $j$  placé dans la spin-orbitale  $i$ . Toute combinaison linéaire de déterminants de Slater vérifie également les propriétés susdites et peut en conséquence être employée pour écrire la fonction d'onde à  $n$  électrons.

### 1.4. Théorème variationnel

Bon nombre de méthodes de la chimie quantique sont dites variationnelles, puisqu'elles ont pour base le théorème variationnel qui fixe une borne inférieure à l'énergie de l'état fondamental d'un système. Soit en effet  $\Psi$  une fonction des coordonnées  $x_i$  des particules du système et de paramètres préalablement choisis  $\lambda_i$ , et soit  $\hat{H}$  l'hamiltonien de ce système, de plus basse valeur propre  $E_1$ . Le théorème variationnel stipule :

$$\frac{\langle \Psi(x_i, \lambda_i) | \hat{H} | \Psi(x_i, \lambda_i) \rangle}{\langle \Psi(x_i, \lambda_i) | \Psi(x_i, \lambda_i) \rangle} = E(x_i, \lambda_i) \geq E_1$$

La résolution approchée de l'équation de Schrödinger se ramène donc à un problème d'optimisation des paramètres  $\lambda_i$  afin d'approcher, autant que possible, la borne

inférieure de l'énergie du système. Selon les méthodes, la nature et le nombre des paramètres est variable.

## 1.5. Théorie de Hartree-Fock

L'approximation de Hartree-Fock constitue la mise en pratique la plus simple du cadre conceptuel précédent. L'état fondamental du système est décrit à l'aide d'un unique déterminant de Slater dépendant paramétriquement des spin-orbitales qui le composent. En vertu du théorème variationnel, le choix optimal des paramètres  $\phi_i$  pour la fonction d'onde  $|\Psi_0(1,2,\dots,n)\rangle = |\phi_1(1)\phi_2(2)\dots\phi_n(n)\rangle$  est celui qui minimise

$$E_0 = \frac{\langle \Psi_0 | \hat{H} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle}$$

L'hamiltonien électronique du système s'écrit :

$$\hat{H} = -\sum_i \frac{1}{2} \nabla_i^2 - \sum_{\lambda,i} \frac{Z_\lambda}{r_{i\lambda}} + \sum_{i<j} \frac{1}{r_{ij}}$$

ce qui implique :

$$E_0 = \sum_i \left\langle \phi_i \left| -\frac{1}{2} \nabla_i^2 - \sum_{\lambda} \frac{Z_\lambda}{r_{i\lambda}} \right| \phi_i \right\rangle + \sum_{i<j} \left( \left\langle \phi_i \phi_j \left| \frac{1}{r_{ij}} \right| \phi_i \phi_j \right\rangle - \left\langle \phi_i \phi_j \left| \frac{1}{r_{ij}} \right| \phi_j \phi_i \right\rangle \right)$$

L'opérateur monoélectronique  $h$  (dit de cœur) décrit l'électron  $i$  en omettant son interaction avec les autres électrons :

$$h_i = -\frac{1}{2} \nabla_i^2 - \sum_{\lambda} \frac{Z_\lambda}{r_{i\lambda}}$$

Deux opérateurs interviennent dans la partie biélectronique : l'opérateur de Coulomb  $J$ , qui multiplie  $\phi$  par l'énergie potentielle d'un électron en  $r_1$  dans le champ créée par la distribution de charge  $|\phi_i|^2$  :

$$J_i(1)\phi(1) = \left[ \int \phi_i^*(2) \frac{1}{r_{12}} \phi_i(2) dr_2 \right] \phi(1)$$

et l'opérateur d'échange  $K$

$$K_i(1)\phi(1) = \left[ \int \phi_i^*(2) \frac{1}{r_{12}} \phi(2) dr_2 \right] \phi_i(1)$$

## 1. Principes de Chimie Quantique

qui, lié à la nature antisymétrique des déterminants de Slater, n'a pas d'équivalent classique. Contrairement à l'opérateur de Coulomb, il s'agit d'un opérateur non local puisque l'action de  $K_i$  en  $r_1$  sur  $\phi$  dépend de la valeur de  $\phi$  sur tout l'espace. Les éléments de matrice des opérateurs  $J_i$  et  $K_i$  sont :

$$J_{ij} = \langle \phi_j(1) | J_i(1) | \phi_j(1) \rangle = \int \phi_j^*(1) \phi_j(1) \frac{1}{r_{12}} \phi_i^*(2) \phi_i(2) dr_2$$

$$K_{ij} = \langle \phi_j(1) | K_i(1) | \phi_j(1) \rangle = \int \phi_j^*(1) \phi_i(1) \frac{1}{r_{12}} \phi_i^*(2) \phi_j(2) dr_2$$

L'énergie du système peut s'écrire  $E_0 = \sum_i \left( h_{ii} + \sum_{i < j} (J_{ij} - K_{ij}) \right)$ . Définissant

l'opérateur de Fock  $F$  par :

$$F = h + \sum_i (J_i - K_i)$$

l'équation de Schrödinger dans l'approximation de Hartree-Fock s'écrit :

$$F_i \phi_i = \varepsilon_i \phi_i$$

où chaque spin-orbitale  $\phi_i$  est fonction propre de l'opérateur de Fock, associée à la valeur propre  $\varepsilon_i$ . L'approximation LCAO (*Linear Combination of Atomic Orbitals*) est alors invoquée pour développer chacune des spin-orbitales dans une base d'orbitales atomiques  $\{\chi_k\}$  préalablement choisie :

$$\phi_i = \sum_k c_{ik} \chi_k$$

Remplaçant dans l'équation de Hartree-Fock, le système linéaire suivant est obtenu :

$$\sum_k c_{ik} \langle \chi_p | F | \chi_k \rangle = \sum_k c_{ik} \varepsilon_i \langle \chi_p | \chi_k \rangle \quad \forall \chi_p$$

c'est-à-dire :

$$F_{pk} - \varepsilon_i S_{pk} = 0$$

Si  $\|F_{pk} - \varepsilon_i S_{pk}\| = 0$ , le système admet des solutions non triviales. Il est résolu de façon itérative :  $F_{pk}^0$  est tout d'abord calculé connaissant un jeu de spin-orbitales  $\phi_i^0$  (dit d'essai) préalablement choisi. Il est alors possible de déterminer un nouveau jeu d'orbitales  $\phi_i^1$ . Le processus se poursuit jusqu'à ce que soit déterminés  $\phi_i$  tels que  $\phi_i = \lim_{n \rightarrow \infty} (\phi_i^n)$ ,

avec une précision fixée *a priori* : c'est la méthode du champ autocohérent (Self Consistent Field).

La méthode Hartree-Fock constitue souvent une très bonne approximation d'ordre zéro, pour un coût modique puisque le temps de calcul se comporte en général comme  $N^3$  ou  $N^4$ ,  $N$  représentant le nombre d'orbitales atomiques qui constituent la base. Elle fournit ainsi des énergies totales dont la précision avoisine 1%. Malheureusement, la contribution d'une liaison covalente à l'énergie totale d'un système est justement de cet ordre de grandeur ; elle ne pourra donc être déterminée de manière fiable par différence entre les énergies Hartree-Fock d'un système et des fragments qui le constituent.

La théorie de Hartree-Fock place chaque électron dans une spin-orbitale, sous l'effet du champ moyen crée par les autres électrons. Elle constitue en cela une approche à particules indépendantes et ne retranscrit pas les phénomènes liés à la corrélation entre les électrons, à l'exception du principe de Pauli introduit par construction dans la fonction d'onde (un seul électron par spin-orbitale ; les électrons de même spin sont donc partiellement corrélés : c'est le trou de Fermi). Dans les faits, les électrons ont tendance à s'éviter : à tout instant, il est plus favorable énergétiquement de les trouver éloignés que proches les uns des autres (c'est le trou de Coulomb). L'absence de ce phénomène de corrélation dynamique nuit à la description des systèmes aux courtes distances d'approche, et a pour conséquence une sous-estimation générale des énergies de liaison. La corrélation non dynamique, quant à elle, résulte de l'inadéquation de l'approche monodéterminante à longue distance et se traduit, entre autres, par des énergies de dissociations et des natures de fragments erronées. Les méthodes post Hartree-Fock détaillées plus loin permettent, au prix d'un compromis entre précision et temps de calcul, de pallier ces manquements.

## 1.6. Méthodes post Hartree-Fock

### *1.6.1. Méthodes de Perturbation*

Soit à résoudre l'équation de Schrödinger indépendante du temps pour un système chimique stationnaire lié :

$$\hat{H}\Psi_n = E_n \Psi_n$$

## 1. Principes de Chimie Quantique

Soit de plus un hamiltonien  $\hat{H}^0$  dont les solutions sont connues :  $\hat{H}^0\Psi_n^0 = E_n^0\Psi_n^0$ .  $\hat{H}$  est supposé ne différer de  $\hat{H}^0$  que par une perturbation  $\hat{H}'$ , elle-même modulée par un paramètre réel  $\lambda$  :

$$\hat{H} = \hat{H}^0 + \lambda\hat{H}'$$

Ce paramètre assure la transition entre les systèmes non perturbé ( $\lambda=0$ ) et perturbé ( $\lambda=1$ ) ; les valeurs et fonctions propres de  $\hat{H}$  en dépendent toutes deux. Il suffit alors de développer  $E_n$  et  $\Psi_n$  en séries de Taylor de  $\lambda$  ; les coefficients des polynomes obtenus s'expriment en fonction de la perturbation  $\hat{H}'$  et des fonctions propres d'ordre zéro  $\Psi_n^0$ . Ils fournissent, à chaque ordre, une correction à l'énergie et aux fonctions d'onde du système non perturbé.

Les méthodes de perturbation de type Møller-Plesset à l'ordre  $n$  (MPn) utilisent l'hamiltonien Hartree-Fock comme référence d'ordre zéro :

$$\hat{H}^0 = \sum_i F(i) = \sum_i \left( h(i) + \sum_{j>i} (J_j - K_j) \right)$$

C'est-à-dire :

$$\hat{H}' = \sum_i \sum_{j>i} \frac{1}{r_{ij}} - \sum_i \sum_{j>i} (J_j - K_j)$$

Les fonctions propres de  $\hat{H}^0$  étant orthonormées :

$$\langle \Psi_k^0 | \hat{H}^0 | \Psi_l^0 \rangle = E_k \delta_{kl}$$

Donc :

$$\langle \Psi_k^0 | \hat{H} | \Psi_l^0 \rangle = \langle \Psi_k^0 | \hat{H}^0 | \Psi_l^0 \rangle + \langle \Psi_k^0 | \hat{H}' | \Psi_l^0 \rangle = E_k \delta_{kl} + \langle \Psi_k^0 | \hat{H}' | \Psi_l^0 \rangle$$

Dans la base des fonctions propres de l'hamiltonien d'ordre zéro, les éléments de matrice non diagonaux ( $k \neq l$ ) sont égaux. Les corrections à l'énergie fournies par la théorie des perturbations sont alors, aux deux premiers ordres :

$$E^1 = \langle \Psi^0 | \hat{H}' | \Psi^0 \rangle$$

$$E^2 = \sum_{j \neq 0} \frac{|\langle \Psi^0 | \hat{H}' | \Psi_j \rangle|^2}{E_0 - E_j}$$

où  $\Psi^j$  représente un état excité du système. Il est possible de montrer que seules les diexcitations ont une influence sur l'expression de  $E^2$  et que, de surcroît, la correction à l'énergie est toujours négative.

Le développement en série de Taylor peut être mené jusqu'à un ordre  $n$  quelconque. Néanmoins, le meilleur compromis entre précision et temps de calcul est généralement atteint pour  $n = 2$  (méthode MP2, dont le temps de calcul varie comme  $N^5$ , où  $N$  est le nombre de fonctions de base). Les méthodes MP3 et MP4 sont beaucoup plus rarement employées.

Les méthodes décrites ci-dessus traitent la corrélation comme une perturbation à la fonction d'onde Hartree-Fock. Cette dernière doit donc être une bonne approximation d'ordre zéro ; en conséquence, les méthodes  $MP_n$  sont inapplicables lorsque cette condition n'est pas réalisée, comme c'est par exemple le cas pour décrire la rupture homolytique d'une liaison en deux systèmes à couche ouverte.

### 1.6.2. Interaction de Configurations

L'approche multidéterminantale est une autre façon de pallier les déficiences de la méthode Hartree-Fock. La fonction d'onde est développée, non plus sur un seul déterminant de Slater, mais sur tous ou une partie des déterminants de Slater obtenus par répartition des  $N$  électrons du système dans un ensemble de  $P$  spin-orbitales orthogonales.

La fonction d'onde Hartree-Fock est tout d'abord déterminée, dans une base d'orbitales atomiques  $\{\chi_p, p \in \{1 \dots P\}\}$ . Les  $P$  orbitales moléculaires ainsi obtenues sont remplies par  $N$  électrons. L'état fondamental  $\Psi_0$  du système (supposé singulet) est obtenu en remplissant chacune des  $N/2$  orbitales de plus basse énergie par deux électrons chacune (elles sont dites occupées), les  $P-N/2$  restantes restant inoccupées (ou virtuelles). Les autres configurations possibles (il y a  $C_N^{2P}$  façons de placer  $N$  électrons dans  $2P$  spin-orbitales) constituent des états électroniques excités. La fonction d'onde s'écrit alors comme combinaison linéaire de l'ensemble des déterminants de Slater  $D_k$  associés à ces configurations, affectés de poids réels  $c_k$  qui constituent les inconnues du problème et dont la détermination exploite le théorème variationnel :

$$\Psi = \sum_k c_k D_k$$

Une interaction de configuration est dite totale lorsque la fonction d'onde est développée sur l'ensemble des  $C_N^{2P}$  déterminants. Dans les faits, ce nombre est si grand que même dans le cas de petits systèmes dans des bases de taille modérée, il faut avoir recours à

## 1. Principes de Chimie Quantique

une troncature de la fonction d'essai n'utilisant qu'un sous-ensemble des déterminants possibles. À cet effet, les déterminants sont classés en fonction du nombre de spin-orbitales dont l'occupation diffère par rapport à la configuration  $\Psi_0$  : un état  $n$ -excité en diffèrera par  $n$  spin-orbitales. La troncature peut alors concerner tous les états excités d'ordre supérieur à une limite fixée. Il est aisé de montrer que les états triexcités et au-delà n'ont aucun couplage direct avec l'état fondamental ; en conséquence, il est fréquent de restreindre une interaction de configurations aux mono- et diexcitations lors de l'étude de l'état fondamental d'un système. La corrélation électronique par rapport à l'état fondamental est alors totalement retranscrite, pour un coût en temps de calcul proportionnel à  $N^6$ . D'autres réductions de l'espace actif sont possibles. Elles peuvent par exemple être fondées sur une connaissance *a priori* de la nature chimique du système et des propriétés à l'étude. Des méthodes de tri systématique des configurations, visant à réduire la dimension de l'espace actif, existent également ; par exemple, l'algorithme CIPSI présenté ci-après.

### 1.6.3. Méthodes Coupled Cluster

Les méthodes d'interaction de configurations tronquées pèchent par le fait qu'elle ne sont pas  $N$ -cohérentes<sup>120</sup> (*size-consistent*). L'énergie d'un système AB obtenue par une interaction de configurations restreinte aux simples et doubles excitations (CISD) n'est pas égale, lorsque la distance A-B tend vers l'infini, à la somme des énergies CISD des monomères A et B considérés séparément. Cela implique en effet que la fonction d'onde CISD de AB tende vers le produit des fonction d'onde CISD de A et de B, qui contient des tri- et quadriexcitations par rapport à l'état fondamental de AB que ne comporte pas la fonction d'onde CISD de AB. En vertu du théorème variationnel, la somme des énergies de A et de B sera donc inférieure à l'énergie de AB à distance infinie.

Les méthodes *coupled cluster* permettent, au prix d'un temps de calcul certes élevé ( $N^7$  ou  $N^8$ ), de rendre compte de cet effet de *size-consistency* en incluant les tri- voire quadriexcitations par rapport à l'état fondamental. Il en résulte généralement une bonne retranscription de la corrélation dynamique. Les méthodes CCSD(T), dans lesquelles les triexcitations sont traitées par perturbation, offrent probablement le meilleur rapport prix-précision.

### 1.6.4. Méthodes MCSCF

De manière générale, la fonction d'onde dans une méthode multiconfigurationnelle s'écrit comme une combinaison linéaire d'un ensemble de configurations donné  $\{\psi_k\}$ , construites à l'aide d'un jeu d'orbitales moléculaires orthogonales :

$$\Psi = \sum_k a_k \psi_k$$

La méthode MCSCF consiste à optimiser non seulement les coefficients d'interaction de configurations  $a_k$ , mais également le jeu d'orbitales moléculaires sous-tendant l'ensemble  $\{\psi_k\}$ . Elle constitue donc une généralisation des méthodes Hartree-Fock et d'interaction de configurations présentées ci-dessus. En pratique, le nombre de déterminants intervenant dans le développement de la fonction d'onde est limité. La méthode la plus utilisée, dite CASSCF (*Complete Active Space Self Consistent Field*), génère l'ensemble complet des déterminants qu'il est possible d'obtenir en répartissant un nombre déterminé d'électrons dans un espace d'orbitales dites actives. Cet espace actif, typiquement limité à une petite dizaine d'orbitales, est complété d'une part par un espace inactif, composé d'orbitales optimisées mais dont l'occupation (fixée à deux) ne varie pas d'un déterminant à l'autre, et d'autre part par un espace externe, à occupation fixe (deux ou zéro électrons) et non optimisé. L'espace actif rassemble le plus souvent les orbitales moléculaires occupées de haute énergie et 'virtuelles' de basse énergie, qui sont d'intérêt pour la chimie traditionnelle. Il en va ainsi du CASSCF dit de valence, dont l'espace actif contient les orbitales occupées de valence et les orbitales virtuelles associées. Le fait d'utiliser un espace actif complet assure, sous réserve que celui-ci soit correctement sous-tendu, que la condition de N-cohérence abordée ci-dessus est vérifiée ; le prix en est le nombre souvent très important de déterminants à inclure dans le développement d'interaction de configuration, dont la majorité est affectée de très faibles poids. La méthode RASSCF (*Restricted Active Space Self Consistent Field*) permet une sélection plus fine de l'espace actif, qui n'est alors plus complet. Enfin, les fonctions d'onde CASSCF ou RASSCF peuvent servir de socle à des traitements plus poussés de la corrélation non dynamique (CASPT2 ou MRCI).

### 1.6.5. Méthode CIPSI

## 1. Principes de Chimie Quantique

La méthode CIPSI (*Configuration Interaction by Perturbative Selected Iterations*)<sup>121,122</sup> est une méthode de perturbation multiréférentielle d'ordre deux, au cours de laquelle les configurations de poids importants sont traitées variationnellement au sein d'un sous-espace construit itérativement. La fonction d'onde obtenue par diagonalisation de la matrice d'interaction de configuration est perturbée par l'ensemble des déterminants qui diffèrent d'une ou de deux spin-orbitales par rapport aux déterminants de référence. Ceux d'entre eux qui participent à la fonction d'onde perturbée au premier ordre au-delà d'un seuil fixé sont inclus, de manière itérative, au sein du sous-espace variationnel. La norme de la correction au premier ordre à la fonction d'onde d'ordre zéro est une jauge de sa qualité : elle est d'autant plus grande que l'influence de la contribution perturbative est importante. À chaque déterminant, candidat à l'incorporation dans l'espace actif, est affecté un poids égal à sa contribution à la norme de la perturbation. Tout déterminant dont le poids est supérieur au seuil précédemment évoqué sera inclus dans l'espace actif lors de l'itération suivante.

Le choix du seuil est empirique. Il doit être suffisamment petit pour que tous les états aient une norme de perturbation similaire lorsque la convergence du processus est atteinte. Cependant, le choix d'un seuil trop petit provoque l'inclusion dans le calcul d'un nombre très important de déterminants, ce qui a pour effet de rendre le calcul irréalisable dans la pratique.

### 1.7. Méthodes semi-empiriques

Le coût en temps de calcul de la méthode Hartree-Fock est proportionnel au nombre de fonctions de base à la puissance quatre, en raison du nombre d'intégrales biélectroniques dont l'évaluation est nécessaire lors de la construction de la matrice de Fock. Les méthodes semi-empiriques ont recours à des approximations afin de limiter le nombre de ces intégrales.

La première étape est de ne traiter explicitement que les électrons de valence, le comportement des électrons de cœur étant simulé par réduction de la charge nucléaire, ou à l'aide d'une fonction combinant la modélisation de la répulsion entre noyaux et entre électrons de cœur. En second lieu, des bases minimales de fonctions de type Slater sont employées. Enfin, les méthodes semi-empiriques ont recours à l'approximation dite de *Zero Differential Overlap* (ZDO)<sup>123</sup> : si  $\mu_A$  et  $\nu_B$  sont deux orbitales atomiques centrées

respectivement sur les atome A et B, alors le produit  $\mu_A \cdot \nu_B$  vaut zéro. Les implications sont les suivantes : (i) la matrice de recouvrement entre orbitales atomiques est la matrice identité ; (ii) les intégrales monoélectroniques impliquant trois centres sont nulles ; (iii) les intégrales biélectroniques à trois et quatre centres, de loin les plus nombreuses, sont négligées. En contrepartie, les intégrales restantes peuvent être calibrées à l'aide de paramètres fixés en comparaison avec des calculs ab initio ou des résultats expérimentaux. Le nombre et la nature de ces paramètres dépend de la méthode semi-empirique elle-même.

### 1.7.1. Méthode CS-INDO

La méthode CS-INDO<sup>124</sup> est dérivée de la méthode INDO (*Intermediate Neglect of Differential Overlap*)<sup>125</sup>. En sus de l'approximation ZDO sur laquelle elle est fondée, celle-ci suppose toute intégrale biélectronique bicentrique négligeable si elle n'est pas coulombienne. Elle impose également que toute intégrale monoélectronique impliquant deux orbitales atomiques centrées sur le même atome mais dont l'opérateur concerne un autre centre soit nulle. Par rapport à sa devancière, la méthode CS-INDO améliore la paramétrisation des intégrales de résonance (les intégrales monoélectroniques d'interaction électron-noyau entre deux orbitales atomiques portées par deux atomes distincts). Ces intégrales dépendent à présent d'un paramètre censé rendre compte du type des deux orbitales en interaction : hybrides  $p_\pi$ , perpendiculaires au plan d'une molécule aromatique, hybrides  $\sigma$ , orientées selon les liaisons chimiques, et paires non liantes n.

La méthode CS-INDO a connu certains succès dans la description des premiers états excités de molécules aromatiques, particulièrement adaptées à l'usage d'orbitales hybrides<sup>116,126,127</sup>.

### 1.7.2. Méthode AM1

La méthode AM1 (*Austin Model 1*)<sup>128</sup> est une application de la méthode ZDO dont les paramètres sont de nature atomique. Elle possède en outre la spécificité d'inclure des fonctions gaussiennes dans les potentiels d'interaction cœur-cœur, ce qui a pour effet de corriger la répulsion exagérée à courte distance qui caractérise d'autres méthodes du même type. Les géométries de grandes molécules, optimisées au niveau AM1, sont souvent remarquablement précises.

### 1.8. Le modèle *valence bond*

Le modèle *valence bond* constitue une généralisation du travail de Heitler et London qui fut, en 1927, la première description quantique de la molécule de dihydrogène<sup>129</sup>. Proche de la conception du chimiste, elle décrit une molécule par ses formes mésomères, chacune d'entre elles consistant d'une assemblée d'atomes reliés par des liaisons localisées résultant de la mise en commun d'électrons de valence appartenant aux atomes. Soit une molécule diatomique AB, et soient  $\chi_A$  et  $\chi_B$  les orbitales atomiques contenant les électrons impliqués dans la liaison. Il existe une unique façon (au facteur de normalisation près) d'associer les électrons par paires tout en respectant le principe de Pauli et l'antisymétrie dans l'échange de deux électrons :

$$\Psi_{VB} = \left| \chi_A \overline{\chi_B} \right| - \left| \overline{\chi_A} \chi_B \right|$$

La fonction d'onde ainsi obtenue est de caractère purement covalent. Il est en outre possible, en théorie des orbitales moléculaires, de décrire le système à l'aide d'une orbitale moléculaire liante  $\sigma$ , doublement occupée, formée par combinaison linéaire de  $\chi_A$  et  $\chi_B$  :

$$\begin{aligned} \Psi_{OM} &= |\sigma\overline{\sigma}| = \left| (\chi_A + \chi_B)(\overline{\chi_A + \chi_B}) \right| \\ &= \underbrace{\left| \chi_A \overline{\chi_B} \right| - \left| \overline{\chi_A} \chi_B \right|}_{\text{forme covalente}} + \underbrace{\left| \chi_A \overline{\chi_A} \right| + \left| \chi_B \overline{\chi_B} \right|}_{\text{formes ioniques}} \end{aligned}$$

Cette fonction d'onde est pour moitié ionique, pour moitié covalente. Ni l'une ni l'autre de ces deux descriptions n'est correcte, le rapport covalent / ionique s'établissant à 0,8 / 0,2 aux distances d'équilibre dans le cas d'une liaison homonucléaire<sup>130</sup>. Le traitement de Heitler et London peut être généralisé de façon à moduler ce rapport :

$$\Psi_{HL} = \lambda \left( \left| \chi_A \overline{\chi_B} \right| - \left| \overline{\chi_A} \chi_B \right| \right) + \mu \left( \left| \chi_A \overline{\chi_A} \right| + \left| \chi_B \overline{\chi_B} \right| \right)$$

La méthode *valence bond*, pourtant intuitivement plaisante, est desservie par le fait que les fonctions d'onde monoélectroniques servant à décrire les formes mésomères ne sont pas orthogonales entre elles, ce qui complique considérablement le calcul des intégrales biélectroniques. Le nombre très important de formes *valence bond* à considérer pour des systèmes de taille conséquente constitue également un handicap majeur. Pour autant, l'utilisation de cette méthode pour de petits systèmes n'est pas dénuée d'intérêt, notamment aux longues distances d'interaction.

## 1.9. Fonctions de base

La majorité des méthodes de chimie quantique moléculaire décrivent la fonction d'onde électronique d'un système par le biais d'une base de fonctions. Il est naturel d'employer à cet effet des orbitales atomiques  $\chi$  centrées sur chacun des atomes, produit d'une fonction radiale  $R$  et d'une harmonique sphérique  $Y_l^m$  :

$$\chi(a) = R(r_a)Y_l^m(\theta_a, \phi_a)$$

où  $a$  est l'atome sur laquelle l'orbitale atomique est centrée. La forme exacte de la fonction radiale ne peut être déterminée que dans le cas d'un atome monoélectronique. Il est en revanche possible d'en caractériser certaines propriétés :

- la fonction  $R$  est continue en 0 (au noyau) mais  $\frac{\partial R}{\partial r_a}$  ne l'est pas : la fonction présente une cuspide (*cusp*) ;
- la solution exacte pour un atome monoélectronique a la forme d'une exponentielle décroissante.

Ces propriétés sont satisfaites par les orbitales dites de Slater (*Slater-type orbital* ou STO), dont la partie radiale est de la forme :

$$R(r_a) = N r_a^n e^{-\zeta r_a}$$

$n$  étant un entier positif ou nul,  $\zeta$  une constante réelle positive et  $N$  un facteur de normalisation lui aussi réel. La fonction d'onde d'un système peut dès lors être décomposée dans une base d'orbitales atomiques, elles-mêmes combinaisons linéaires d'une ou plusieurs STO.

Un problème apparaît cependant dans le cas de molécules polyatomiques. Un nombre d'atomes supérieur ou égal à quatre implique l'évaluation d'intégrales à quatre centres ( $rs|tu$ ) qui sont au nombre d'environ  $N^4/8$  dans un traitement SCF,  $N$  étant la dimension de la base. Or, le calcul des intégrales à trois et quatre centres entre STO est numérique donc très coûteux, ce qui limite l'intérêt des bases de STO aux systèmes de petite taille.

Boys proposa en 1950 de pallier les défauts des STO en employant des orbitales gaussiennes (*Gaussian-type orbital* ou GTO). La fonction radiale d'une gaussienne cartésienne prend la forme :

## 1. Principes de Chimie Quantique

$$R(r_a) = N x_a^i y_a^j z_a^k e^{-\alpha r_a}$$

$N$  étant une constante de normalisation réelle,  $i, j$  et  $k$  des entiers positifs ou nuls qui vérifient  $i + j + k = l$  où  $l$  est le nombre quantique orbital, et  $\alpha$  une constante réelle positive. Il existe par définition de  $i, j$  et  $k$  une gaussienne de type  $s$  ( $l = 0$ ), trois de type  $p$  ( $l = 1$ ), six de type  $d$  ( $l = 2$ )... L'absence remarquable du nombre quantique principal  $n$  dans la définition ci-dessus implique que des gaussiennes de même  $l$  ne sont différenciées que par la valeur de l'exposant  $\alpha$ . Les gaussiennes sphériques, combinaison linéaire des gaussiennes cartésiennes, ont le même comportement angulaire que les orbitales réelles et sont en conséquence fréquemment employées.

Les fonctions gaussiennes, ne possédant pas la cuspide requise à l'origine, offrent en général une piètre représentation d'une orbitale atomique aux petites valeurs de  $r_a$ . Celle-ci peut être améliorée en utilisant une combinaison linéaire de gaussiennes pour rendre compte d'une seule orbitale atomique. De sorte, lors d'un traitement SCF, il faudra évaluer bien plus d'intégrales entre GTO qu'entre STO à qualité de calcul égale (environ huit fois davantage). Cependant, l'évaluation d'intégrales entre GTO est bien moins coûteuse qu'entre STO, puisque le produit de deux gaussiennes de centres distincts est une gaussienne centrée en un troisième point ; les intégrales à trois ou quatre centres se réduisent donc à des intégrales bicentriques.

Il est donc d'usage d'employer un ensemble de fonctions de base  $\{\chi_i\}$  dont chacune s'exprime comme combinaison linéaire à coefficients  $d_{ij}$  constants d'un petit nombre de gaussiennes  $\{g_j\}$ , de mêmes valeurs de  $i, j, k$  et  $l$  mais de valeurs de  $\alpha$  différentes :

$$\chi_i = \sum_j d_{ij} g_j$$

$\chi_i$  est dénommée gaussienne contractée,  $g_j$  gaussienne primitive et  $d_{ij}$  coefficient de contraction. Les coefficients de contraction étant fixés, le nombre de coefficients à optimiser au cours d'une procédure SCF est considérablement réduit par l'emploi de gaussiennes contractées.

La définition d'une gaussienne cartésienne, et la hiérarchie gouvernant la constitution d'une base, seront très utiles dans cette étude lorsqu'il conviendra de calculer des intégrales de recouvrement entre orbitales moléculaires (deuxième partie). En revanche, la nomenclature des bases de gaussiennes ne sera pas abordée ici ; elle peut être trouvée dans de nombreux textes de chimie quantique, dont <sup>131</sup>.



## 2. Potentiel Modèle Intermoléculaire

**L**e terme ‘potentiel modèle intermoléculaire’ recouvre généralement l’expression de l’énergie d’interaction entre molécules sous la forme d’une expression analytique. Les potentiels modèle développés au laboratoire prennent le parti d’un chemin médian entre les méthodes *ab initio*, précises mais coûteuses, et les potentiels plus ‘simplistes’ visant des systèmes de grande taille dont une description semi-quantitative est suffisante.

Deux approches mènent à l’élaboration d’un potentiel modèle. La première consiste à se doter d’une forme de potentiel choisie *a priori*, dépendant de paramètres déduits de calculs *ab initio* ou de résultats expérimentaux. La seconde, plus raffinée, subdivise le potentiel en termes distincts associés aux divers types d’interaction, qui sont paramétrés séparément. Si la première approche favorise la simplicité et le coût en temps de calcul, la seconde est bien plus flexible et précise. Ces deux approches ne sont d’ailleurs pas directement concurrentes puisque la seconde fournit plus d’information que la première : au prix d’un potentiel plus complexe, elle offre en effet la découpe de l’énergie d’interaction en ses composantes élémentaires, que le théoricien peut mettre à profit pour affiner son analyse des interactions intermoléculaires et de la façon dont elles influencent les phénomènes à l’étude. De cette même découpe émane également la possibilité d’employer différents niveaux d’approximation dans le traitement des différents termes, en fonction de leur nature (classique ou quantique), de leur portée (longue, moyenne ou courte), et du nombre d’entités simultanément impliquées (deux corps ou  $n$  corps). En cela, la deuxième approche prend à nouveau le pas sur la première, qui traite généralement le potentiel comme une somme de termes à deux corps (contenant implicitement et de manière non triviale tout terme à  $n$  corps) et dont la fiabilité n’est assurée que dans une gamme de distances restreinte (celle caractéristique du système à l’équilibre, dont les propriétés furent utilisées lors du paramétrage). Les paramètres de la seconde approche, bien que plus nombreux, revêtent donc un sens physique qui rend envisageable leur transférabilité à d’autres espèces. La première approche mise au contraire sur l’encapsulation du sens physique dans aussi peu de paramètres que possible, avantage compensé par le nombre de valeurs distinctes que peuvent adopter ces paramètres d’un système à un autre (puisque ils sont trop effectifs pour être transférables).

C'est généralement la taille du système à étudier qui guide le choix d'une méthode au détriment de l'autre, puisque sur ce terrain les domaines d'application des deux méthodes sont disjoints. La complexité du potentiel dicte en effet la taille maximale des systèmes envisageables, à plus forte raison lorsqu'il est nécessaire d'évaluer les dérivées du potentiel par rapport aux degrés de liberté du système. De plus, pour des systèmes de grande taille, la précision d'un potentiel grossier peut se révéler bien suffisante, notamment si l'étude est axée sur des différences entre énergies d'interaction de différentes configurations plutôt que sur la détermination d'énergies d'interaction absolues.

Le présent travail, de par la taille relativement modeste des systèmes mis en jeu et la complexité des interactions en présence, a pris le parti de la seconde approche grâce à un potentiel modèle dont la précision peut atteindre celle de calculs *ab initio*<sup>113</sup>. Il est présenté ci-dessous, et sera complété et paramétré au cours de cette étude.

## 2.1. Méthode de perturbation d'échange

Soit un système constitué de deux molécules A et B et décrit par un hamiltonien  $\hat{H}$  :

$$\hat{H} = \hat{H}_A + \hat{H}_B + \hat{H}_{AB}$$

$\hat{H}_A$  et  $\hat{H}_B$  sont les hamiltoniens des molécules A et B isolées et  $\hat{H}_{AB}$  traduit les interactions entre ces dernières. A séparation infinie,  $\hat{H} = \hat{H}_A + \hat{H}_B = \hat{H}_0$ . L'hamiltonien d'interaction peut donc être considéré comme une perturbation par rapport au système décrit par le produit des fonctions d'onde des molécules A et B isolées, et traité, comme tel, en théorie des perturbations. Le principe de Pauli pour l'échange entre électrons de A et B n'est cependant pas satisfait, de sorte que le traitement perturbatif diverge à courte distance. La théorie de perturbation d'échange pallie ce problème en introduisant dans la fonction d'onde d'ordre zéro un antisymétriseur. À l'issue du traitement, les termes intervenant dans l'expression de l'énergie d'interaction peuvent être identifiés à des composantes de sens physique distinguable :

$$E_{\text{int}} = \underbrace{E_{\text{elec}} + E_{\text{rep-ech}}}_{\text{premier ordre}} + \underbrace{E_{\text{pol}} + E_{\text{pol-ech}} + E_{\text{disp}} + E_{\text{disp-ech}}}_{\text{second ordre}}$$

$E_{\text{elec}}$  est l'énergie électrostatique,  $E_{\text{rep-ech}}$  l'énergie de répulsion d'échange,  $E_{\text{pol}}$  l'énergie de polarisation,  $E_{\text{pol-ech}}$  l'énergie de polarisation d'échange,  $E_{\text{disp}}$  l'énergie de

## 2. Potentiel Modèle Intermoléculaire

dispersion,  $E_{\text{disp-ech}}$  l'énergie de dispersion d'échange. Les termes d'échange assurent la convergence du développement à courte distance. Il est maintenant nécessaire d'associer à chaque terme une forme analytique permettant son calcul.

### 2.1.1. *Energie Electrostatique*

L'énergie électrostatique, terme à équivalent classique, est due à l'interaction électrostatique entre les noyaux et nuages électroniques des deux molécules, représentés par leurs moments permanents. Ce terme découle du premier ordre de perturbation, est actif à longue distance et peut être attractif ou repulsif. En pratique, le calcul de l'énergie électrostatique se résume à une interaction entre distributions de multipôles situées sur l'une et l'autre des deux entités en interaction. Le nombre, le degré et la localisation spatiale desdits multipôles, de même que l'observable qu'ils sont censés retranscrire, restent au choix du théoricien.

Deux philosophies existent quant à la nature des données dont sont tirées les distributions de multipôles. La première privilégie la retranscription correcte du potentiel électrostatique, la seconde exploite la connaissance de la matrice densité à une particule associée à la fonction d'onde *ab initio*<sup>132</sup>. Le nombre et le positionnement des multipôles sur les molécules, s'il est intuitif en ce qui concerne les noyaux, l'est bien moins dans le cas des nuages électroniques. Bien qu'il soit naturel de restreindre autant que possible le nombre de multipôles employés, il a été montré que la réduction d'une molécule à un unique site est inacceptable. L'anisotropie du champ électrique permanent de la molécule, plus particulièrement ressenti aux courtes distances, ne peut être correctement retranscrit de cette façon. Qui plus est, aux courtes et moyennes distances et à précision égale, le gain de temps dû à la limitation du nombre de sites est contrebalancé par l'augmentation de l'ordre des multipôles placés sur ces derniers<sup>132,133</sup>.

Le choix d'un développement multipolaire multicentrique semble donc de rigueur. La procédure retenue dans cette étude associe les orbitales atomiques par paires. De la nature et de la position de chaque orbitale sont déduits le site et l'ordre maximal du multipôle correspondant. Il a été montré qu'en exploitant les propriétés des gaussiennes, il est possible d'obtenir une représentation exacte de la densité électronique *ab initio* à l'aide d'un nombre fini de multipôles d'ordres finis. La qualité de la distribution dépend donc uniquement de la qualité du calcul source, et justifie l'importance du choix de ce dernier, en fonction du

système à étudier. Le calcul pratique de la composante électrostatique de l'énergie d'interaction nécessite cependant de travailler avec un nombre de sites d'ordre aussi limité que possible. La procédure de Vigné-Maeder et Claverie<sup>134</sup> propose, à cet égard, une troncature systématique au-delà du quadrupôle et la réduction du nombre de sites *via* la redistribution des multipôles supprimés sur ceux qui sont conservés. Si la représentation de la densité électronique n'est dès lors plus exacte, les erreurs typiques commises sur le potentiel électrostatique à courtes et longues distances (respectivement 2 % et 0,1 %) soulignent la légitimité de l'approximation. Le nombre de sites reste cependant plus important que dans la procédure de Stone<sup>133</sup>, qui lui privilégie l'ordre maximal des multipôles (hexadécapole).

### 2.1.2. Energie de Polarisation Dipolaire

L'énergie de polarisation dipolaire a pour origine l'interaction des sites polarisables d'une entité avec le champ électrique local auquel ces sites sont soumis. Celui-ci est créé, en l'absence de champ électrique externe, par les moments (permanents et induits) de l'ensemble des autres entités de l'édifice à l'étude. L'énergie de polarisation constitue en cela un terme à  $n$  corps déterminé de façon autocohérente : le champ électrique créé par l'ensemble des moments permanents est responsable de l'apparition de moments induits qui y contribuent également. Le champ électrique, donc l'énergie de polarisation, ne peuvent donc être déterminés que par itérations successives. Bien qu'il soit possible de ne tenir compte du champ qu'au premier ordre, l'autocohérence se montre dans certains systèmes d'une importance cruciale<sup>135</sup>.

L'énergie de polarisation dipolaire est reliée au champ électrique par l'intermédiaire d'un tenseur de polarisabilité qui est généralement considéré comme une caractéristique propre à chaque molécule, et dont la valeur peut être déterminée expérimentalement ou *ab initio*. Ce tenseur moléculaire est généralement anisotrope (c'est-à-dire non scalaire) : un champ électrique orienté par exemple le long de l'axe  $x$  engendrera une polarisation de composantes *a priori* non nulles le long de chacun des trois axes :

$$P_x = \alpha_{xx} E_x \quad P_y = \alpha_{yx} E_x \quad P_z = \alpha_{zx} E_x$$

Il en est de même pour les axes  $y$  et  $z$ , de sorte que :

$$\begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} = \begin{pmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \quad \text{ou} \quad \vec{P} = \boldsymbol{\alpha} \vec{E}$$

## 2. Potentiel Modèle Intermoléculaire

Il existe un repère dit principal dans lequel le tenseur  $\alpha$  est diagonal. Le cas isotrope correspond à  $\alpha_{xx}=\alpha_{yy}=\alpha_{zz}=\alpha$ <sup>136</sup>. Le cas général ( $\alpha_{xx}\neq\alpha_{yy}\neq\alpha_{zz}$ ), correspond à un tenseur dont l'anisotropie est mesurée par la grandeur  $\Delta\alpha$ <sup>137</sup>:

$$|\Delta\alpha|^2 = \frac{3\text{tr } \alpha^2 - (\text{tr } \alpha)^2}{2} \text{ dans tout repère,}$$
$$= \frac{1}{2} [(\alpha_{xx} - \alpha_{yy})^2 + (\alpha_{xx} - \alpha_{zz})^2 + (\alpha_{yy} - \alpha_{zz})^2] \text{ dans le repère principal d'inertie.}$$

Le tenseur de polarisabilité moléculaire n'est cependant pas directement exploitable pour le calcul précis de l'énergie de polarisation ; il occulte le fait que les molécules de tailles typiques peuvent consister en plusieurs sous-ensembles de propriétés différentes, dont les influences ne peuvent être moyennées aux courtes et moyennes distances d'approche (sans mentionner le problème du placement d'un tenseur moyen sur une molécule non ponctuelle...). En outre, le champ électrique n'est généralement pas constant d'une extrémité à l'autre de la molécule. Il est donc souhaitable d'introduire la notion de polarisabilité sous une forme plus locale. L'énergie de polarisation d'une entité A s'écrit :

$$E_{\text{pol}}^A = -\frac{1}{2} \sum_{\text{sites } i \in A} \alpha_i |E_i|^2$$

Chaque site  $i$  de la molécule A dispose de son tenseur  $\alpha_i$  centré en un point où règne un champ  $E_i$ . L'obtention de ces tenseurs locaux doit se faire par le biais d'un modèle puisque seul le tenseur de polarisabilité moléculaire est accessible expérimentalement. Plusieurs modèles ont été proposés à cet effet, dont le point commun réside dans une hypothèse d'additivité selon laquelle la somme des tenseurs locaux dans le repère moléculaire doit fournir un tenseur global compatible avec une valeur de référence (expérimentale ou *ab initio*).

L'étude fondatrice de Claverie<sup>132</sup> emploie un modèle mis au point par Lefèvre<sup>138</sup>, visant à prédire le tenseur de polarisabilité d'une vaste gamme de molécules de la chimie organique traditionnelle, à partir de tenseurs de sites supposés additifs et transférables. Le tenseur de polarisabilité moléculaire est tout d'abord déterminé expérimentalement, par une conjonction de trois types de mesures : polarisabilité moyenne, facteur de dépolarisation de la lumière, et biréfringence. La molécule est divisée en sites atomiques et de liaisons ; chaque type d'atome dispose d'un tenseur local scalaire, chaque type de liaison d'un tenseur, diagonal dans le repère associé à la liaison, qui présente une composante longitudinale ( $\alpha_{xx}$ ) et deux composantes transversales égales ( $\alpha_{yy}$  et  $\alpha_{zz}$ ). Ces tenseurs de sites sont supposés spécifiques à un type d'atome ou de liaison, c'est-à-dire transférables entre espèces

moléculaires. La valeur des tenseurs de sites est ajustée de façon à rendre compte du tenseur de polarisabilité de chacune des molécules du *pool* considéré.

Claverie ajoute au modèle de Lefèvre une hypothèse d'isotropie des tenseurs ; chaque tenseur de site est remplacé par le tenseur scalaire moyen  $\bar{\alpha}$  correspondant :

$$\bar{\alpha} = \frac{1}{3}(\alpha_{xx} + \alpha_{yy} + \alpha_{zz})$$

où  $\alpha_{xx}$ ,  $\alpha_{yy}$  et  $\alpha_{zz}$  sont les composantes diagonales du tenseur de site dans le repère principal qui lui est associé. La somme de ces tenseurs scalaires locaux fournit donc le tenseur moléculaire moyen, lui aussi scalaire. Si cette hypothèse supplémentaire est acceptable pour des entités de taille réduite ou globalement sphériques (ions, molécule d'eau<sup>135</sup>...), elle se montre le plus souvent grossière.

L'extension la plus intuitive consiste à rétablir l'anisotropie des tenseurs de sites dans l'acception de Lefèvre. Bien que ce traitement soit sans doute justifiable pour un certain nombre d'espèces chimiques, il ne l'est pas pour les molécules aromatiques : celles-ci affichent une anisotropie hors-plan que ne peut retranscrire un modèle à deux composantes ( $\alpha_{yy} = \alpha_{zz}$ ) tel que celui de Lefèvre. Mulder<sup>139</sup> propose une généralisation en deux étapes de ce modèle. En premier lieu, les tenseurs de liaisons disposent à présent de trois paramètres distincts : une composante longitudinale, une composante verticale (c'est-à-dire perpendiculaire au plan du cycle aromatique) et une composante transversale. En second lieu, les systèmes  $\sigma$  et  $\pi$ , qui diffèrent par leurs propriétés de polarisabilité, disposent de leurs propres jeux de paramètres. En revanche, l'hypothèse d'additivité est conservée et la philosophie d'ajustement des tenseurs de sites reste voisine de celle employée par Lefèvre, le *pool* d'espèces étant ici restreint au benzène et à l'ensemble des azines monocycliques. Cette approche a été retenue et adaptée dans le présent travail, de façon à rendre compte de l'énergie de polarisation dans les agrégats aromatiques étudiés.

### ***2.1.3. Energie de Dispersion, de Répulsion de de Dispersion-Echange***

Ces composantes de l'énergie d'interaction sont issues du second ordre du traitement de perturbation d'échange décrit précédemment, et interviennent à courte et moyenne portée. Elles s'écrivent, en première approximation, comme une somme de termes à

## 2. Potentiel Modèle Intermoléculaire

deux corps entre le site atomique  $i$  sur la molécule 1 et le site atomique  $j$  sur la molécule 2. Dans le formalisme de Kitaigorodski<sup>140</sup>, leur expression analytique est :

$$E_{\text{rep}} + E_{\text{disp}} + E_{\text{disp-ech}} = \sum_{i \in 1} \sum_{j \in 2} k_i k_j \left( G_{ij} \left( C e^{-\gamma z} + C^{\text{de}} e^{-\gamma^{\text{de}} z} \right) - \left( \frac{C_6}{z^6} + \frac{C_8}{z^8} + \frac{C_{10}}{z^{10}} \right) \right)$$

$$z = \frac{r_{ij}}{r_{ij}^0} \text{ où } r_{ij} \text{ est la distance séparant les sites } i \text{ et } j \text{ et } r_{ij}^0 = 2\sqrt{R_i R_j} \text{ la moyenne}$$

quadratique des rayons de Van der Waals<sup>141</sup> associés aux atomes  $i$  et  $j$ .

$$G_{ij} = \left( 1 - \frac{Q_i}{n_i} \right) \left( 1 - \frac{Q_j}{n_j} \right) \text{ fait intervenir le rapport entre les 'charges' de dispersion-répulsion } Q_i$$

et  $Q_j$  et le nombre  $n_i$  et  $n_j$  d'électrons de valence des atomes  $i$  et  $j$ , respectivement. Les paramètres  $C_6$ ,  $C_8$  et  $C_{10}$  gouvernent l'importance relative des différentes composantes de dispersion et sont des valeurs standards déterminées par Stogryn<sup>142</sup> pour l'interaction hélium-hélium. Enfin, les paramètres standards de répulsion  $C$ ,  $C^{\text{de}}$ ,  $\gamma$  et  $\gamma^{\text{de}}$  proviennent des travaux de Claverie, Langlet et Caillet<sup>143,144</sup>. Pour un type d'atome donné, le paramétrage s'effectue donc au moyen des deux seuls paramètres  $R$  et  $k$ .

### 2.1.4. Traitement du Transfert de Charge : Méthode Valence Bond

Le transfert de charge constitue un terme à courte portée, d'origine purement quantique. Dans un agrégat neutre  $AB$ , il a pour origine le couplage de la structure électronique neutre  $AB$  avec les formes zwitterioniques  $A^+B^-$  et  $A^-B^+$ , dont l'effet global est une stabilisation du système. Si l'influence de ce phénomène est négligeable dans les systèmes neutres à l'état électronique fondamental (du fait, entre autres, des longues distances d'équilibre), il n'en va pas ainsi dans le cas des systèmes chargés. Le potentiel modèle utilisé ici ne présente pas, contrairement à d'autres travaux, de terme analytique dû au transfert de charge<sup>145</sup>; ce phénomène est inclus *via* le formalisme *valence bond*, comme présenté plus loin dans ce travail.



## 3. Description d'un Supersystème : Méthodes à Hamiltonien Effectif

Les méthodes *ab initio* présentées plus haut, si elles assurent une description généralement fidèle des propriétés électroniques de petits systèmes, ne sont techniquement pas applicables à l'étude d'espèces de taille plus importante, particulièrement lorsqu'il est question de leurs états électroniques excités. Dans ce contexte, l'idée de segmenter le système en un nombre de constituants élémentaires dont l'étude est plus aisée, s'impose naturellement. Un agrégat peut ainsi être décomposé en un ensemble de monomères. En revanche, les états électroniques d'un agrégat ne se résument pas à la simple superposition des états correspondants des monomères le constituant. Ainsi, il est connu de longue date que les états excités d'un cristal ne sont pas directement identifiables à ceux des atomes ou molécules qui le composent<sup>146</sup>. L'explication tient à l'existence de couplages reliant les états électroniques des monomères, de sorte que la fonction d'onde d'un état  $\Psi$  du supersystème doit en réalité être écrite comme combinaison linéaire des fonctions d'onde  $\psi$  de ces états :

$$\Psi = \sum_{\text{monomères } M} \sum_{\text{états } S \in M} C_M^S \psi_M^S$$

La détermination de l'état électronique d'un supersystème se décompose donc en deux étapes : l'étude des états des monomères constitutifs d'une part, et leur mise en interaction *via* des termes de couplage d'autre part. La méthode et le niveau de calcul utilisés lors de la première étape sont intimement liés à la nature des monomères et des états électroniques (fondamental ou excités) à l'étude ; elle fournit l'ensemble des fonctions d'onde  $\{\psi_M^S\}$ . La deuxième étape s'apparente à une interaction de configuration entre ces états ; elle fournit l'ensemble des coefficients  $C_M^S$  du développement des états du supersystème sur les états de base de l'interaction de configuration. Cette procédure revient à négliger les états impliquant plus d'une entité élémentaire, c'est-à-dire essentiellement les états à transfert de charge  $A^+B^-$ .

L'hamiltonien électronique exact du supersystème peut s'écrire sous la forme d'une somme d'hamiltoniens électroniques exacts de chacun des monomères et d'un

hamiltonien d'interaction, qui couple le mouvement des électrons  $i$  du monomère  $n$  à ceux des électrons  $j$  et noyaux  $Z$  du monomère  $m$ :

$$\hat{H} = \sum_n \hat{H}_{el}^n + \hat{V}$$

$$\text{avec } \hat{V} = \sum_n \left[ \sum_{\substack{Z \in m \\ m < n}} \frac{1}{R_{iZ}} + \sum_{i \in n} \sum_{\substack{j \in m \\ m < n}} \frac{1}{r_{ij}} \right]$$

L'action de l'hamiltonien d'interaction  $\hat{V}$  est traité comme une perturbation par rapport aux états propres des hamiltoniens des monomères. Cette approche, dite perturbative, est d'un coût bien moindre que l'approche variationnelle totale, puisqu'elle remplace avantageusement une interaction de configurations de grande taille par un calcul de même type mais de taille plus réduite.

Le choix des états propres des monomères à inclure dans l'interaction de configuration est dicté par l'importance relative des écarts énergétiques entre états des monomères et les couplages qui les lient, comme le met en évidence la figure 1.1 sur un modèle de bandes :

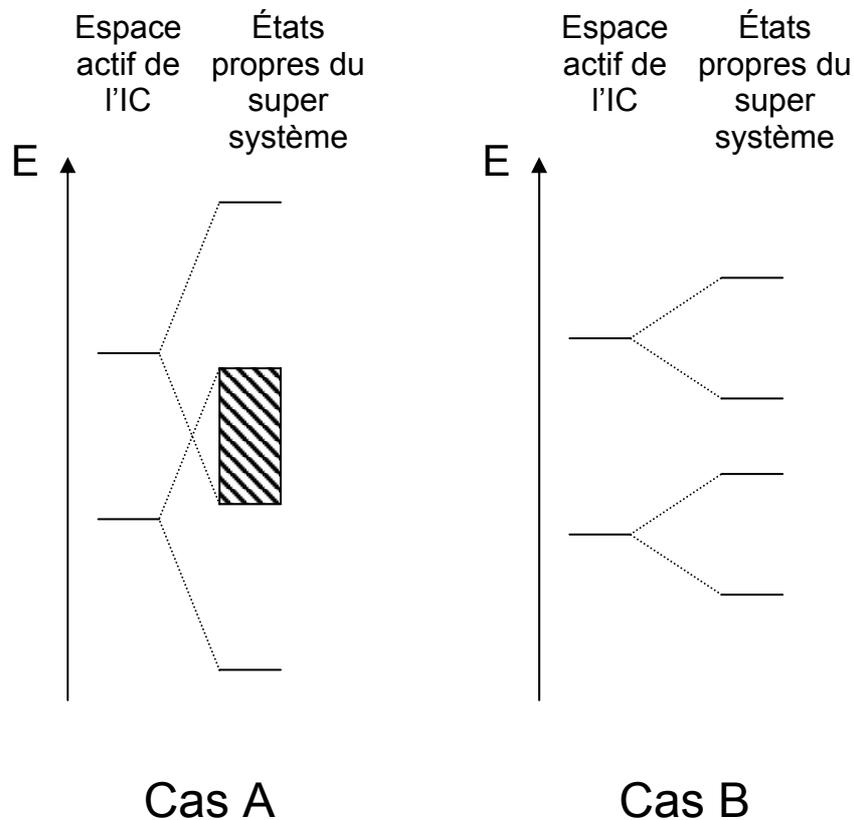


Figure 1.1

### 3. Méthodes à Hamiltonien Effectif

L'interaction entre états identiques de l'espace actif induit une bande d'états propres du supersystème dont la largeur varie en fonction du nombre des états initiaux. Le raisonnement qui suit se place dans le cas simple où seul le couplage entre plus proches voisins n'est pas nul. Deux états identiques de l'espace actif donnent alors naissance à deux états propres du supersystème distants de deux fois la valeur du couplage, tandis qu'un nombre infini d'états identiques en interaction résulte en une bande d'états propres de largeur égale à quatre fois la valeur du couplage. S'il y a recouvrement entre bandes, comme dans le cas A, il conviendra de prendre en compte l'intégralité des états de l'espace actif intervenant dans chacune des bandes. Au contraire, si la séparation énergétique entre états de l'espace actif est suffisante pour que nul recouvrement entre bandes n'ait lieu (cas B), il suffit d'employer comme base de l'interaction de configuration les configurations à l'origine de la bande qui contient le ou les états propres du supersystème à l'étude.

La construction de l'hamiltonien effectif, définie par le niveau de calcul de ses éléments de matrice diagonaux et non diagonaux, est liée au problème à l'étude et sera détaillée, sur des domaines d'applications précis, en parties deux et trois de ce mémoire.



## 4. Etude de Surfaces de d'Energy Potentielle

### 4.1. Surfaces d'énergie potentielle

Une surface d'énergie potentielle représente la variation de l'énergie potentielle d'un système en fonction des degrés de liberté associés à ses constituants. C'est donc, dans la plupart des cas, une application  $f: \mathfrak{R}^N \rightarrow \mathfrak{R}$ ,  $N$  étant le nombre de degrés de liberté décrivant le système. Soit  $Q = \{q_i, i \in [1..N]\}$ ,  $Q \in \mathfrak{R}^N$ , la représentation vectorielle de ces degrés de liberté  $q_i$ . L'ensemble des  $\{Q, f(Q)\}$ ,  $Q \in \mathfrak{R}^N$ , est un sous-espace de  $\mathfrak{R}^{N+1}$  dont un exemple pour  $N=2$  est représenté en figure 1.2.

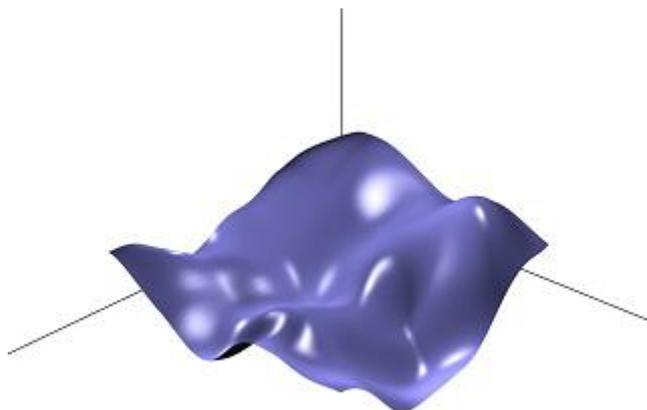


Fig. 1.2

Une surface d'énergie potentielle est caractérisée par un ensemble de points remarquables. Les minima et maxima ont en commun la propriété que  $\frac{\partial f}{\partial q_i} = 0, \forall i \in [1..N]$ . Ils diffèrent par le signe des valeurs propres de leur hessien  $H$  :

$$H = \left[ \frac{\partial^2 f}{\partial q_i \partial q_j} \right]$$

La matrice hessienne est réelle et symétrique. Sa diagonalisation fournit N valeurs propres  $\{\lambda_k, k \in [1..N]\}$ . Soit l'indice de Morse  $I_N$  défini par la relation :

$$I_N = \text{card}\{\lambda_k, \lambda_k \leq 0, k \in [1..N]\}$$

Le cas  $I_N=0$  correspond à un minimum. Pour  $I_N=N$ , c'est un maximum. Tout cas intermédiaire  $I_N=p$ ,  $p \in ]1..N[$  est dénommé point selle d'ordre p, c'est-à-dire un minimum selon N-p directions et un maximum dans les p restantes.

La présente étude, de par sa nature, s'attachera plus particulièrement à la recherche des minima et des points selle d'ordres peu élevés.

Les premiers définissent les structures stables du système chimique à l'étude ; leur nombre est généralement supérieur à un même dans le cas d'un petit nombre de degrés de liberté. Il est alors pratique de distinguer le minimum global, associé à l'énergie la plus faible, des minima locaux restants. La détermination du minimum global d'une surface d'énergie potentielle nécessite la connaissance de tous les minima de cette surface, ce qui est irréalisable en pratique puisque les degrés de liberté évoluent dans un espace qui, même borné, est dense dans  $\mathfrak{R}$ .

Les seconds revêtent également un intérêt aux yeux du chimiste, puisque, séparant deux minima sur la surface, ils fournissent la barrière énergétique à franchir pour qu'une interconversion entre les deux structures soit possible. La comparaison de ces barrières et de l'énergie apportée par le milieu permet d'expliquer simplement un grand nombre de phénomènes chimiques (chiralité ou racémisation, réactivité, formes mésomères et tautomères...).

## 4.2. Degrés de liberté d'un ensemble de corps rigides

Les diverses moutures de potentiels modèle employées au long de cette étude ont en commun la particularité de s'appliquer à des molécules rigides. Le nombre de degrés de liberté d'une assemblée de M telles entités indéformables est alors de  $6(M-1)$ , contre  $3N-6$  si ce même système est décrit par ses N atomes constitutifs.

Chaque molécule possède son propre repère dit *body fixed*, dont l'origine se situe au centre de masse et dont les axes sont positionnés par rapport aux axes du repère du laboratoire (*space fixed*) par trois angles dits angles d'Euler. Elle est donc décrite par trois translations et trois rotations par rapport au repère *space fixed*, soit six coordonnées. La notation des angles d'Euler adoptée ici est conforme à la *y-convention*<sup>147</sup> (Figure 1.3).

#### 4. Etude de Surfaces d'Énergie Potentielle

L'emploi d'angles d'Euler, s'il est répandu et intuitif, se heurte à deux difficultés. La première est la pléthore de conventions existantes. La seconde provient d'une singularité connue sous le nom de *gimbal lock* : lorsque les axes  $z$  et  $z'$  sont confondus,  $\theta$  est nul et  $\psi$  indéfini. L'utilisation des quaternions pour décrire les rotations (voir paragraphe 5.3 ci-dessous) permet de remédier à ce problème. Dans le cas où l'utilisation d'angles d'Euler est indispensable, la redéfinition des repères *body fixed* peut également intervenir de manière anecdotique.

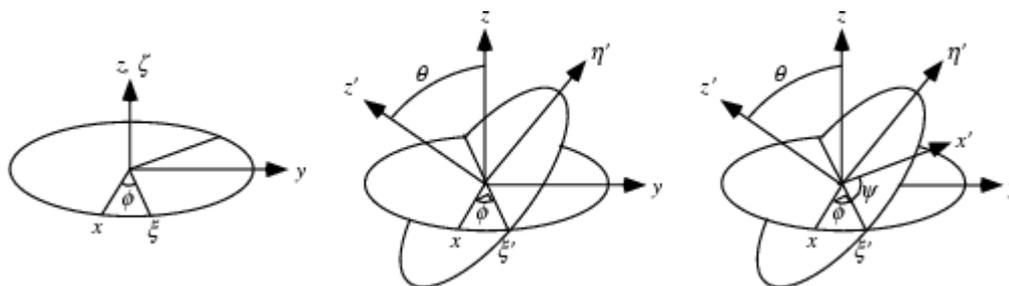


Fig. 1.3.

#### 4.3. Classification des méthodes d'optimisation

La recherche des minima d'une surface d'énergie potentielle est une représentation d'une plus vaste classe de problèmes dits problèmes d'optimisation. Transversale à toute discipline scientifique ou technique, cette classe a pour objet la détermination d'un jeu optimal de paramètres minimisant une fonction, dite fonction de coût ou de pénalité, et se subdivise en sept catégories dont aucune n'est mutuellement exclusive. Cette hiérarchie est présentée sur la figure 1.3 :

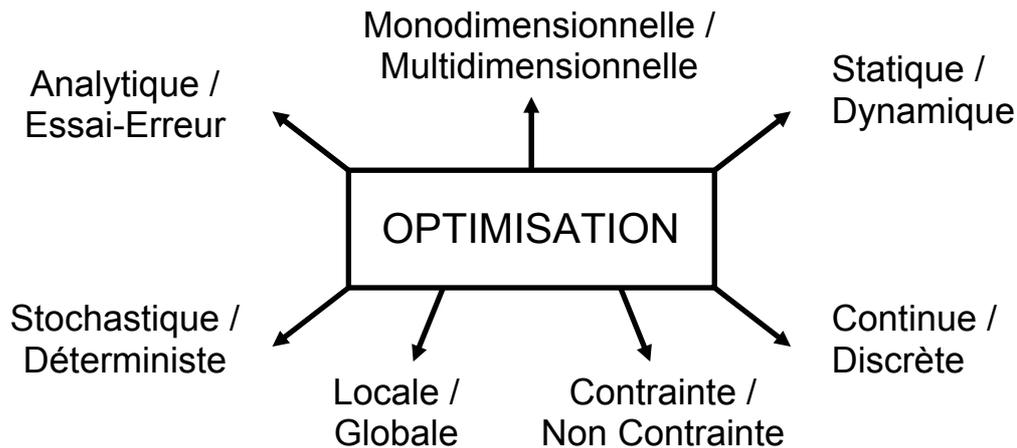


Fig 1.3.

- L'optimisation est dite analytique lorsque la fonction coût peut s'écrire analytiquement en fonction des paramètres à optimiser. L'exploitation de cette relation entre paramètres et coût permet d'orienter les recherches (pour une fonction de classe  $C^1$  sur son domaine de définition, les propriétés du gradient de la fonction par rapport à ses paramètres pourront par exemple être exploités). Lorsque aucune idée préconçue de l'action des paramètres sur le coût n'est connue (la fonction est une 'boîte noire'), l'optimisation procède au contraire par essai-erreur.
- Selon le nombre de paramètres mis en jeu, l'optimisation sera mono- ou multidimensionnelle. La sensibilité des différentes méthodes d'optimisation à l'augmentation de dimensionnalité est très variable, de sorte que la dimension du problème est souvent un facteur primordial de choix.
- Un problème statique ne dépend pas du temps. Dans une optimisation dynamique, au contraire, la fonction coût et les paramètres à optimiser dépendent du temps. Un problème dynamique est donc bien plus complexe à résoudre qu'un problème statique, ne serait-ce qu'à cause de l'augmentation de la dimensionnalité.
- Un problème d'optimisation peut également être classifié par la nature des paramètres qui y entrent en jeu. Un ensemble de paramètres discrets appelle une optimisation discrète, aussi nommée combinatoire puisque la solution du problème est une combinaison des valeurs possibles des paramètres et appartient à un ensemble dénombrable. Des paramètres continus (formant par exemple un ensemble qui est dense dans un corps) devront au contraire être traités à l'aide de méthodes continues.

#### 4. Etude de Surfaces d'Énergie Potentielle

- Une optimisation est dite contrainte lorsque des égalités ou inégalités entre paramètres sont incluses dans la fonction de coût. La contrainte peut être une fonction linéaire ou non des paramètres. Il est parfois possible de convertir un problème contraint en un problème non contraint, par exemple à l'aide d'un changement de variable ; les méthodes d'optimisation s'accommodent mieux d'un problème non contraint que d'un problème contraint.
- Les méthodes locales minimisent systématiquement la fonction coût au cours d'une série d'étapes successives. Partant d'un point  $P_k$  de l'espace des configurations,  $P_k \equiv (Q_k = \{q_i, i \in [1..N]\}, f(Q_k))$ , une étape élémentaire fournit le point  $P_{k+1} \equiv (Q_{k+1} = \{q'_i, i \in [1..N]\}, f(\{Q_{k+1}\}))$  tel que  $f(Q_{k+1}) \leq f(Q_k)$ . Il peut par exemple s'agir d'une minimisation unidimensionnelle le long de la coordonnée  $j$ , soit  $P_{k+1} \equiv (Q_{k+1} = \{q_1, q_2, \dots, q_j + \lambda_j, \dots, q_N\}, f(\{Q_{k+1}\}))$ . Chaque étape rapproche le système du minimum local le plus proche ; rien n'indique cependant que ce minimum soit le minimum global. Les méthodes globales, au contraire, prétendent échantillonner de manière aussi exhaustive que possible la surface d'énergie potentielle. A cet effet, elles ne cherchent pas systématiquement à diminuer la valeur de la fonction coût à chaque pas, mais lui laissent la liberté d'augmenter de sorte que le système puisse évoluer vers un bassin d'attraction distinct, éventuellement plus profond. Elles offrent donc une plus grande opportunité de converger vers le minimum global, même si rien ne peut prouver que celui-ci soit jamais atteint.
- Les méthodes stochastiques, dites aussi statistiques ou probabilistes, échantillonnent la surface de façon pseudo-aléatoire le long d'une chaîne de Markov<sup>148-150</sup>, c'est-à-dire une suite de tirs aléatoires où la probabilité de réussite d'un tir donné ne dépend pas du résultat du tir précédent. Généralement globales, ces méthodes autorisent l'énergie du système à augmenter en vertu d'une statistique d'équilibre boltzmanienne. Une étape donnée d'un processus déterministe ne dépend au contraire que de l'étape qui la précède, donc, par récurrence, d'un jeu de paramètres initial. De telles méthodes sont souvent locales.

La classification abordée ci-dessus justifie en grande partie le choix des méthodes retenues dans le cadre de la présente étude, qui sont présentées plus en détail dans les paragraphes suivants.

## 4.4. Méthodes globales

### *4.4.1. Recuit Simulé*

La méthode de recuit simulé<sup>151</sup> est une méthode globale stochastique à paramètres continus ou discrets qui fait décrire à un point initial P sur la surface d'énergie potentielle une trajectoire aléatoire gouvernée par une loi de distribution de Boltzmann régie par une température fictive T. Un ensemble de configurations est sélectionné, la probabilité de sélection étant fournie par la loi de distribution. Chaque configuration évolue selon une chaîne de Markov générée au moyen d'un critère de Metropolis<sup>152</sup> : la probabilité associée au passage d'une configuration d'énergie  $E_i$  à une configuration d'énergie  $E_{i+1}$  s'écrit :

$$p_{i+1} = e^{-\frac{E_{i+1}-E_i}{kT}}$$

Toute nouvelle configuration telle que  $E_{i+1} \leq E_i$  est donc automatiquement acceptée ; une nouvelle configuration d'énergie supérieure à la précédente ne l'est qu'en cas de tirage favorable. La température fictive est diminuée après un nombre donné d'évolutions, selon un programme de recuit fixé à l'avance. La diminution de la température s'accompagne d'une chute vers les zones de basse énergie de la surface d'énergie potentielle et restreint les probabilités d'échange entre bassins d'attraction. Lorsque la proportion de refus avoisine 100%, la convergence est atteinte. Une diminution progressive de la température s'accompagne d'une meilleure exploration de la surface de potentiel et augmente la probabilité de converger vers un minimum de basse énergie qui peut être le minimum global. Au contraire, une chute rapide de la température (une 'trempe') a de grandes chances d'emprisonner le système dans le bassin d'attraction initial.

La méthode de recuit simulé a connu un succès certain dans un nombre important de domaines. Ses principaux avantages sont d'être peu avide (c'est-à-dire de ne pas favoriser abusivement une décroissance rapide de l'énergie qui confinerait le système à un minimum local) et d'offrir une logique dans la prise de décision (les décisions les plus drastiques sont prises pendant la première phase, lorsque la température fictive est élevée ; leur irrévocabilité s'accroît au fur et à mesure que la température diminue mais peut être raffinée par les décisions de moindre ampleur qui sont prises ultérieurement).

Dans cette étude, la méthode de recuit simulé a été appliquée à l'exploration globale des surfaces d'énergie potentielle d'agrégats aromatiques, neutres et chargés.

### 4.4.2. Algorithme génétique

L'algorithme génétique est une méthode d'optimisation globale, conçue à l'origine par John Holland<sup>153,154</sup>, et dont la popularité ne cesse de s'accroître. Elle constitue la base des algorithmes dits évolutionnaires, qui prétendent émuler les principes de l'évolution biologique. L'algorithme supervise l'évolution d'une population composée d'un grand nombre d'individus au cours de générations successives, sous l'action de règles de sélection définissant le niveau d'adéquation de chacun.

La population ou *pool*, dont la taille  $N_{\text{pool}}$  est choisie par l'utilisateur, est constituée d'individus qui représentent autant de points sur la surface d'énergie potentielle. Chaque individu dispose d'un patrimoine génétique constitué de l'ensemble des paramètres le caractérisant, encapsulés dans un chromosome. Son adéquation est mesurée par la valeur de la fonction coût au point correspondant : l'individu est d'autant plus adéquat que cette valeur est basse. L'encodage des paramètres dans le chromosome dépend de leur nature (discrète ou continue). L'algorithme génétique binaire s'applique à un ensemble de paramètres discrets qu'il convertit en entiers binaires de longueur *ad hoc*. L'algorithme génétique continu manipule quant à lui des chromosomes de paramètres continus. Une fois définies les règles d'encodage, la première génération est constituée aléatoirement sur le domaine de définition des paramètres. Les individus sont alors classés par adéquation (c'est-à-dire par ordre croissant de la fonction coût).

La seconde étape consiste à choisir les individus qui constitueront les géniteurs de la génération suivante ; il s'agit en général des éléments les plus adéquats. Un groupe de  $N_{\text{gen}}$  géniteurs est isolé.  $\frac{1}{2}(N_{\text{pool}} - N_{\text{gen}})$  accouplements entre deux éléments de ce groupe sont menés, chaque union donnant naissance à deux enfants qui, avec leurs parents, constituent la nouvelle génération (dont la population est dans ce cas égale à celle de sa devancière). Dans l'algorithme binaire, le patrimoine génétique des enfants est généré à partir de celui des deux parents en choisissant un point de coupure sur le chromosome de ceux-ci et en panachant les deux morceaux, comme l'indique la figure 1.4. Il est à noter que la coupure peut avoir lieu à l'intérieur de la zone du chromosome correspondant à un paramètre. Dans le patrimoine génétique des enfants, celui-ci diffèrera de celui du père ou de la mère, puisqu'il est constitué des *bits* de poids fort de l'un et des *bits* de poids faible de l'autre ; ceci a pour effet d'accroître la diversité de la population. Dans l'algorithme continu, cet effet est émulé à l'aide d'une

combinaison linéaire à coefficients aléatoires des paramètres des deux parents. La diversité est encore accrue par un phénomène de mutation, dont l'occurrence est fixée par l'utilisateur et qui transforme aléatoirement une partie du patrimoine génétique de la nouvelle génération (par la négation d'un ou plusieurs *bits* par exemple). L'introduction de nouvelles valeurs de paramètres augmente la proportion explorée de la surface d'énergie potentielle, donc les chances de converger vers le minimum global.

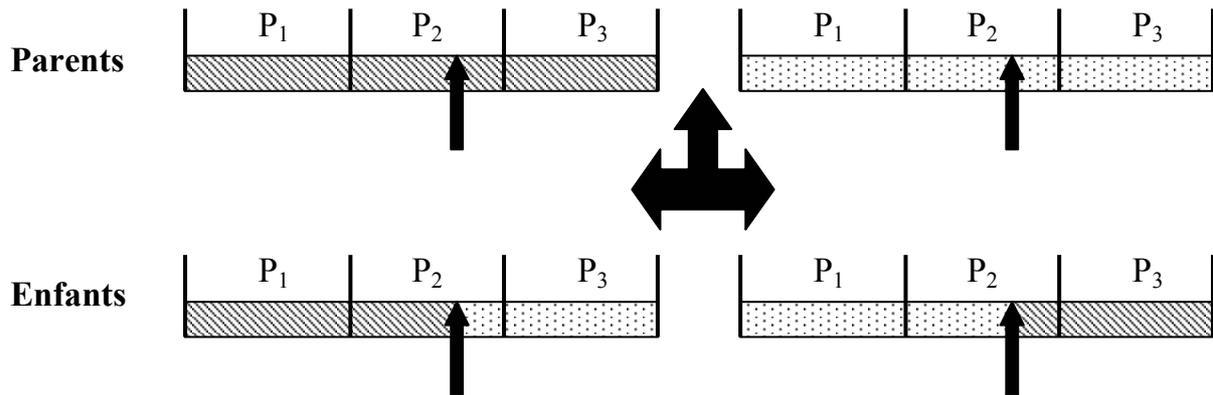


Fig 1.4.

La convergence est atteinte lorsque la population est constituée de  $N_{pool}$  représentations d'un même individu, et que toute mutation induit des résultats défavorables. La moyenne et l'écart-type des valeurs de la fonction coût sur l'intégralité de la population fournissent des renseignements à cet égard.

L'algorithme génétique est particulièrement adapté à l'étude de fonctions coûteuses en temps de calcul et à paramètres multiples, qui sont traitées comme des 'boîtes noires'. Tout comme le recuit simulé, cette méthode est peu avide et fournit un bon échantillonnage de surfaces complexes. Elle est en outre excessivement flexible, chacune des étapes décrites ci-dessus présentant un grand nombre de variantes ou de raffinements (sénescence de la population, appariement par tournois, etc...). Elle a trouvé naturellement sa place, dans le cadre de cette étude, pour la détermination de paramètres moyens du potentiel modèle reproduisant au mieux les énergies *ab initio* d'un nombre important d'agrégats distincts. Sur un tel cas d'optimisation d'une fonction coût complexe et sans expression analytique, aucune autre méthode ne peut rivaliser avec l'algorithme génétique.

#### 4.4.3. Croissance Monte Carlo

#### 4. Etude de Surfaces d'Énergie Potentielle

La méthode de croissance Monte Carlo ne constitue pas à proprement parler un procédé d'optimisation, mais a pour ambition de fournir un ensemble de configurations représentatif d'une surface d'énergie potentielle. Cet ensemble pourra faire l'objet, lors d'une seconde étape, d'une série d'optimisations locales fournissant les divers minima de la surface.

Il s'agit d'une méthode globale stochastique à paramètres continus, développée à l'origine par Garel<sup>155,156</sup> pour l'étude conformationnelle de macromolécules. Elle permet la construction d'un système par adjonction progressive de ses éléments constitutifs à la façon d'un arbre. Au critère énergétique des méthodes statiques standard – en l'occurrence une statistique boltzmanienne –, elle associe un critère topologique de construction qui assure une exploration raisonnée de la surface d'énergie potentielle. Du bon choix de ce critère dépend l'efficacité de l'exploration, donc l'intérêt de la méthode. La méthodologie retenue pour l'étude des agrégats moléculaires qui font l'objet de ce travail est celle formulée par Bertolus<sup>114</sup>. Une des molécules de l'agrégat initial est aléatoirement choisie et joue le rôle d'origine lors du placement de la nouvelle entité. Les coordonnées de translation de celle-ci par rapport à l'origine sont sélectionnées aléatoirement à l'intérieur d'un domaine fixé par l'utilisateur. Les trois degrés de liberté angulaires restants sont fixés par l'intermédiaire d'un algorithme Monte Carlo, de façon à ce que la molécule nouvellement ajoutée interagisse de façon attractive avec au moins l'une des autres entités de l'agrégat.

Les configurations retenues pour un agrégat de taille  $n$  forment un ensemble de cardinal  $M_n$ , constitué de  $g_n$  géométries distinctes d'énergie  $E_n^i, i \in [1..g_n]$ , et dont l'occurrence  $o_n^i, i \in [1..g_n]$  est donnée par :

$$o_n^i = A e^{-\frac{E_n^i}{kT}}$$

La nouvelle entité est rajoutée à chacune de ces géométries, un nombre de fois égal à son occurrence dans l'agrégat de taille  $n$ . Le poids statistique correspondant au passage de la  $i^{\text{ième}}$  configuration de l'agrégat de taille  $n$  à la  $j^{\text{ième}}$  configuration de l'agrégat de taille  $n+1$  s'écrit:

$$\omega_{n+1}^{i,j} = e^{-\frac{E_{n+1}^j - E_n^i}{kT}}$$

L'occurrence de la géométrie  $j$  est alors  $o_{n+1}^j = \omega_{n+1}^{i,j} o_n^i$ , arrondie à l'entier le plus proche ; un ensemble boltzmannien de configurations est ainsi automatiquement généré pour l'agrégat de taille  $n+1$ .

L'attrait principal de la méthode de croissance Monte Carlo est de se restreindre aux régions de la surface d'énergie potentielle qui présentent réellement un intérêt au sens chimique du terme. L'exploration est orientée par le biais du critère topologique : le choix de l'entité origine influe par exemple sur la forme de l'agrégat (sphérique ou filamentaire). En revanche, le choix d'une distance à l'origine comprise dans un intervalle donné présuppose une symétrie sphérique qui peut ne pas être appropriée si l'extension spatiale de la molécule origine est fortement dépendante de la direction (molécules planes par exemple). Si, dans le cadre de ce travail, la méthode a pu être raisonnablement appliquée à l'exploration des surfaces d'énergie potentielle d'agrégats neutres et chargés de benzène, elle a pour cette raison été remplacée par le recuit simulé dans le cas du naphthalène et de l'anthracène.

#### 4.5. Méthodes locales

Les méthodes d'optimisation locales présentées ci-dessous requièrent l'évaluation du gradient de la fonction  $f$  à optimiser (supposée  $C_1$  sur son domaine de définition), c'est-à-dire du vecteur  $G$  défini par :

$$G = \left( \frac{\partial f}{\partial q_i}, i \in [1..N] \right)$$

où les  $q_i$  représentent les  $N$  degrés de liberté du système. Lorsqu'une expression analytique de la fonction  $f$  est disponible,  $G$  est défini par la valeur au point considéré de chacune des fonctions  $\partial f / \partial q_i$  dont les expressions analytiques sont connues. Il est également possible (voire nécessaire si une expression analytique des dérivées de  $f$  ne peut être employée) de calculer numériquement la valeur du gradient en un point donné, au moyen d'approximations telles que des différences finies ; cela revient à remplacer une évaluation de la fonction  $\partial f / \partial q_i$  par au moins deux évaluations de la fonction  $f$ . Cette approche, moins précise et plus coûteuse que la première, n'est usitée qu'à défaut d'expression analytique pour  $f$ ; le potentiel modèle employé dans cette étude n'y aura donc pas recours.

La recherche de minima s'effectue le long de directions privilégiées fournies par le vecteur  $v = (v_i, i \in [1..N])$ , défini par :

$$v = -DG$$

## 4. Etude de Surfaces d'Énergie Potentielle

La constitution de la matrice D différencie les méthodes présentées ci-dessous.

### 4.5.1. Méthode de la plus grande pente (steepest descent)

La matrice D pour cette méthode est l'identité de rang N. Le système évolue dans la direction inverse du gradient, c'est-à-dire le long de l'axe de plus grande pente. L'avantage d'une convergence initiale rapide est contrecarré par le désavantage d'une convergence finale poussive, puisque deux directions de recherche successives sont toujours mutuellement orthogonales.

### 4.5.2. Méthodes à gradient conjugué

Il s'agit en réalité d'une classe de méthodes qui ont en commun l'actualisation de la matrice D à chaque itération en fonction de la valeur de cette matrice à l'itération précédente. Au prix d'un espace de stockage supérieur à celui de la méthode de plus grande pente, il est possible de tenir compte des informations accumulées au cours des itérations précédentes. La méthode employée dans ce travail est celle due à Broyden, Fletcher, Goldfarb et Shanno<sup>157-159</sup> et fait l'approximation que la fonction à optimiser est localement quadratique. Il a été montré que le minimum d'une telle fonction est accessible en N étapes à partir de la connaissance de la matrice hessienne inverse, N étant le nombre de paramètres à optimiser. L'algorithme BFGS construit par itérations une série de matrices dont la limite est précisément la hessienne inverse, qu'il exploite pour constituer la matrice D. Le lecteur est renvoyé aux références<sup>160,161</sup> pour plus de détails sur le sujet. Grâce à cet artifice, la convergence finale est notablement accélérée (d'autant plus que l'approximation quadratique est raisonnable). Le choix de cette méthode pour la présente étude se trouve ainsi pleinement justifié.

## 4.6. Méthodes hybrides et extensions

Les méthodes d'optimisation constituent un vaste sujet, en perpétuelle évolution. À l'interface de plusieurs communautés (scientifiques et informatiques), ce domaine profite

pleinement de l'interdisciplinarité et de la créativité qui en résulte. C'est ainsi qu'un nombre croissant de méthodes dites hybrides, visant à combiner les avantages des diverses approches, ont vu le jour ces dernières années. Il est par exemple intéressant de combiner la convergence initiale rapide d'une méthode *steepest descent* avec la convergence finale d'une méthode à gradient conjugué prenant le relais lors de la deuxième phase de l'optimisation. De même, les avantages respectifs des méthodes de recuit simulé et d'algorithme génétique peuvent être combinées avec succès, offrant l'accès à une convergence accélérée et à une meilleure exploration des surfaces d'énergie potentielle.

D'autres méthodes plus sophistiquées, issues de la recherche en intelligence artificielle, commencent à trouver leur place dans la communauté scientifique et peuvent être panachées avec les méthodes présentées ci-dessus<sup>162</sup>. Il en est ainsi des réseaux neuronaux et de la logique floue, qui permettent une approche heuristique du problème. Ces outils puissants donnent à la machine une capacité d'analyse d'un ensemble de données, ainsi que la faculté d'en extraire des tendances, et ce en toute autonomie. Il est également possible d'apprendre à la machine un ensemble de règles floues lui permettant de prendre des décisions raisonnées. Ceci facilite grandement l'orientation de la recherche de points remarquables sur une surface d'énergie potentielle et la compréhension de sa topologie par l'esprit humain, dont la faculté d'analyse de grands ensembles est limitée. Il est clair que des méthodes de ce type vont jeter un jour nouveau sur les problèmes d'optimisation dans les années à venir.

### 5. Dynamique Moléculaire Classique

Une simulation de dynamique moléculaire classique menée sur un système de  $N$  particules consiste à explorer l'espace des phases à  $6N$  dimensions correspondant, par intégration des équations différentielles couplées qui régissent le mouvement des particules. Les fondements théoriques présentées dans ces lignes constituent un bref survol de cette technique de simulation, que le lecteur pourra approfondir grâce à des ouvrages tels que <sup>163</sup> ou <sup>115</sup>.

#### 5.1 Généralités

Les simulations présentées ici ont été menées dans l'ensemble microcanonique dit aussi  $(N,V,E)$  : le nombre de particules  $N$ , le volume  $V^\dagger$  et l'énergie  $E$  du système sont des constantes du mouvement. L'énergie potentielle du système est calculée au moyen du potentiel modèle intermoléculaire décrit ci-dessus ; il en résulte que les particules susdites sont représentées par les molécules rigides. Les modes de vibration intramoléculaires (souvent des modes 'durs') sont donc négligés *de facto*. Il en va de même pour les effets quantiques, ce qui est d'autant plus raisonnable que l'agrégat est gros, le nombre de degrés de liberté important, la température élevée et les modes de vibration 'mous'. Les modes de vibration intermoléculaires vérifiant en général fort bien cette condition, l'emploi de la dynamique moléculaire classique est justifié même dans le cas des agrégats de tailles réduites envisagés dans cette étude. Il convient néanmoins de garder à l'esprit que cette approximation gagne en légitimité lorsque la taille des systèmes augmente.

#### 5.2. Mouvement de translation des molécules

Le mouvement de translation du centre de masse de chaque molécule rigide est régi par les équations de Newton :

$$m \frac{d^2 \vec{r}}{dt^2} = \vec{F}$$

---

<sup>†</sup> Cette nomenclature est classique mais mieux adaptée aux systèmes périodiques, le volume d'un agrégat étant une notion délicate à définir...

où  $\vec{F}$  est la résultante des forces s'appliquant au centre de masse,  $m$  la masse de la molécule, et  $\vec{r}$  le rayon-vecteur du centre de masse dans le repère du laboratoire. L'intégration numérique de cette équation au cours du temps est réalisée au moyen de l'algorithme de Verlet<sup>164</sup>, dont les étapes sont les suivantes :

$$\begin{cases} \vec{v}(t + \delta t/2) = \vec{v}(t - \delta t/2) + \delta t \frac{\vec{F}(t)}{m} \\ \vec{r}(t + \delta t) = \vec{r}(t) + \delta t \cdot \vec{v}(t + \delta t/2) \\ \vec{v}(t) = \frac{1}{2}(\vec{v}(t + \delta t/2) + \vec{v}(t - \delta t/2)) \end{cases}$$

La position de la molécule à chaque pas est déterminée à l'aide de sa vitesse aux demi-pas voisins. La précision en  $\delta t^3$  de cet algorithme lui confère de bonnes propriétés de conservation de l'énergie pour des pas de temps relativement longs.

### 5.3. Mouvement de rotation des molécules

Le mouvement de rotation de chaque molécule autour de son centre de masse est régi par les équations d'Euler :

$$\begin{cases} I_1 \frac{d\omega_1}{dt} - \omega_2\omega_3(I_2 - I_3) = N_1 \\ I_2 \frac{d\omega_2}{dt} - \omega_3\omega_1(I_3 - I_1) = N_2 \\ I_3 \frac{d\omega_3}{dt} - \omega_1\omega_2(I_1 - I_2) = N_3 \end{cases}$$

où  $(I_1, I_2, I_3)$  sont les trois moments d'inertie principaux de la molécule,  $(\omega_1, \omega_2, \omega_3)$  les trois composantes de sa vitesse angulaire de rotation dans le repère principal d'inertie, et  $(N_1, N_2, N_3)$  les trois composantes, dans ce même repère, du moment de la résultante des forces appliquées au centre de masse. Bien qu'il soit possible d'exprimer le vecteur vitesse angulaire  $\vec{\omega}$  au moyen des trois angles d'Euler employés dans les méthodes d'optimisation et d'exploration de surfaces décrites ci-dessus, l'utilisateur se heurte au problème du *gimbal lock*. Il est possible de le contourner en utilisant des quaternions comme paramètres de rotation. Un quaternion est un vecteur  $q = (q_0, q_1, q_2, q_3) \in \mathfrak{R}^4$ <sup>165-167</sup>. Un quaternion normé, c'est-à-dire tel que  $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$ , représente une rotation autour d'un axe, et peut être aisément converti en angles d'Euler (et *vice versa*). Il devient alors possible de relier les

## 5. Dynamique Moléculaire Classique

quaternions, leur dérivées, et le vecteur  $\vec{\omega}$  à l'aide de la relation suivante, exempte de singularité :

$$\begin{pmatrix} \frac{\partial q_0}{\partial t} \\ \frac{\partial q_1}{\partial t} \\ \frac{\partial q_2}{\partial t} \\ \frac{\partial q_3}{\partial t} \end{pmatrix} = \frac{1}{2} \bar{Q} \begin{pmatrix} 0 \\ \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix}$$

où  $\bar{Q}$  est une matrice 4x4 dépendant uniquement des quatre composantes de  $q$ . De même, le moment de la résultante des forces s'écrit en fonction des quaternions et des forces :

$$\begin{pmatrix} 0 \\ N_1 \\ N_2 \\ N_3 \end{pmatrix} = \frac{1}{2} {}^t\bar{Q} \begin{pmatrix} -\frac{\partial U}{\partial q_0} - \alpha q_0 \\ -\frac{\partial U}{\partial q_1} - \alpha q_1 \\ -\frac{\partial U}{\partial q_2} - \alpha q_2 \\ -\frac{\partial U}{\partial q_3} - \alpha q_3 \end{pmatrix}$$

où  ${}^t\bar{Q}$  est la transposée de la matrice  $\bar{Q}$  ci-dessus,  $U$  est le potentiel intermoléculaire,  $(-\partial U/\partial q_i), i \in [0..3]$  les forces dérivant du potentiel intermoléculaire, et  $(\alpha q_i), i \in [0..3]$  une force corrective, sans sens physique, visant à assurer la nullité de la composante  $N_0$ .

L'intégration numérique des deux équations précédentes s'effectue à l'aide de l'algorithme implicite de Svanberg<sup>168</sup> :

$$\begin{cases} q(t + \frac{\delta t}{2}) = q(t) + \frac{1}{2} \Xi(t + \frac{\delta t}{2}) \cdot \frac{\delta t}{2} & (1) \\ \Xi = \bar{Q} \cdot \omega^b & (2) \\ \omega^b = I^{-1} L^b & (3) \\ q(t + \frac{\delta t}{2}) = q(t) + \frac{1}{2} \Xi(t) \cdot \frac{\delta t}{2} & (4) \end{cases}$$

l'indice supérieur  $b$  désignant une quantité relative au repère moléculaire (*body fixed*).

L'équation (1) est résolue itérativement : partant de la valeur de  $q$  en  $t + \frac{\delta t}{2}$  donnée par (4), la valeur de  $\Xi$  en  $t + \frac{\delta t}{2}$  est déduite en utilisant (2) et (3) puis fournit une

nouvelle valeur pour  $q(t + \frac{\delta t}{2})$  via (1). Le processus est itéré jusqu'à convergence. Une fois la valeur de  $q$  en  $t + \frac{\delta t}{2}$  connue, sa valeur en  $t + \delta t$  est déterminée par:

$$q(t + \delta t) = q(t) + \frac{1}{2} \Xi(t + \frac{\delta t}{2}) \delta t$$

La supériorité de ce schéma d'intégration par rapport à l'algorithme de Fincham classique<sup>169</sup> a été démontrée<sup>168</sup>. Il autorise en pratique l'augmentation significative de la durée du pas d'intégration sans pour autant nuire à la conservation de l'énergie.

#### 5.4. Fonction d'autocorrélation des vitesses

L'obtention des fréquences de vibration est souvent une étape nécessaire lors de l'étude d'un système. L'analyse de la raideur des modes de vibration fournit des informations intéressantes quant aux propriétés et à la réactivité du système. Le calcul des fréquences constitue également une aide précieuse lors de l'interprétation de données expérimentales, issues de la spectroscopie infrarouge notamment. Expérimentalement, les différences d'énergie entre deux états électroniques sont déterminées entre leurs états vibrationnels de nombre quantique de vibration nul, tandis qu'au niveau théorique elles sont calculées de fond de puits à fond de puits. La connaissance des fréquences de vibration permet, par le calcul de l'énergie de point zéro, d'assurer la liaison avec les résultats expérimentaux. Enfin, au niveau théorique cette fois, la détermination des fréquences de vibration d'un système en un point remarquable de sa surface d'énergie potentielle permet de s'assurer de la nature de ce point (minimum ou point selle).

L'obtention des fréquences de vibration d'un système de grande taille est très coûteuse par les méthodes *ab initio*, puisqu'elle requiert l'évaluation de la matrice hessienne associée. Dès lors, il est bien plus avantageux d'employer à cet effet les informations de trajectoires retirées d'un calcul de dynamique moléculaire classique, en l'occurrence la transformée de Fourier de la fonction d'autocorrélation des vitesses des particules<sup>170</sup>.

Soient  $E_i = \{x_i(t), i \in [1..N]\}$ ,  $N$  ensembles statistiques pour une variable  $x(t)$ . La fonction d'autocorrélation de cette variable s'exprime comme une moyenne sur les  $E_i$  :

$$C(\tau) = \overline{x_i(t)x_i(t+\tau)} = \frac{1}{N} \sum_{i=1}^N x_i(t)x_i(t+\tau)$$

## 5. Dynamique Moléculaire Classique

L'emploi d'un temps de corrélation suffisant pour garantir une bonne exploration de l'espace des phases autorise l'application de l'hypothèse ergodique, qui stipule que la moyenne statistique sur l'ensemble des  $E_i$  est équivalent à la moyenne temporelle sur un seul ensemble. La fonction d'autocorrélation s'écrit alors :

$$C(\tau) = \overline{x_i(t)x_i(t+\tau)} = \langle x_i(t)x_i(t+\tau) \rangle = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T x(t)x(t+\tau)dt$$

La fonction d'autocorrélation représente, sur l'intervalle  $[0, T]$ , l'influence de la valeur de  $x(t)$  sur celle de  $x(t+\tau)$ <sup>171</sup>.

Pratiquement, la fonction d'autocorrélation des vitesses issue d'une dynamique de pas d'intégration  $\Delta t$  est une moyenne normalisée de corrélations à des instants différents :

$$\left\{ \begin{array}{l} C_k(j\Delta t) = \frac{1}{N} \left( \sum_{i=k}^{k+N} \vec{v}(i\Delta t) \cdot \vec{v}[(i+j)\Delta t] \right) \\ C(j\Delta t) = \frac{\sum_{k'} C_{k'}(j\Delta t)}{\sum_{k'} C_{k'}(0)} \end{array} \right.$$

Le spectre de vibration  $S(\omega)$  est obtenu par transformée de Fourier rapide<sup>172,173</sup> de cette fonction :

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

avec une résolution  $\Delta\omega = \frac{1}{2cN\Delta t}$  (en  $\text{cm}^{-1}$ ). Du fait qu'en pratique, la transformée porte sur un intervalle de temps fini, des fréquences parasites sont susceptibles de contaminer le spectre. Ce phénomène peut être limité par l'usage de filtres de convolution, en l'occurrence le filtre de Kaiser<sup>174,175</sup>.

### 5.5. Thermostats

L'ensemble microcanonique  $(N, V, E)$  considéré jusqu'à présent n'est qu'un ensemble de mécanique statistique parmi d'autres, tels  $(N, V, T)$  – canonique – et  $(N, P, T)$  – isotherme-isobare–. Dans la limite thermodynamique (taille infinie, densité  $N/V$  constante), les propriétés moyennes d'un système ne dépendent pas de l'ensemble dans lequel il est

décrit<sup>165</sup>. Cette limite n'étant en pratique jamais atteinte, le choix d'un ensemble thermodynamique n'est pas anodin.

L'ensemble microcanonique représente un système isolé entouré d'une paroi imperméable à la chaleur. L'ensemble canonique, au contraire, peut être perçu comme un sous-ensemble d'un système physique étendu (aussi dénommé bain thermique), avec lequel il peut échanger de l'énergie mais pas de particules. Cet ensemble affiche un certains nombres d'avantage par rapport à l'ensemble microcanonique, notamment son analogie avec l'expérience – la grande majorité des expériences conduites en laboratoire s'effectuent à pression et température constantes.

La température  $T$  d'un système à  $g$  degrés de liberté est reliée à la moyenne de son énergie cinétique par le biais du théorème d'équipartition :

$$T = \frac{2}{gk_B} \left\langle \sum_{i=1}^N m_i v_i^2 + \bar{\omega}_i \cdot I_i \cdot \bar{\omega}_i \right\rangle$$

chacune des molécules possédant une vitesse  $v_i$ , une masse  $m_i$ , une vitesse angulaire  $\bar{\omega}_i$  et un tenseur d'inertie  $I_i$ . Une alternative consiste, dans le cas de molécules rigides, à définir séparément des températures de translation et de rotation :

$$\begin{cases} T_{\text{trans}} = \frac{2}{g_{\text{trans}} k_B} \left\langle \sum_{i=1}^N m_i v_i^2 \right\rangle \\ T_{\text{rot}} = \frac{2}{g_{\text{rot}} k_B} \left\langle \sum_{i=1}^N \bar{\omega}_i \cdot I_i \cdot \bar{\omega}_i \right\rangle \end{cases}$$

Il est aisé de fixer l'énergie cinétique initiale d'un système à l'aide d'une distribution de Boltzmann à la température désirée. En revanche, pour peu que la configuration initiale soit éloignée de l'équilibre, elle possèdera une énergie potentielle excédentaire qui, convertie en énergie cinétique au cours d'une dynamique à énergie constante, produira une élévation indésirable de la température. Un mécanisme d'évacuation de cette énergie cinétique excédentaire, nommé thermostat, est donc indispensable.

Deux exemples de tels mécanismes, mis en application au cours de cette étude, sont présentés dans ces lignes. Le premier est simple et a pour ambition de fournir, à l'issue d'une phase d'équilibration préliminaire à la dynamique proprement dite, un agrégat à l'équilibre thermodynamique. Le second, plus complexe, permet une meilleure description de l'ensemble canonique et peut être employé pour réguler la température du système en cours de dynamique.

### 5.5.1. Thermostat de Berendsen

Le thermostat de Berendsen<sup>176</sup> est une représentation de la classe la plus simpliste de mécanismes de régulation, connue sous la dénomination de *velocity scaling*, et qui consiste à multiplier les vitesses des particules par un facteur correctif, à intervalles de temps réguliers au cours de la dynamique.

L'algorithme de Berendsen fait agir sur chaque molécule une force de friction (proportionnelle à la vitesse), dont le but est de ramener, à chaque instant, la température du système de sa valeur actuelle  $T$  vers la valeur de référence  $T_0$ . Les équations du mouvement s'écrivent alors :

$$\begin{cases} m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i + m_i \gamma \left( \frac{T_{\text{trans}}}{T_{\text{trans}}^0} - 1 \right) \frac{d\vec{r}_i}{dt} \\ I_i \frac{d\vec{\omega}_i}{dt} - \vec{\omega}_i \cdot I_i \cdot \vec{\omega}_i = \vec{N}_i - \gamma \left( \frac{T_{\text{rot}}}{T_{\text{rot}}^0} - 1 \right) I_i \cdot \vec{\omega}_i \end{cases}$$

La constante d'amortissement  $\gamma$  module l'intensité du couplage avec le bain thermique. Ceci revient à multiplier l'intégralité des vitesses (de translation comme de rotation), par un facteur d'échelle  $\lambda$  :

$$\lambda_{\text{trans}} = \sqrt{1 + \frac{\Delta t}{\tau} \left( \frac{T_{\text{trans}}^0}{T_{\text{trans}}} - 1 \right)} \quad \text{et} \quad \lambda_{\text{rot}} = \sqrt{1 + \frac{\Delta t}{\tau} \left( \frac{T_{\text{rot}}^0}{T_{\text{rot}}} - 1 \right)}$$

où  $\tau = (2\gamma)^{-1}$  est le temps caractéristique associé au couplage avec le bain, et  $\Delta t$  est l'intervalle de temps séparant deux ajustements des vitesses.

Cette technique de régulation, efficace mais simpliste, génère des trajectoires sans sens physique et ne peut donc être raisonnablement employée qu'au cours de la phase d'équilibration d'un système.

### 5.5.2. Thermostat de Nosé-Hoover

Le thermostat de Nosé-Hoover présente un couplage plus réaliste entre un système et un bain thermique. Il sera abordé ici dans le formalisme de Hoover<sup>165,177</sup>, reformulation des travaux initiaux de Nosé<sup>178,179</sup>. Comme précédemment, une force de friction assure l'apport

ou la dissipation de l'énergie cinétique nécessaire à la conservation de la température lors de la simulation :

$$\left\{ \begin{array}{l} m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i + m_i \zeta_{\text{trans}} \frac{d\vec{r}_i}{dt} \quad (1) \\ I_i \cdot \frac{d\vec{\omega}_i}{dt} - \vec{\omega}_i \cdot I_i \cdot \vec{\omega}_i = \vec{N}_i - \zeta_{\text{rot}} \cdot I_i \cdot \vec{\omega}_i \quad (2) \\ \frac{d\zeta_{\text{trans}}}{dt} = \frac{g_{\text{trans}}}{Q_{\text{trans}}} k_B (T_{\text{rot}} - T_{\text{rot}}^0) \quad \text{et} \quad \frac{d\zeta_{\text{rot}}}{dt} = \frac{g_{\text{rot}}}{Q_{\text{rot}}} k_B (T_{\text{trans}} - T_{\text{trans}}^0) \quad (3) \end{array} \right.$$

$\zeta_{\text{trans}}$  et  $\zeta_{\text{rot}}$  sont des variables de couplage avec le bain thermique, qui présentent ici la particularité d'être dynamiques puisque régies par les équations d'évolution (3). Celles-ci font intervenir les paramètres  $Q_{\text{trans}}$  et  $Q_{\text{rot}}$ , analogues à des 'masses' fictives. Le couplage avec le bain thermique induisant des oscillations sans sens physique, de période  $t_0 = 2\pi\sqrt{Q/2gk_B T}$ , le choix des paramètres  $Q$  est dicté par deux critères. En premier lieu, l'inclusion d'un nombre suffisant d'oscillations dans la simulation est indispensable pour garantir un échantillonnage correct de la portion de l'espace des phases qui est exploré lors d'une oscillation. De ce choix dépend la qualité des propriétés moyennes calculées. En second lieu,  $Q$  doit être tel que la période  $t_0$  soit grande par rapport à la durée de vie typique des fonctions de corrélation dynamiques. Ce critère, particulièrement important pour le calcul de propriétés dynamiques telles que les fonctions d'autocorrélation des vitesses, assure le découplage entre la dynamique fictive du bain thermique et la dynamique moléculaire à proprement parler.

## 5.6. Principe de Gauss

Il est parfois souhaitable, au cours d'une simulation de dynamique moléculaire, de limiter la portion de l'espace des phases explorée par le système à celle qui correspond le mieux à la nature et aux propriétés de ce système. Il peut par exemple se révéler intéressant de conserver une coordonnée généralisée (distance, angle ou angle dièdre entre atomes) au cours de la dynamique. Il peut également être souhaitable de conserver une quantité thermodynamique (température, enthalpie, énergie libre, etc...). Ces exemples constituent autant de contraintes qui, reliant entre elles les coordonnées généralisées et leurs dérivées

## 5. Dynamique Moléculaire Classique

temporelles, limitent le nombre de degrés de liberté effectifs du système à l'étude. De manière générale, une contrainte  $G$  s'écrit :

$$G\left(\vec{r}_i, \frac{\partial \vec{r}_i}{\partial t}, t\right) = 0$$

Une contrainte est dite holonome si elle est intégrable. Elle est dite non holonome si elle ne l'est pas, c'est-à-dire si la force de contrainte associée fournit un travail. Les contraintes non holonomes dépendent en général des dérivées temporelles des coordonnées généralisées. La conservation de grandeurs thermodynamiques implique généralement de telles contraintes : les thermostats présentés ci-dessus en sont de bons exemples. Les contraintes holonomes ne dépendent en général que des seules coordonnées généralisées et interviennent dans la conservation de critères topologiques (distances, angles et dièdres). Elles sont rhéonomes lorsque leur équation fait explicitement référence au temps, scléronomes dans le cas contraire.

Le principe de Gauss<sup>180</sup>, dit de moindre contrainte, est une façon élégante de définir une contrainte par son action sur le système. Il introduit la courbure  $C$ , fonction des accélérations des particules et définie par :

$$C = \frac{1}{2} \sum_{i=1}^N m_i \left( \frac{\partial^2 \vec{r}_i}{\partial t^2} - \frac{\vec{F}_i}{m_i} \right)^2$$

Le principe de Gauss stipule que les accélérations réellement observées minimisent  $C$ . Ainsi, le mouvement contraint de particules doit rester aussi proche que possible du mouvement non contraint, pour lequel  $C = 0$ .

Soit une contrainte générale  $G$ . L'équation de contrainte différentielle :

$$\vec{n}\left(\vec{r}_i, \frac{\partial \vec{r}_i}{\partial t}, t\right) \cdot \frac{\partial^2 \vec{r}_i}{\partial t^2} = s\left(\vec{r}_i, \frac{\partial \vec{r}_i}{\partial t}, t\right)$$

obtenue par différenciation de  $G\left(\vec{r}_i, \frac{\partial \vec{r}_i}{\partial t}, t\right) = 0$ , est l'équation du plan de contrainte, de normale  $\vec{n}$  et de distance à l'origine  $s$ . Le mouvement contraint doit demeurer sur le plan, de sorte que la contrainte a pour effet de projeter les accélérations non contraintes sur celui-ci. La force de contrainte est donc colinéaire à  $\vec{n}$  et l'équation du mouvement devient :

$$\sum_{i=1}^N m_i \frac{\partial^2 \vec{r}_i}{\partial t^2} = \sum_{i=1}^N \vec{F}_i - \lambda \vec{n}$$

où  $\lambda$  est un multiplicateur de Gauss donné par :

$$\lambda = \frac{\vec{n} \cdot \sum_{i=1}^N \vec{F}_i}{\vec{n} \cdot \vec{n}}$$

Le principe de Gauss a pour avantage de ne se référer explicitement qu'à la différentielle de la contrainte et non à la contrainte elle-même. Il s'est révélé fort utile, dans la présente étude, pour astreindre les centres de gravité respectifs des monomères d'un agrégat de naphthalène colonnaire à un axe, au cours d'une simulation de dynamique moléculaire.

### 6. Dynamique Moléculaire avec Transitions Électroniques

**B**ien que puissantes et fréquemment employées, les méthodes de dynamique moléculaire fondées sur l'intégration numérique des équations classiques du mouvement se heurtent à des limites conceptuelles : les comportements quantiques sont négligés et le mouvement des noyaux est gouverné par une unique surface d'énergie potentielle adiabatique, dans l'approximation de Born-Oppenheimer. Elles se révèlent donc inadéquates dans nombre de problèmes affichant un ou plusieurs degrés de liberté quantiques. S'il est possible de traiter quantiquement lesdits degrés de liberté tout en conservant l'approche classique pour les autres dans le cas où les variables quantiques demeurent dans un état adiabatique tout au long de la simulation, cette stratégie échoue lorsque des transitions non adiabatiques entrent en jeu. Les probabilités de transitions quantiques résultent des trajectoires classiques des particules, elles-même influencées par les populations relatives des différents états quantiques. L'incorporation correcte de ces effets est cruciale pour la description de phénomènes tels la photochimie, les transitions non radiatives, le transfert d'excitation ou de charge et la relaxation vibrationnelle.

#### 6.1. Méthode de *Trajectory Surface Hopping*

Les méthodes de dynamique moléculaire avec transitions électroniques visent à pallier les limitations de l'approche classique présentées ci-dessus. Celle présentée ici, connue sous le nom de *trajectory surface hopping* (TSH), est due aux travaux de Tully<sup>181-184</sup>. Particulièrement adaptée aux états électroniques issus de méthodes à hamiltonien effectif, elle présente la particularité de ne considérer qu'une partie des degrés de liberté quantiques d'un système (existence de plusieurs états électroniques couplés de manière non adiabatique), négligeant par exemple l'effet tunnel ou les mouvements vibrationnels de point zéro dont la prise en compte est coûteuse en temps de calcul. L'interconnexion entre les degrés de liberté classiques et quantiques est quant à elle retranscrite de manière autocohérente.

La méthode TSH apporte un degré de sophistication supplémentaire par rapport aux méthodes à champ moyen de type Ehrenfest<sup>185,186</sup>. Dérivées du traitement *time-dependent self-consistent field* (TDSCF)<sup>187,188</sup>, celles-ci décrivent le mouvement des particules classiques

dans le champ moyen crée par les particules quantiques, et vice versa ; il en résulte que les particules classiques évoluent sur une surface d'énergie potentielle correspondant à la moyenne des états quantiques du système. Les méthodes à champ moyen échouent donc dans les systèmes caractérisés par des couplages forts et une interdépendance marquée entre forces classiques et densités des états quantiques. La méthode TSH leur substitue une approche multiconfigurationnelle, où le système est décrit comme une superposition d'états Born-Oppenheimer ; chacun de ces états est défini par le produit d'une partie classique propre à l'état quantique correspondant, et d'une partie quantique dépendant paramétriquement des coordonnées classiques. Elle retranscrit donc la corrélation classique-quantique de manière moins globale.

Soit l'hamiltonien total  $\hat{H}$  du système :

$$\hat{H} = \hat{T}_R + \hat{H}_0(\vec{r}, \vec{R})$$

$\hat{T}_R$  est l'opérateur énergie cinétique des noyaux et  $\hat{H}_0$  l'hamiltonien électronique.

Soit à présent une base orthonormale quelconque  $\{\phi_i(\vec{r}, \vec{R}), i \in [1..n]\}$  de fonctions d'onde électroniques qui dépendent paramétriquement de la position des noyaux, et qui sera constituée dans cette étude par les fonctions propres adiabatiques du système à la géométrie actuelle du système. En supposant que les noyaux évoluent sur une trajectoire  $\vec{R} = \vec{R}(t)$  donnée qui dépend du temps, l'hamiltonien électronique  $\hat{H}_0$  en dépend également.

$\Psi(\vec{r}, \vec{R}, t)$ , solution de  $\hat{H}_0$ , est décomposée dans la base des  $\{\phi_i(\vec{r}, \vec{R})\}$  :

$$\Psi(\vec{r}, \vec{R}, t) = \sum_j c_j(t) \phi_j(\vec{r}, \vec{R})$$

les coefficients  $c_j$  étant à valeurs complexes. Cette fonction d'onde doit vérifier l'équation de Schrödinger dépendante du temps, ce qui donne :

$$i\hbar \dot{c}_k = \sum_j c_j (V_{kj} - i\hbar \dot{\vec{R}} \cdot \vec{d}_{kj}) \quad \forall k \in [1..n] \quad (1)$$

avec les notations :

$$\begin{cases} V_{ij}(\vec{R}) = \langle \phi_i(\vec{r}, \vec{R}) | H_0(\vec{r}, \vec{R}) | \phi_j(\vec{r}, \vec{R}) \rangle \\ \vec{d}_{ij}(\vec{R}) = \langle \phi_i(\vec{r}, \vec{R}) | \vec{\nabla}_R \phi_j(\vec{r}, \vec{R}) \rangle \end{cases}$$

L'état électronique réel du système à l'instant  $t$  est une combinaison linéaire des états propres de  $\hat{H}_0$  de coefficients d'expansion  $c_j$ , dont l'évolution temporelle est régie par le

## 6. Dynamique Moléculaire avec Transitions Électroniques

système d'équations différentielles couplées du premier ordre (1). En revanche, la dynamique classique des noyaux ne s'effectue que sur l'une des surfaces d'énergie potentielle  $\phi_i$ , la combinaison linéaire de surfaces n'ayant aucun sens. Le système sera donc considéré dans un état propre  $\phi_i$  de  $\hat{H}_0$ . Les équations (1), intégrées au cours du mouvement, fournissent les coefficients complexes  $c_j$  qui déterminent les probabilités de changement d'états (voir ci-dessous). La transition vers un état  $\phi_j, j \neq i$ , est donc promue par deux facteurs : les couplages électroniques  $V_{ij}$  et non adiabatiques  $\dot{\vec{R}} \cdot \vec{d}_{ij}$ .

Il est pratique de remplacer les  $c_j$  par la matrice densité associée, d'éléments :

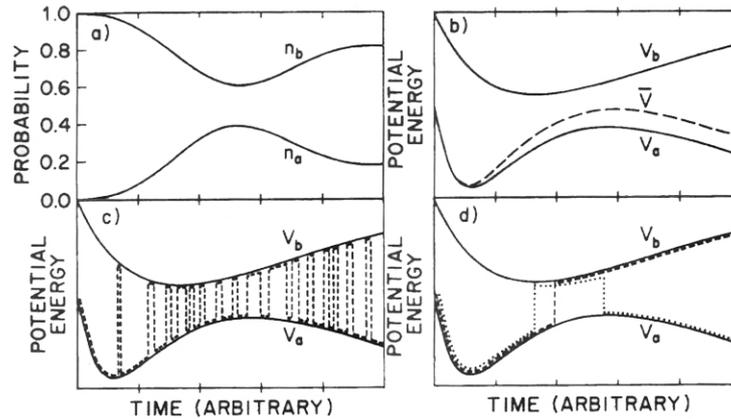
$$a_{kj} = c_k c_j^*$$

Les éléments diagonaux  $a_{kk}$ , réels, représentent la population du  $k^{\text{ième}}$  état électronique du système, et évoluent au cours du temps. Les éléments non diagonaux  $a_{kj}$  définissent quant à eux la cohérence entre états. La probabilité  $P_{ij}$  associée à la transition de l'état  $i$  vers l'état  $j$  peut s'exprimer en fonction des populations relatives des états concernés et de leurs dérivées temporelles ; le lecteur est renvoyé à la référence<sup>182</sup> pour l'expression analytique de ces probabilités. L'ensemble des probabilités de transition d'un niveau  $i$  donné vers tous les niveaux  $j$  subdivise l'intervalle  $[0,1]$ , de sorte que l'état électronique vers lequel le saut a lieu (le cas échéant) peut être déterminé par tirage d'un nombre aléatoire compris dans cet intervalle.

L'algorithme de saut de surface employé ici ajoute à ce formalisme une approximation supplémentaire. Soit par exemple un système à deux niveaux, de populations respectives  $a_{11}$  et  $a_{22}$ , comportant  $N$  trajectoires et dont l'évolution est telle que :

$$\begin{cases} a_{11}(t + \delta t) < a_{11}(t) \\ a_{22}(t + \delta t) > a_{22}(t) \end{cases}$$

Au cours du temps  $\delta t$ , davantage de trajectoires passent de l'état 1 vers l'état 2 que dans le sens inverse. La condition de *minimum switch* impose un nombre  $N(a_{22}(t + \delta t) - a_{22}(t))$  de trajectoires de 1 vers 2 et un nombre nul de trajectoires de 2 vers 1. L'intérêt de cette approche est illustrée à l'aide de la figure 1.5, présentée et expliquée ci-dessous :


 Fig. 1.5. (tirée de <sup>182</sup>)

Les populations normalisées des deux niveaux (dénommés ici  $a$  et  $b$ ) évoluent comme indiqué en (a). En (b) sont schématisées les surfaces d'énergie potentielle associées aux deux niveaux, ainsi que, en pointillés, le potentiel 'moyen' pondéré par la populations des états, qui n'a aucun sens physique dès que le système émerge de la zone de croisement évité. Sans condition de *minimum switch* (c), la probabilité de transition est grande à chaque pas d'intégration, de sorte qu'un grand nombre de sauts ont lieu au cours de la traversée de la zone de croisement évité. L'état électronique final obtenu en sortie de cette zone dépend du nombre de sauts ayant effectivement eu lieu, qui varie d'une trajectoire à l'autre. L'effet global est de faire évoluer le système sur la surface 'moyenne' déjà présentée en (b). La condition de *minimum switch* (d), au contraire, réduit le nombre de transitions effectives tout en conservant la distribution statistique correcte de population des états.

La transition d'un état électronique  $i$  vers un état  $j$  s'accompagne d'une variation de l'énergie potentielle qui doit être compensée, dans l'ensemble microcanonique, par un réajustement de l'énergie cinétique des particules, de façon à assurer la conservation de l'énergie totale. Les différents modes de vibration sont d'autant plus influencés par le réajustement qu'ils sont colinéaires au vecteur directeur du couplage non-adiabatique au point de saut  $\vec{d}_{ij}(\vec{R}) = \langle \phi_i(\vec{r}, \vec{R}) | \vec{\nabla}_R \phi_j(\vec{r}, \vec{R}) \rangle$ . Le passage d'un état excité à un autre d'énergie plus basse s'accompagne de la conversion d'énergie électronique en énergie cinétique qui est répartie dans les différents modes comme indiqué ci-dessus ; ce type de transition est toujours possible. En revanche, la transition vers un état excité d'énergie plus élevée doit donner lieu à la conversion d'énergie cinétique en énergie électronique. Si les modes concernés (ceux qui sont le plus colinéaires au vecteur  $\vec{d}_{ij}$ ) ne peuvent céder un énergie suffisante, le saut sera refusé.

## 6.2. Mise en œuvre

La méthode de trajectory surface hopping implémentée dans cette étude emploie comme fonctions d'onde  $\{\phi_i(\vec{r}, \vec{R})\}$  la base des états électroniques adiabatiques du système, c'est-à-dire les fonctions propres de  $\hat{H}_0$ , avec pour conséquence que  $V_{ij}(\vec{R}) = \langle \phi_i(\vec{r}, \vec{R}) | \hat{H}_0(\vec{r}, \vec{R}) | \phi_j(\vec{r}, \vec{R}) \rangle = \delta_{ij} \varepsilon_i$ , où  $\varepsilon_i$  désigne la  $i^{\text{ème}}$  valeur propre de  $\hat{H}_0$ . Ces états sont en effet les seuls à posséder un sens physique réel dans les zones de croisement évités où plusieurs surfaces de potentiels parviennent en vicinité, étant donnés les temps d'intégration caractéristiques employés.

- Un état adiabatique initial  $i$  est préparé ; la valeur initiale correspondante des coefficients  $c_j$  est  $(0_1, 0_2, \dots, 1_i, \dots, 0_n)$ . Les valeurs initiales des coordonnées  $q$  et des moments  $p$  des noyaux sont également supposées connues.
- Les équations classiques du mouvement sont intégrées de  $t$  à  $t+\Delta t$ , comme décrit au paragraphe 5, le long de la surface d'énergie potentielle de l'état électronique actuel  $i$ . Il en résulte une nouvelle configuration des noyaux  $\vec{R}(t+\Delta t)$ .
- L'hamiltonien  $\hat{H}_0$  est diagonalisé et fournit les valeurs et vecteurs propres  $\{\varepsilon_i(\vec{r}, \vec{R}(t+\Delta t))\}$  et  $\{\phi_i(\vec{r}, \vec{R}(t+\Delta t))\}$ ,  $i \in [1..n]$ . Un vecteur propre normé étant défini au signe près, il convient de s'assurer que la convention de signe est la même pour  $\phi_i(\vec{r}, \vec{R}(t))$  et  $\phi_i(\vec{r}, \vec{R}(t+\Delta t))$ , sous peine de changements brutaux dans les couplages non adiabatiques. Ceci est réalisé par projection du vecteur  $\phi_i(\vec{r}, \vec{R}(t+\Delta t))$  sur chacun des vecteurs  $\phi_i(\vec{r}, \vec{R}(t))$ ,  $i \in [1..n]$ . Soit  $m$  tel que  $|\phi_1(\vec{r}, \vec{R}(t+\Delta t)) \cdot \phi_m(\vec{r}, \vec{R}(t))| = \text{Max}(|\phi_1(\vec{r}, \vec{R}(t+\Delta t)) \cdot \phi_i(\vec{r}, \vec{R}(t))|, i \in [1..n])$ . Si  $\phi_1(\vec{r}, \vec{R}(t+\Delta t)) \cdot \phi_m(\vec{r}, \vec{R}(t)) < 0$ , le signe des composantes de  $\phi_1(\vec{r}, \vec{R}(t+\Delta t))$  est modifié. Cette approche permet de tenir compte d'un éventuel changement dans l'ordre des états entre  $t$  et  $t+\Delta t$  (si  $m \neq 1$ ). Elle ne permet pas, en revanche, de remédier au problème de définition des vecteurs propres associés à un espace propre de dimension supérieure à un. Aucun cas de dégénérescence de valeurs propres n'ayant été mis en évidence au cours de cette étude (de par l'emploi d'une représentation adiabatique), cette lacune ne prête pas à conséquence. Un traitement possible dans de tels cas est présenté dans<sup>189</sup>.

- Le couplage non adiabatique  $\dot{\vec{R}} \cdot \vec{d}_{ij}$  est évalué à l'instant  $t+\Delta t/2$  :

$$\left[ \dot{\vec{R}} \cdot \vec{d}_{ij} \right] \left( t + \frac{\Delta t}{2} \right) = \frac{1}{2\Delta t} \left( \langle \phi_i(t) | \phi_j(t + \Delta t) \rangle - \langle \phi_i(t + \Delta t) | \phi_j(t) \rangle \right) \text{ pour tout } j.$$

- Les équations (1) sont intégrées sur l'intervalle  $\Delta t$  en plusieurs étapes de durée  $\delta t$  telle que  $\delta t \ll \Delta t$ . En effet, les phénomènes quantiques se produisent sur un temps caractéristique plus court que leurs homologues classiques, ce qui justifie un temps d'intégration inférieur. Les valeurs de  $V_{ij}$  et  $\dot{\vec{R}} \cdot \vec{d}_{ij}$  aux étapes intermédiaires sont déterminées par inter- et extrapolation linéaire. L'intégration à proprement parler emploie un algorithme de Runge-Kutta d'ordre quatre<sup>160,161</sup>, réputé pour sa robustesse.
- Les probabilités de saut  $P_{ij}$  sont déterminées pour tout  $j$ , à partir des valeurs de  $c_j$  à l'instant  $t+\Delta t$ . Le tirage d'un nombre aléatoire détermine, le cas échéant, la surface  $k$  vers laquelle le saut a lieu.
- Le vecteur  $\vec{d}_{ij}$  est calculé à l'aide du théorème de Hellmann-Feynman non diagonal<sup>190</sup> :

$$\vec{d}_{ik} = \frac{\langle \phi_i | \vec{\nabla}_R \hat{H}_0 | \phi_k \rangle}{\varepsilon_k - \varepsilon_i}$$

- Les vitesses sont réajustées. Si la transition a lieu vers un état électronique d'énergie plus élevée que l'état d'origine, la présence d'énergie vibrationnelle suffisante est contrôlée, faute de quoi le saut échoue.
- Le processus est itéré jusqu'à la fin de la simulation.

Il convient de noter que le déroulement de la simulation dépend de la graine utilisée lors de l'initialisation du générateur de nombres pseudo-aléatoire. Une statistique doit donc en toute rigueur être menée sur les conditions d'initialisation du générateur, en sus de la statistique sur les positions et vitesses initiales des noyaux requise par la dynamique classique. Un échantillonnage suffisamment exhaustif de ces deux types de conditions initiales peut donc se révéler fort coûteux.

## 7. Développement et Programmation

L'utilisation de l'outil informatique, et plus particulièrement la conception et l'implémentation d'algorithmes et de logiciels, joue un rôle fondamental dans le domaine de la chimie computationnelle. Ce rôle est encore plus crucial dans le cadre de travaux méthodologiques qui requièrent la mise au point et l'implémentation de modèles non disponibles dans les logiciels commerciaux ou académiques classiques. Or, bien souvent, l'aspect technique du travail est passé sous silence au profit de ses implications scientifiques. Pourtant, les deux domaines sont généralement interdépendants voire indissociables : par exemple, le choix d'une méthode pour un problème donné relève en général du meilleur compromis entre sophistication et coût en temps de calcul, ce dernier aspect dépendant du logiciel utilisé, du compilateur qui l'a généré, de la conception du programme source et de la machine sur laquelle il est exécuté.

Une part importante de ce travail de thèse a consisté en programmation et développement d'outils implémentant les diverses méthodes présentées ci-dessus. Les principales avancées effectuées sont les suivantes :

- Ajout de l'anisotropie dans le modèle d'énergie de polarisation en potentiels modèle. L'énergie de polarisation et ses dérivées par rapport aux degrés de liberté sont calculées, de manière très générale, à l'aide de tenseurs de polarisabilité locaux dont la constitution est laissée au choix de l'utilisateur. Divers modèles pour ces tenseurs de site ont été mis au point en fonction des systèmes chimique étudiés.
- Ajout du formalisme de transfert de charge dans le potentiel modèle, impliquant le calcul d'intégrales de recouvrement entre orbitales moléculaires, et leurs dérivées, en temps réel et pour une géométrie quelconque de l'agrégat étudié.
- Conception d'un algorithme génétique à paramètres continus.
- Ajout de contraintes et de thermostats au programme de dynamique moléculaire classique.
- Développement d'un logiciel implémentant la méthode TSH, en transfert de charge ou d'excitation.

Compte tenu du travail de développement réalisé et de l'expérience acquise, il a semblé intéressant de conclure la présente partie par un court aperçu des possibilités des langages de programmation modernes (compilés ou interprétés) qui ont été mis en œuvre dans cette étude.

## 7.1. Langages compilés

Le langage FORTRAN 77 reste très employé par la communauté des chimistes computationnels, principalement en raison du nombre important de programmes et routines accumulés au cours des années. Il se montre en outre généralement plus rapide que d'autres langages, particulièrement pour la gestion des tableaux et des réels à virgule flottante longs, et est conceptuellement très simple. Il souffre néanmoins de graves manquements :

- L'absence de variables globales. Pour étendre la visibilité d'une variable au programme entier, il convient dès lors de la passer en argument lors d'appel de sous-programmes ou de l'inclure dans un bloc COMMON. La première solution rend le programme peu lisible par l'accumulation du nombre d'arguments d'appel des sous-programmes ; la seconde est peu élégante (il faut répéter la déclaration COMMON dans chaque unité) et dangereuse (aucun contrôle n'est réalisé sur la disposition correcte des données du bloc). L'instruction COMMON a d'ailleurs été déclarée périmée par la norme FORTRAN 90.
- L'absence de structures de données. La création de types complexes est pourtant pratique pour rendre compte d'une dépendance ou d'une hiérarchie entre données.
- Les tableaux statiques. Les tableaux doivent être dimensionnés une fois pour toutes dans le source. Lorsque leur dimension ne peut être connue à l'avance, il est donc d'usage de les surdimensionner, ce qui constitue un gaspillage de mémoire, voire une source potentielle de dysfonctionnements en cas de débordement.
- L'instruction GOTO. L'emploi de cette instruction est généralisé en FORTRAN 77. Bien qu'elle apparaisse souvent comme une solution de facilité, cette instruction brise la logique d'un algorithme et constitue une source infinie de bogues. Son emploi est souvent révélateur d'une

## 7. Développement et Programmation

mauvaise conception de l'algorithme. Comme COMMON, elle a été rendue obsolète par la norme FORTRAN 90.

- L'instruction IMPLICIT. Une autre solution de facilité, rendue très dangereuse par le fait que le langage FORTRAN 77 est peu typé (aucune vérification de type n'a lieu). La norme FORTRAN 90 recommande l'emploi d'IMPLICIT NONE, qui impose de déclarer toutes les variables. Outre le gain en sécurité, la déclaration systématique permet d'éviter l'accumulation de variables redondantes qui pénalisent les performances.
- Le format fixe du source, vestige de l'ère où les cartes perforées tenaient lieu de mémoire de masse, est peu élégant et peut, en cas de malchance, se révéler source de bogues.
- Autres instructions désuètes (GOTO ventilé, IF calculé, EQUIVALENCE, BLOCK DATA...).

En résumé, le langage FORTRAN 77 privilégie la simplicité au détriment de la rigueur et encourage les mauvaises habitudes de programmation dans lesquelles, à moins d'une autodiscipline de fer, il est aisé de tomber.

La norme FORTRAN 90, bien que restant totalement compatible avec sa devancière, constitue une évolution radicale du langage qui présente bien des avantages :

- La modularité. Bien qu'encore procédural, FORTRAN 90 introduit la notion de module, proche de la programmation orientée objet. Il devient ainsi possible de segmenter un programme en modules indépendants composés de variables et de procédures, et d'exporter une partie de ces méthodes qui pourront être appelés par d'autres parties du programme. L'encapsulation favorise la genericité (les modules sont transférables d'un programme à l'autre) et facilite la conception et l'évolution du code (les modules constituent autant de 'boîtes noires' qui peuvent être modifiés en toute indépendance). En outre, la surdéfinition d'opérateurs permet de se doter de classes de base (vecteurs, matrices...) dont la manipulation est aisée. La déclaration de variables globales au sein d'un module est également autorisée.
- Les pointeurs. Ce sont des outils très puissants, qui ont fait le succès de langages tels que le C ou le C++. Absents en FORTRAN 77, ils peuvent

être émulés par des instructions telles que COMMON ou EQUIVALENCE. FORTRAN 90 les introduit de manière rigoureuse ; bien que moins généraux qu'en C ou C++, ils sont d'emploi facile et permettent souvent de substantiels gains en performance.

- L'allocation dynamique. Les tableaux peuvent à présent être dimensionnés à l'exécution, ce qui évite gaspillages de mémoire et risques de débordement. Ils peuvent en outre être désalloués lorsqu'ils ne sont plus nécessaires.
- Diverses facilités syntaxiques : format de source libre, section de tableaux..., qui rendent le code plus concis ou lisible.

La conversion des programmes existant au laboratoire en FORTRAN 90 avant leur extension a grandement facilité le cycle de développement. L'usage de modules, en particulier, devrait être étendu à tout développement futur ; il permet en effet de compartimenter les composantes disjointes d'un programme (méthodes d'optimisation ou d'exploration de surfaces d'énergie potentielle, méthodes de dynamique moléculaire classique ou semi-classique, versions plus ou moins complexes du potentiel modèle...), de sorte que le théoricien puisse, en un minimum de temps et d'effort, choisir les plus adaptées au problème en cours.

## 7.2. Langages interprétés

Les langages compilés sont indispensables pour mener à bien des calculs en un minimum de temps. Ils se révèlent en revanche peu pratiques pour bien des opérations (gestion des systèmes de fichiers, des tâches, de la communication entre processus...). Or, dans bien des cas, ces opérations ne sont pas limitantes en temps de calcul. Dès lors, l'emploi d'un langage de scripts puissant prend tout son intérêt. Le langage Python a été employé au cours de cette étude. Ses avantages sont les suivants :

- Il est peu typé ; il est ainsi possible d'englober dans un tableau ou une *hash-table* des valeurs de types différents.
- Il est semi-compilé, donc plus rapide qu'un langage de script classique.
- Sa gestion des fichiers et des chaînes de caractère est d'une puissance telle qu'il est aisé d'effectuer, en quelques lignes, des opérations qui seraient excessivement lourdes en FORTRAN. Ces fonctionnalités permettent

## 7. Développement et Programmation

l'analyse lexicale des fichiers de sortie ou la génération automatique de fichiers d'entrée d'un calcul.

- Il encapsule toutes les fonctionnalités du système d'exploitation : gestion des tâches, communication par *pipeline* entre tâches, synchronisation... Il est ainsi idéalement adapté à servir de lien entre unités compilées.
- Sa syntaxe est simple et peu contraignante.
- Il est orienté objet.
- Il est modulaire et extensible ; il existe par exemple un module dédié à l'algèbre linéaire.
- Il est gratuit, bien documenté (<http://www.python.org>), et son emploi se généralise dans de nombreux domaines scientifiques et techniques.

L'emploi du langage Python dans cette étude a résulté en un gain de temps appréciable, en automatisant nombre de tâches répétitives et en rendant possible l'exécution de tâches multi-étapes complexes. Le traitement et la synthèse de grands ensembles de résultats se seraient également révélés impraticables sans son utilisation.

L'apprentissage approfondi d'un langage de scripts moderne tel Python est un investissement rapidement rentabilisé pour tout travail méthodologique.



*Deuxième Partie*

LE TRANSFERT DE  
CHARGE



*Or, du sommet à la base, les mesures de la Grande Pyramide, en pouces égyptiens, sont de 161 000 000. Combien d'âmes humaines ont vécu sur la terre depuis Adam jusqu'à ce jour ? Une bonne approximation donnerait quelque chose entre 153 000 000 et 171 000 000.*

*Piazzzi Smyth, *Our Inheritance in the Great Pyramid* (1880)*

*'Messieurs, je vous invite à aller mesurer ce kiosque. Vous verrez que la longueur de l'éventaire est de 149 cm, c'est-à-dire un cent-milliardième de la distance Terre-Soleil. La hauteur postérieure divisée par la largeur de l'ouverture fait  $176 : 56 = 3,14$ . La hauteur antérieure est de 19 décimètres, c'est-à-dire égale au nombre d'années du cycle lunaire grec. La somme des hauteurs des deux arêtes antérieures et des deux arêtes postérieures fait  $190 \times 2 + 176 \times 2 = 732$ , qui est la date de la victoire de Poitiers. L'épaisseur de l'éventaire est de 3,10 centimètres et la largeur de l'encadrement de l'ouverture de 8,8 centimètres. En remplaçant les nombres entiers par la lettre alphabétique correspondante, nous aurons  $C_{10}H_8$ , qui est la formule de la naphthaline.*

*—Fantastique, dis-je, vous avez essayé ?*

*—Non, dit Agliè. Un certain Jean-Pierre Adam l'a fait sur un autre kiosque. J'imagine que tous les kiosques de la loterie ont plus ou moins les mêmes dimensions. Avec les nombres on peut faire ce qu'on veut.*

*Umberto Eco, *Le Pendule de Foucault**



## 1. Description du Transfert de Charge

# 1. Description du Transfert de Charge

### 1.1. Introduction

**S**'il est généralement aisé de concevoir, lors de la mise au point d'un potentiel modèle, une formulation analytique pour des contributions à l'énergie d'interaction telles que l'énergie électrostatique ou de dispersion-repulsion, il n'en est pas de même en ce qui concerne le terme à transfert de charge. Il convient en effet de rendre correctement compte de ses caractéristiques physiques, notamment son comportement avec la distance, son anisotropie spatiale et le fait qu'il s'agit d'un terme à  $n$  corps. Compte tenu des applications envisagées pour ce potentiel, il faut également qu'il soit suffisamment simple pour pouvoir être calculé en temps réel, dans des simulations de dynamique moléculaire par exemple.

Plusieurs modélisations du phénomène de transfert de charge ont été proposées dans la littérature. La plus aboutie d'entre elles est sans doute celle employée par Jortner dans l'ADN<sup>98,99,191</sup>. Elle est fondée sur l'observation expérimentale que la constante de vitesse  $k$  associée à la réaction de transfert de la charge entre un donneur D et un accepteur A le long de la double hélice (appelée 'pont') peut dépendre beaucoup, ou au contraire très peu, de la distance D-A, en fonction notamment de la séquence des brins. Ceci conduit Jortner à postuler la coexistence de deux mécanismes en compétition.

Le premier d'entre eux, dénommé super-échange, est de type bicentrique mono-étape. Il obéit à l'équation de Marcus-Levich-Jortner<sup>17</sup> :

$$k = \frac{2\pi}{\hbar} V_0^2 F_{DA} e^{-\beta R}$$

où  $F_{DA}$  est la densité Franck-Condon entre les états vibroniques respectifs du donneur et de l'accepteur et  $R$  la distance entre ces molécules. Le facteur  $V_0$  dépend quant à lui des couplages donneur-premier site du pont et accepteur-dernier site du pont, ainsi que de la différence énergétique entre les états vibroniques de D et A. Ce mécanisme est caractérisé par une décroissance exponentielle avec la distance donneur-accepteur, et tend à se produire lorsqu'il existe un couplage non résonant entre états vibroniques initiaux et finaux.

Le second mécanisme, dit de *hopping*, est quant à lui de type multi-étapes et s'effectue *via* la chaîne séparant D et A. Il est favorisé par l'existence de couplages résonants entre états vibroniques du donneur, de l'accepteur et du pont. Le transfert s'effectue par l'intermédiaire de chacune des bases du pont ; les constantes de vitesse régissant chaque étape dépendent du couplage et de la densité Franck-Condon entre états vibroniques de sites voisins. Analogue au modèle de polaron dans les solides monodimensionnels<sup>192</sup>, ce modèle est caractérisé par une faible dépendance vis-à-vis de la distance D-A (typiquement en  $N^{-1}$ ,  $N$  étant le nombre de molécules constituant le pont et  $\eta = 1$  ou  $2^{98}$ ).

Une autre approche, fréquemment employée, consiste à décrire le transfert de charge par le biais d'un hamiltonien effectif dont le but est de restreindre la complexité et la dimensionnalité du système à l'étude. Les travaux de ce type se différencient par la nature de l'hamiltonien effectif et les modèles employés lors de sa constitution. Le modèle VB/HF de Castet et Ducasse<sup>193,194</sup> introduit ainsi un hamiltonien effectif de type Pariser-Pople-Parr<sup>195,196</sup> écrit dans la base des formes *valence bond* à charge localisée. Les paramètres impliqués dans le calcul des éléments diagonaux (énergies de site) et non diagonaux (couplages responsables du transfert de la charge) sont tirés de calculs Hartree-Fock de monomères et de dimères. Ce modèle a été appliqué avec succès à l'ADN<sup>197-199</sup>. Dans ce même système, Orlandi et Troisi<sup>200,201</sup> partitionnent pour leur part l'hamiltonien effectif en deux sous-espaces, dont l'un regroupe les états électroniques du donneur et de l'accepteur de même que les couplages afférents.

La détermination des divers couplages électroniques entre états du donneur, de l'accepteur et du pont est l'étape délicate de chacune des méthodologies précédemment évoquées. Leur calcul *ab initio* dans des systèmes de tailles importantes est coûteux<sup>202</sup>. L'existence éventuelle de sites inéquivalents complique encore la situation, en multipliant le nombre de couplages distincts qu'il faudra évaluer ; or, le désordre structural dans l'ADN, responsable de la levée de dégénérescence, joue un rôle primordial dans le mécanisme de transfert par super-échange<sup>89</sup>. Déterminés dans un premier temps à l'aide de méthodes semi-empiriques<sup>203</sup>, les couplages peuvent à l'heure actuelle être déterminés *ab initio* au niveau Hartree-Fock, mis à la portée de gros systèmes par les progrès technologiques informatiques. Se pose cependant le problème de rendre compte d'un couplage entre formes diabatiques (à charge localisée sur un site) à partir d'une fonction d'onde adiabatique (à charge *a priori* délocalisée) ; ce couplage est généralement pris égal à la moitié du *splitting* entre états ioniques adiabatiques, lui-même déterminé, en vertu du théorème de Koopmans<sup>120</sup>, comme la

## 1. Description du Transfert de Charge

différence d'énergie entre deux orbitales moléculaires occupées du neutre issues d'un calcul SCF<sup>96,97,100,204</sup>. Même étendu pour traiter les systèmes à forte séparation de charge<sup>205</sup>, ce modèle demeure assez simple. La détermination directe des couplages non adiabatiques recherchés est possible (des calculs d'interaction de configurations ont été employés à cet effet au cours du présent travail), mais généralement coûteuse. Si l'emploi de méthodes de la fonctionnelle de la densité pour le calcul des couplages semble prometteur<sup>206</sup>, une modélisation précise des couplages responsables du transfert de charge, dans l'ADN comme dans les agrégats moléculaires au sens large, est loin d'être dénuée d'intérêt.

Une première solution consiste à décrire le transfert de charge comme une contribution à l'énergie d'interaction intermoléculaire, retranscrite comme les autres contributions par des formules analytiques plus ou moins complexes à l'intérieur d'un potentiel modèle applicable à l'exploration globale de surfaces d'énergie potentielle, les méthodes Monte Carlo ou les simulations de dynamiques moléculaires. Deux exemples de telles modélisations sont présentés ci-après.

La première fut introduite par Gresh<sup>207</sup> en 1985, et suppose le terme à transfert de charge proportionnel à une intégrale de recouvrement entre orbitales et à la différence de potentiel d'ionisation entre les deux entités concernées. Si l'emploi d'intégrales de recouvrement introduit l'indispensable anisotropie spatiale propre au transfert de charge, ce modèle n'en reste pas moins fondé sur des interactions additives à deux corps ; de surcroît, sa complexité le rend difficilement généralisable.

Une autre approche fut présentée par Stone<sup>208</sup> en 1993 ; le transfert de charge y est décrit comme une somme de termes à deux corps décroissant exponentiellement avec la distance. Bien moins complexe que le modèle de Gresh, ce modèle souffre cependant de son isotropie spatiale et se cantonne à l'hypothèse d'additivité de termes à deux corps.

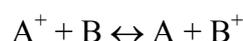
Contrairement à ces modèles, les méthodes adoptant le formalisme d'hamiltonien effectif découplent les énergies de site (énergies associées aux fonctions d'onde formant la base de l'hamiltonien) des couplages électroniques responsables du transfert. Les premiers peuvent n'être constitués de quantités locales (potentiels d'ionisation<sup>200,201</sup> ou énergies orbitales<sup>199</sup>, énergies de répulsion de site<sup>198,199</sup>...). Idéalement, ils doivent cependant retranscrire les effets d'environnement liés à la géométrie du système, permettant la prise en compte du désordre structural. Les seconds sont classiquement modélisés comme proportionnels à une intégrale de recouvrement entre orbitales moléculaires des monomères concernés, le facteur de proportionnalité constituant un paramètre généralement tiré de calculs *ab initio*.

Bien que délicate, la mise au point d'un modèle de transfert de charge est indispensable pour décrire nombre de systèmes. L'exemple de l'ADN a déjà été évoqué. Il en va de même dans les agrégats aromatiques ionisés. Les agrégats aromatiques neutres sont caractérisés par une énergie d'interaction très faible, essentiellement due aux contributions de dispersion et de répulsion ; une fois ionisés, en revanche, l'énergie d'interaction dans ces agrégats augmente d'un ordre de grandeur, tandis que la structure géométrique subit un réarrangement drastique. Il sera vérifié, au cours de cette étude, que ces profondes modifications sont la conséquence du transfert de charge au sein de l'agrégat, dont l'importance est telle qu'elle en dicte les propriétés.

S'il adopte la même philosophie et les mêmes objectifs que ses devanciers, notamment par le choix d'un hamiltonien effectif et l'assimilation des couplages électroniques à une intégrale de recouvrement, le modèle de transfert de charge décrit dans ces pages relève d'une interaction qui décroît exponentiellement avec la distance, est à  $n$  corps, spatialement anisotrope et ajustée sur des calculs de couplage *ab initio* de bonne qualité. Il offre de surcroît la prise en compte des effets d'environnement (de polarisation notamment). Il est présenté ci-dessous.

### 1.2. Le modèle à hamiltonien effectif

Le transfert de charge entre deux molécules A et B, au sein d'un agrégat portant une charge globale +1, s'écrit :



Pour peu que l'énergie des formes à charge localisée ( $A^+B$ ) et ( $AB^+$ ) soient proches (comme c'est le cas si les molécules A et B sont équivalentes), le mélange des deux formes stabilise l'état fondamental de l'agrégat ( $AB$ )<sup>+</sup>. L'interaction entre formes à charge localisée (dites aussi formes mésomères) conduit naturellement à l'emploi de la méthode *valence bond*, succinctement présentée en première partie. La description des états électroniques de l'agrégat à l'aide d'un hamiltonien effectif, exprimé dans la base de ces formes à charge localisée, est à présent abordée.

Le transfert de charge est la conséquence d'un transfert d'électron au sein de l'agrégat, qui est décrit à l'aide du potentiel effectif  $V$  :

## 1. Description du Transfert de Charge

$$V(\vec{r}) = \sum_m V_m(\vec{r})$$

qui consiste en une somme des contributions  $V_m$  de toutes les entités de l'agrégat. De manière générale, la séparation de  $V$  en ses composantes n'est pas triviale ; elle ne l'est que dans le cas d'un transfert d'électron bimoléculaire entre molécules supposées isolées. Les contributions  $V_m$  traduisent l'action du cortège électronique total de chaque fragment sur l'électron transféré. L'équation de Schrödinger monoélectronique pour cet électron s'écrit alors :

$$(T_{el} + V_m(\vec{r}))\phi_m(\vec{r}) = E_m\phi_m(\vec{r})$$

La première valeur propre  $E_m$  est associée au site  $m$  de l'agrégat, et dénommée en conséquence énergie de site. Les fonctions d'onde  $\{\phi_m\}$  fournissent une représentation diabatique du problème. Elles ne forment généralement pas une base orthonormée, c'est-à-dire qu'il subsiste des recouvrements tels que  $\langle\phi_m|\phi_n\rangle \neq \delta_{mn}$ .

L'équation de Schrödinger électronique du supersystème s'écrit quant à elle :

$$(T_{el} + V)|\Psi\rangle = E|\Psi\rangle$$

Dans la base diabatique des fonctions d'onde  $\{\phi_m\}$ , la fonction d'onde monoélectronique  $|\Psi\rangle$  s'écrit :

$$|\Psi\rangle = \sum_m c_m |\phi_m\rangle$$

en multipliant l'équation de Schrödinger à gauche par  $\langle\phi_n|$  :

$$\langle\phi_n|T_{el} + \sum_k V_k|\Psi\rangle = E\langle\phi_n|\Psi\rangle$$

c'est-à-dire :

$$\sum_m c_m \left( E_n \langle\phi_m|\phi_n\rangle + \sum_{k \neq m} \langle\phi_n|V_k|\phi_m\rangle \right) = E \sum_n c_n \langle\phi_m|\phi_n\rangle$$

Cette équation contient d'une part des intégrales de recouvrement entre formes diabatiques  $\langle\phi_m|\phi_n\rangle$ , et d'autre part des intégrales à trois centres de la forme  $\langle\phi_n|V_k|\phi_m\rangle$ ,  $k \neq m \neq n$ . Deux approximations sont alors faites. Il est tout d'abord supposé que les fonctions d'onde diabatiques  $\{\phi_m\}$  forment une base orthogonale, c'est-à-dire que  $\langle\phi_m|\phi_n\rangle = \delta_{mn}$ . En second lieu, les intégrales à trois centres, qui sont généralement bien plus petites que les intégrales à deux centres, sont également négligées.

Dans ces conditions, les éléments de matrice de l'hamiltonien électronique du supersystème s'écrivent :

$$\langle \phi_m | \hat{H}_{el} | \phi_n \rangle = \underbrace{\delta_{mn} \left( E_m + \sum_{k \neq m} \langle \phi_m | V_k | \phi_m \rangle \right)}_{\text{éléments diagonaux}} + \underbrace{(1 - \delta_{mn}) \langle \phi_m | T_{el} + V_m + V_n | \phi_n \rangle}_{\text{éléments non diagonaux}}$$

Compte tenu de l'équation de Schrödinger monoélectronique définissant  $E_n$  et  $\phi_n$ , les éléments de matrice non diagonaux de l'hamiltonien peuvent être réécrits :

$$\begin{aligned} \langle \phi_m | T_{el} + V_m + V_n | \phi_n \rangle &= \frac{1}{2} \langle \phi_m | (T_{el} + V_m) + (T_{el} + V_n) + (V_m + V_n) | \phi_n \rangle \\ &= \frac{1}{2} \langle \phi_m | E_m + E_n + (V_m + V_n) | \phi_n \rangle \\ &= \frac{1}{2} \langle \phi_m | V_m + V_n | \phi_n \rangle = V_{mn} \end{aligned}$$

La grandeur  $V_{mn}$  est généralement dénommée intégrale de transfert ou couplage inter-état, voire tout simplement couplage. L'hamiltonien électronique du complexe s'écrit donc finalement :

$$\hat{H}_{el} = \sum_m E_m | \phi_m \rangle \langle \phi_m | + \sum_{m,n} V_{mn} | \phi_m \rangle \langle \phi_n |$$

La détermination, puis la résolution, de l'équation de Schrödinger associée à un agrégat dans la base diabatique des formes à charge localisée nécessite donc le calcul de deux types de termes : les énergies de site  $E_m$  et les intégrales de transfert  $V_{mn}$ . Il est intéressant de noter que le modèle d'hamiltonien effectif permet la prise en considération de mécanismes de transfert simultanés ou concurrents, comme ceux de super-échange et de *hopping* qui coexistent dans l'ADN (voir paragraphe 1.1).

Les méthodes employées pour le calcul de chaque type d'élément de matrice, adaptées à la nature chimique des systèmes à l'étude, font l'objet des paragraphes suivants.

### 1.3. Calcul des éléments diagonaux de l'hamiltonien effectif

Les éléments diagonaux de l'hamiltonien électronique d'un agrégat, dans la base diabatique des formes à charge localisée, constituent autant d'énergies de site égales à l'énergie de la forme *valence bond* correspondante. Le calcul de ces énergies a été effectué au moyen du potentiel modèle polarisable décrit plus haut : la retranscription des effets d'environnement liés à la géométrie du supersystème est ainsi rendue possible. La

## 1. Description du Transfert de Charge

paramétrisation du potentiel constitue une étape préliminaire indispensable, dans l'optique de son application à des agrégats de molécules aromatiques.

### 1.3.1. Procédure d'ajustement

L'énergie d'interaction déterminée à l'aide du potentiel modèle est la somme de trois contributions indépendantes : les énergies d'interaction électrostatique, de polarisation dipolaire et de dispersion-répulsion. Chacune d'entre elles peut être ajustée indépendamment des deux autres, selon la procédure suivante :

- L'énergie électrostatique résulte de l'interaction entre distributions multipolaires multacentriques, déduites de la matrice densité électronique à une particule calculée *ab initio* pour chacun des monomères de l'agrégat. Il s'agit de la partie la plus précise du potentiel modèle : les erreurs typiques constatées par rapport à l'*ab initio* sont inférieures à 1%<sup>112</sup>. Le choix de la méthode et de la base de gaussiennes du calcul *ab initio* dont est déduite la distribution constituent les paramètres de la contribution électrostatique.
- L'énergie de polarisation dipolaire est calculée de manière autocohérente à partir de tenseurs de polarisabilité de site (atomes et milieux de liaisons), qui dépendent de la nature du site (type d'atome et de liaison) et du nombre d'électrons impliqués (présence éventuelle de doublets non liants sur les atomes, multiplicité variable des liaisons). Ces tenseurs locaux constituent autant de paramètres, dont l'ajustement s'effectue par comparaison avec (i) le tenseur de polarisabilité moléculaire expérimental ou *ab initio*, qui doit être aussi proche que possible de la somme des tenseurs locaux de la molécule, et (ii) l'énergie d'interaction avec une charge ou un cation peu polarisable, calculée *ab initio* à moyenne et longue distance, et comparée à la somme de la contribution électrostatique ajustée lors de l'étape précédente et de la contribution de polarisation dont l'ajustement est en cours. L'emploi d'une charge polarisante assure la nullité des composantes de dispersion et de répulsion dont l'influence pourrait contaminer l'ajustement, mais peut fournir des résultats erronés si la probabilité de présence d'électrons de la molécule polarisable au voisinage de la charge n'est pas nulle, comme ce peut être le cas si la base de gaussiennes employée compte des orbitales diffuses. Étant donné l'ordre de grandeur des contributions de dispersion et de répulsion aux distances envisagées, et compte tenu du fait que la

prédiction correcte de la polarisabilité d'une molécule dépend fortement de la qualité de la base et de l'inclusion de diffuses en son sein, le choix d'un cation semble plus indiqué et a été retenu. La comparaison avec le tenseur moléculaire teste la validité de l'hypothèse d'additivité des tenseurs locaux, tandis que la comparaison avec l'énergie d'interaction *ab initio* sonde le degré d'adéquation des tenseurs locaux.

- L'énergie de dispersion-répulsion est sans doute la plus complexe à modéliser, ne serait-ce que par sa nature purement quantique. Le fait que cette contribution regroupe plusieurs effets la rend également délicate à appréhender. Pour ces raisons, elle est ajustée en dernier lieu, une fois les contributions électrostatiques et de polarisation correctement mises en place. Le terme de dispersion-répulsion dépend de deux ensembles de paramètres atomiques qui sont fonction de la nature de l'atome. Ces paramètres sont ajustés par comparaison de l'énergie d'interaction obtenue à l'aide du potentiel modèle avec son équivalent *ab initio*, au voisinage de la position d'équilibre, supposant correctes les contributions électrostatiques et de polarisation étalonnées au cours des deux étapes précédentes.

### 1.3.2. Modèle électrostatique

Comme mentionné ci-dessus, la contribution électrostatique à l'énergie d'interaction est supposée égale à l'énergie d'interaction entre jeux de multipôles distribués sur chacune des entités de l'agrégat. Ces développements multipolaires multicentriques sont déduits de la matrice densité à une particule obtenue par un calcul *ab initio* réalisé sur chacun des monomères. Il a été montré<sup>209,210</sup> que la méthode MP2 en base triple- $\zeta$  6-311++G(3d,3p) fournit une retranscription fidèle des propriétés de la molécule de benzène. L'algorithme employé pour le développement multipolaire multicentrique de la densité électronique *ab initio* s'accommode cependant mal de la présence de fonctions de base diffuses, qui donnent lieu à des singularités dans la matrice de recouvrement entre orbitales atomiques pouvant conduire à des résultats aberrants. En conséquence, les développements multipolaires multicentriques de toutes les molécules aromatiques traitées ont été tirés de calculs MP2 en base 6-311G(2d,2p), qui reproduisent fidèlement le quadrupôle du benzène (il diffère de 2% de sa valeur en base 6-311++G(3d,3p)).

## 1. Description du Transfert de Charge

### 1.3.3. Modèle de polarisation

Les molécules aromatiques considérées ici sont dotées d'une grande anisotropie spatiale (elles sont planes). Une anisotropie au sein même de ce plan peut par ailleurs exister pour des molécules telles que le naphthalène ou l'anthracène (coexistence d'un petit et d'un grand axe de symétrie). Il est donc probable que le tenseur de polarisabilité de telles entités retranscrive lui aussi cette anisotropie, de sorte que le modèle de polarisabilité isotrope classiquement associé à l'emploi du présent potentiel modèle ne puisse convenir. Les expériences et les calculs *ab initio* montrent en effet une polarisabilité hors-plan bien plus faible que la polarisabilité dans le plan, pour l'ensemble des espèces considérées. L'anisotropie petit axe-grand axe dans le plan moléculaire se révèle également de plus en plus marquée, du benzène à l'anthracène.

Le modèle isotrope associe à chaque atome et chaque liaison, selon son type et le nombre d'électrons impliqués, un tenseur de polarisabilité à symétrie cylindrique, dont la liaison constitue l'axe de révolution et doté d'une composante longitudinale ( $\alpha_{//}$ ) et deux composantes transversales égales ( $\alpha_{\perp}$ ). Le tenseur de polarisabilité local, positionné sur l'atome ou au milieu de la liaison, est le tenseur scalaire  $\bar{\alpha} = \frac{1}{3}(\alpha_{//} + 2\alpha_{\perp})$ . La somme de ces tenseurs locaux (elle aussi scalaire) doit avoisiner le tenseur moléculaire moyen, calculé *ab initio* ou déterminé expérimentalement.

L'extension la plus évidente de ce modèle consiste à conserver la symétrie cylindrique des tenseurs locaux. Dans les systèmes aromatiques, cependant, il n'est pas acceptable d'associer un seul et même paramètre aux deux composantes transversales des tenseurs de polarisabilité locaux. En effet, dans les molécules aromatiques pour lesquelles toutes les liaisons appartiennent à un même plan, ces deux directions ne sont en rien équivalentes : l'une d'entre elle appartient au plan, l'autre y est perpendiculaire. L'anisotropie hors-plan mentionnée ci-dessus ne saurait être retranscrite dans le cadre de ce modèle. Cette observation avait poussé Mulder et van Dijk<sup>139</sup> à proposer, dès 1979, un modèle de polarisabilité propre aux molécules aromatiques et hétéroaromatiques. Celui-ci part du constat que coexistent, dans ces espèces, deux systèmes électroniques distincts : un système  $\sigma$  situé dans le plan moléculaire, polarisable à-peu-près de la même façon dans toutes les directions de l'espace, et un système  $\pi$  perpendiculaire au plan, très polarisable le long des liaisons mais peu polarisable perpendiculairement au plan. Les tenseurs atomiques et de liaisons résultent

donc de la somme des contributions tensorielles des électrons  $\sigma$  et  $\pi$ . La figure 2.1 résume les caractéristiques de ces modèles de polarisabilité successifs.

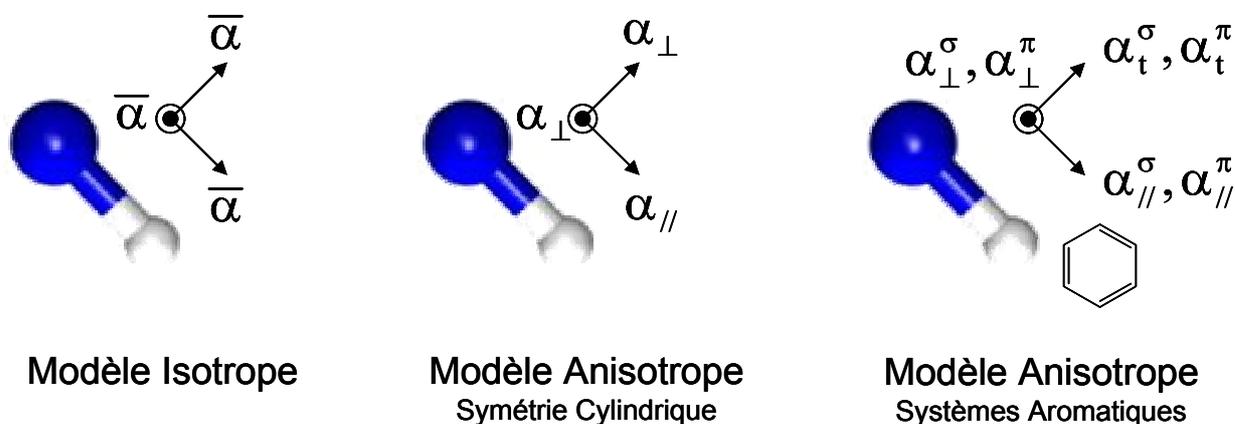


Figure 2.1

Le modèle de Mulder n'est cependant pas directement applicable à un potentiel modèle tel que celui utilisé ici ; son but n'est pas de fournir des tenseurs de site réalistes mais uniquement d'assurer que la somme desdits tenseurs soit le plus proche possible du tenseur moléculaire expérimental. Le problème est alors surdéterminé, c'est-à-dire que plusieurs jeux de paramètres distincts satisfont le cahier des charges. Ceci pousse Mulder à choisir pour certains des paramètres (notamment la polarisabilité  $\pi$  transversale  $\alpha_t^\pi$ ) des valeurs arbitraires sans sens physique<sup>139</sup>. Il a été procédé, dans le cadre de cette étude, à une redistribution des différentes contributions, dans le but d'obtenir des tenseurs de site raisonnables. L'ajustement a été effectué en référence à l'énergie d'interaction *ab initio* dans le système benzène – ion sodium (voir plus bas), la valeur du tenseur moléculaire global du benzène ayant été préservée autant que possible au cours du processus.

Les tenseurs moléculaires obtenus avec le modèle isotrope et le modèle anisotrope adapté aux systèmes aromatique sont comparés à leurs équivalents *ab initio*, par leurs valeurs moyennes et leur anisotropies (définies plus haut), dans le tableau 2.1. La méthode MP2 en base triple- $\zeta$  6-311G++ a été employée dans le cas du benzène et du naphthalène ; ce niveau de calcul, s'étant révélé trop coûteux dans le cas de l'anthracène, a été substitué pour ce système par une approche *coupled-perturbed Hartree-Fock*<sup>211</sup> dans cette même base.

## 1. Description du Transfert de Charge

	<b>Modèle Isotrope</b>	<b>Modèle Anisotrope</b>		<b><i>Ab Initio</i></b>	
	$\bar{\alpha}$ (Å <sup>3</sup> )	$\bar{\alpha}$ (Å <sup>3</sup> )	$\Delta\alpha$ (Å <sup>3</sup> )	$\bar{\alpha}$ (Å <sup>3</sup> )	$\Delta\alpha$ (Å <sup>3</sup> )
<b>Benzène</b>	10,6	10,5	5,8	9,9	5,5
<b>Naphtalène</b>	16,2	16,3	9,5	15,9	10,4
<b>Anthracène</b>	27,6	27,8	20,5	26,1	24,4

Les deux modèles, isotrope comme anisotrope, fournissent une valeur très correcte du tenseur moléculaire. En revanche, l'importante anisotropie qui caractérise les tenseurs moléculaires ne peut évidemment être retranscrite qu'à l'aide du modèle anisotrope.

La prédiction correcte du tenseur moléculaire global, si elle est nécessaire pour justifier l'hypothèse d'additivité des tenseurs locaux, n'est pas directement exploitée dans les calculs de potentiels modèle. Certes, elle assure un comportement correct de l'énergie de polarisation à la limite de distances très supérieures aux dimensions moléculaires ; aux distances d'équilibre typiques, une molécule n'en reste pas moins caractérisée, dans le présent modèle, par une distribution de tenseurs de polarisabilité locaux.

Afin de tester la qualité de cette distribution, les énergies d'interaction sont comparées à celles issues de calculs *ab initio*, sur le système benzène-Na<sup>+</sup>. Ce système-test est approprié à l'étude du modèle de polarisation pour deux raisons. Tout d'abord, la charge portée par l'ion sodium polarise le nuage électronique de la molécule de benzène, l'ion Na<sup>+</sup> étant quant à lui peu polarisable en retour ( $\alpha_{\text{Na}} = 0,15 \text{ \AA}^3$ <sup>212</sup>) ; cet ion se comporte donc comme une charge-test jugeant la polarisabilité du benzène. De plus, les premiers potentiels d'ionisation respectifs du benzène et du sodium ( $\text{PI}_{\text{Na}} = 4,96 \text{ eV}$ <sup>212</sup> et  $\text{PI}_{\text{benzène}} = 9,25 \text{ eV}$ <sup>213</sup>) sont suffisamment éloignés pour que tout transfert de charge entre les deux entités (que le potentiel modèle ne saurait prendre en considération) soit proscrit. Deux types de configurations de l'agrégat benzène-Na<sup>+</sup> sont considérées. Dans les structures dites planes, l'ion sodium est dans le plan de la molécule de benzène. Dans les structures dites pontées, l'ion se situe sur l'axe perpendiculaire au plan du benzène et passant par son centre de gravité. Ces deux types d'isomère coexistent effectivement au sein des agrégats benzène-ion ; ils permettent ici de jauger les différentes composantes des tenseurs de site.

La figure 2.2 compare, sur divers agrégats plans et pontés, les énergies d'interaction *ab initio* (MP2/6-311G++(3d,3p)-pseudopotentiel petit cœur Na<sup>+214</sup> + correction de BSSE) et en potentiels modèle anisotrope et isotrope, une fois ajustés les paramètres gouvernant le calcul des tenseurs locaux.

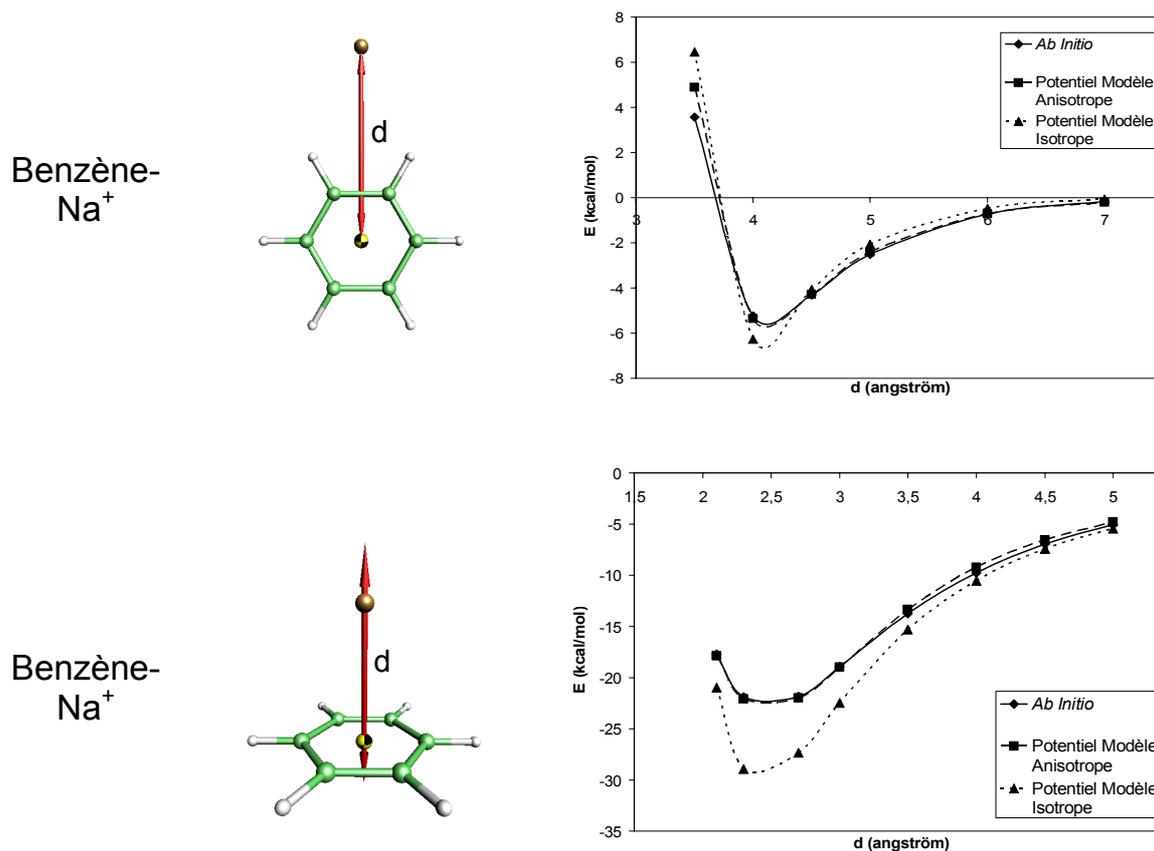


Figure 2.2

L'accord avec le calcul *ab initio* est bien meilleur avec le potentiel modèle anisotrope qu'avec son homologue isotrope. Ceci est particulièrement flagrant dans le cas des isomères pontés, pour lesquels l'erreur commise en isotrope peut s'élever à 8 kcal/mol. Dans ces structures, le potentiel modèle isotrope surestime la stabilisation due à l'énergie de polarisation ; en effet, l'emploi d'un tenseur isotrope surévalue la polarisabilité hors-plan, en réalité bien plus faible que les deux autres composantes.

Les énergies de stabilisation des isomères plans sont bien moindres que celle des configurations pontées : l'ion sodium n'interagit de manière effective qu'avec quelques atomes de la molécule de benzène, contre tous les atomes dans les isomères pontés. L'écart absolu entre niveaux de calcul est en conséquence moins marqué ici ; celui observé entre les potentiels isotrope et anisotrope tient essentiellement à une meilleure paramétrisation du modèle de polarisabilité anisotrope, alors que le modèle isotrope emploie les valeurs par défaut de Lefèvre<sup>138</sup>.

Le modèle de polarisation développé et implémenté, dans le cadre de cette étude, pour les molécules aromatiques, est un indéniable progrès par rapport à la version isotrope.

## 1. Description du Transfert de Charge

Son application à l'étude d'agrégats neutres et chargés de telles molécules est légitime. Des essais d'application à des molécules hétéroaromatiques (mono-, di- et triazines) ont cependant mis en lumière un manque de généralité, probablement causé par l'invalidité de l'hypothèse de transférabilité entre paramètres atomiques dans ces systèmes, et par la modélisation très délicate de la contribution tensorielle des électrons de paires libres. Plutôt que de multiplier les types d'atome ou les paramètres (le modèle anisotrope en compte déjà beaucoup), l'approche proposée par Chipot<sup>215</sup> pourrait constituer une voie d'amélioration du modèle. Elle emploie la théorie *Atoms in Molecules* de Bader<sup>216</sup> pour générer une grille d'énergies d'induction, utilisée par la suite pour ajuster un jeu de polarisabilités distribuées. Elle pourrait en conséquence fournir un ensemble de tenseurs de site adaptés à chaque molécule, au prix d'un unique calcul *ab initio*, et sans ajustement manuel de quelque paramètre que ce soit.

### 1.3.4. Modèle de dispersion-répulsion

Comme indiqué en première partie, le modèle de dispersion-répulsion s'articule autour d'un jeu de paramètres atomiques supposés transférables entre type d'atomes. Les molécules concernées par cette étude sont constituées d'atomes de carbone aromatique et d'hydrogène. Les paramètres de l'hydrogène, déterminés lors d'une précédente étude<sup>113</sup> sur d'autres systèmes chimiques, ne seront pas modifiés ici. Restent donc les deux paramètres  $R$  et  $k$  associés à l'atome de carbone aromatique.

Le dimère du benzène a été choisi comme système modèle pour la mise au point de ces paramètres. La structure optimale pour cet agrégat est en effet caractérisée par une énergie d'interaction très faiblement attractive (environ 2 kcal/mol), consistant en grande partie de composantes dispersives et répulsives (l'article inséré ci-après<sup>217</sup> fournit de plus amples précisions à ce sujet). Le système (benzene<sub>2</sub>) représente donc un cas d'école pour la mise au point des paramètres d'un modèle de dispersion-répulsion. Un problème subsiste néanmoins : la description *ab initio* de ce système requiert l'emploi de la méthode CCSD-T, les niveaux de calcul inférieurs (MP2 notamment) se révélant incapables d'atteindre un degré suffisant de précision ; le lecteur est renvoyé à l'article<sup>217</sup> pour plus de détails. Hobza *et al*<sup>218</sup> ont publié ce qui est à ce jour l'étude la plus exhaustive sur le dimère de benzène, comportant un échantillonnage de configurations sur la surface d'énergie potentielle CCSD-T. Ces calculs coûteux ont été employés pour calibrer le potentiel modèle. Ils consistent en 37 géométries distinctes d'énergies respectives  $E_i^{\text{CCSD-T}}$ . Soit  $E_i^{\text{PM}}$  l'énergie obtenue à l'aide du potentiel

modèle en ces points de l'espace des configurations. La fonction coût à optimiser s'écrit alors :

$$f(R, k) = \sqrt{\sum_{i=1}^{37} (E_i^{\text{PM}}(R, k) - E_i^{\text{CCSD-T}})^2}$$

Les valeurs de  $R$  et  $k$  optimisant la fonction  $f$  sont déterminées à l'aide d'un algorithme génétique à paramètres continus ( $1,0 \leq R \leq 3,0$ ,  $0,5 \leq k \leq 1,5$ , 5000 individus, taux de géniteurs 50 %, taux de mutation 10 %, 81 générations avant convergence). Les valeurs optimales après optimisation sont  $R_{\text{opt}} = 1,78$  et  $k_{\text{opt}} = 1,11$ .

La recherche des paramètres optimaux a mis en lumière une difficulté conceptuelle dans le modèle de dispersion-répulsion employé ici : les paramètres  $k$  et  $R$  ne peuvent être considérés comme indépendants. Ceci se traduit par la non-unicité du doublet  $(R_{\text{opt}}, k_{\text{opt}})$ . Le paramètre  $R$  pouvant être assimilé au rayon de van der Waals de l'atome, le jeu de paramètres retenu est celui qui, pour une valeur raisonnablement proche de ce rayon (qui est de 1,8 Å pour l'atome de carbone<sup>141</sup>), optimise la valeur de la fonction coût. La figure 2.3 est une représentation de la surface d'énergie potentielle associée à la fonction coût  $f$ , dans une certaine gamme de valeurs de  $R$  et de  $k$  (pour des raisons de temps de calcul, le nombre de configurations pris en compte dans  $f$  a été réduite à 8 configurations pour les besoins de cette représentation). La présence d'une 'vallée' quasiment linéaire sur cette surface est caractéristique de l'interdépendance entre les paramètres.

## 1. Description du Transfert de Charge

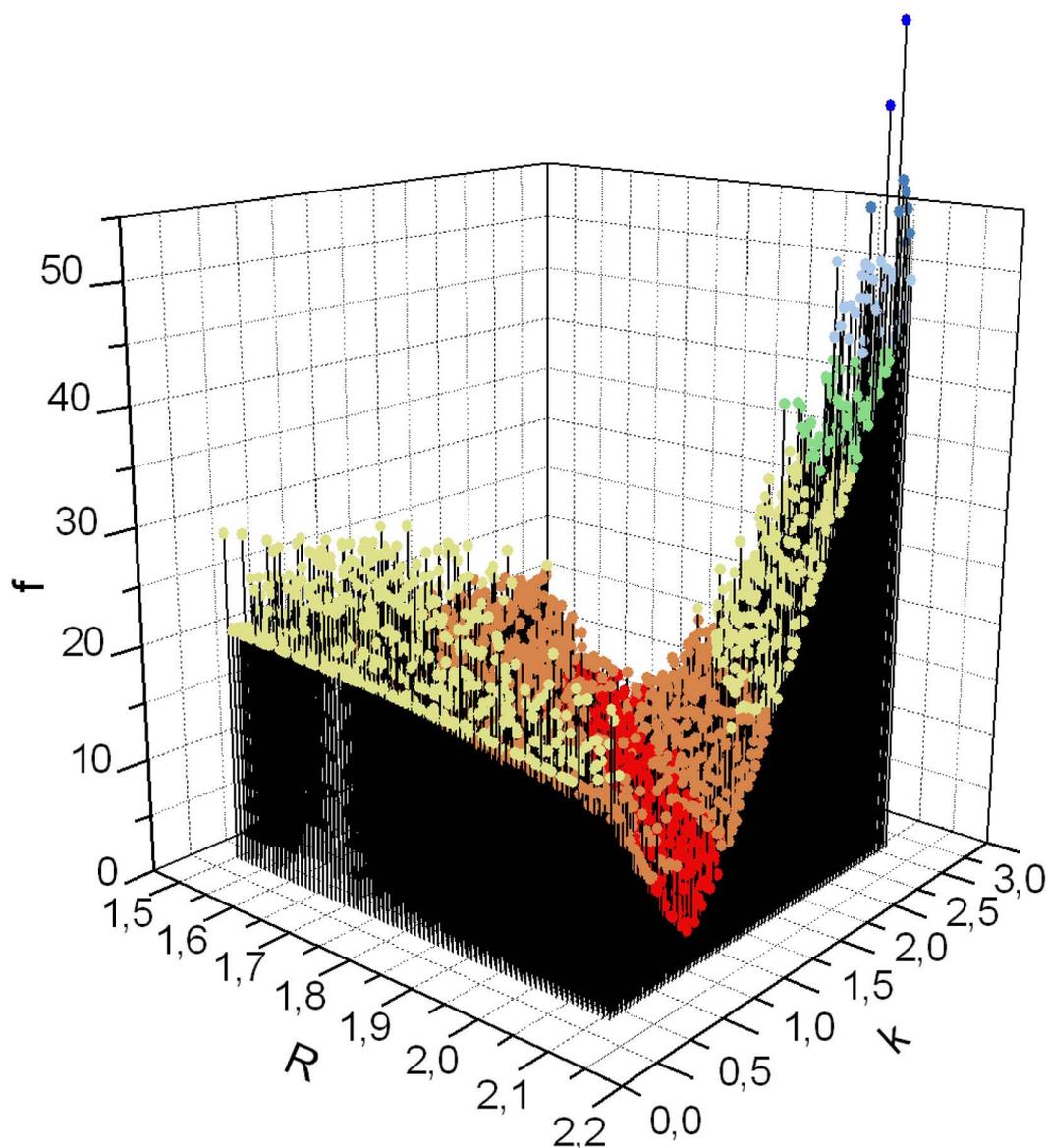


Figure 2.3

La figure 2.4 présente les énergies d'interaction caractérisant le dimère de benzène de géométrie empilée<sup>217</sup>, à diverses distances intermoléculaires, calculées *ab initio* (SCF et MP2/6-311G++(3d,3p) + correction de BSSE, CCSD-T<sup>218</sup>) et par le potentiel modèle ajusté. La méthode SCF prédit un comportement dissociatif, tandis que la correction MP2 conduit à une importante surestimation des énergies de stabilisation. La méthode CCSD-T et le potentiel modèle fournissent quant à eux des résultats de bonne qualité.

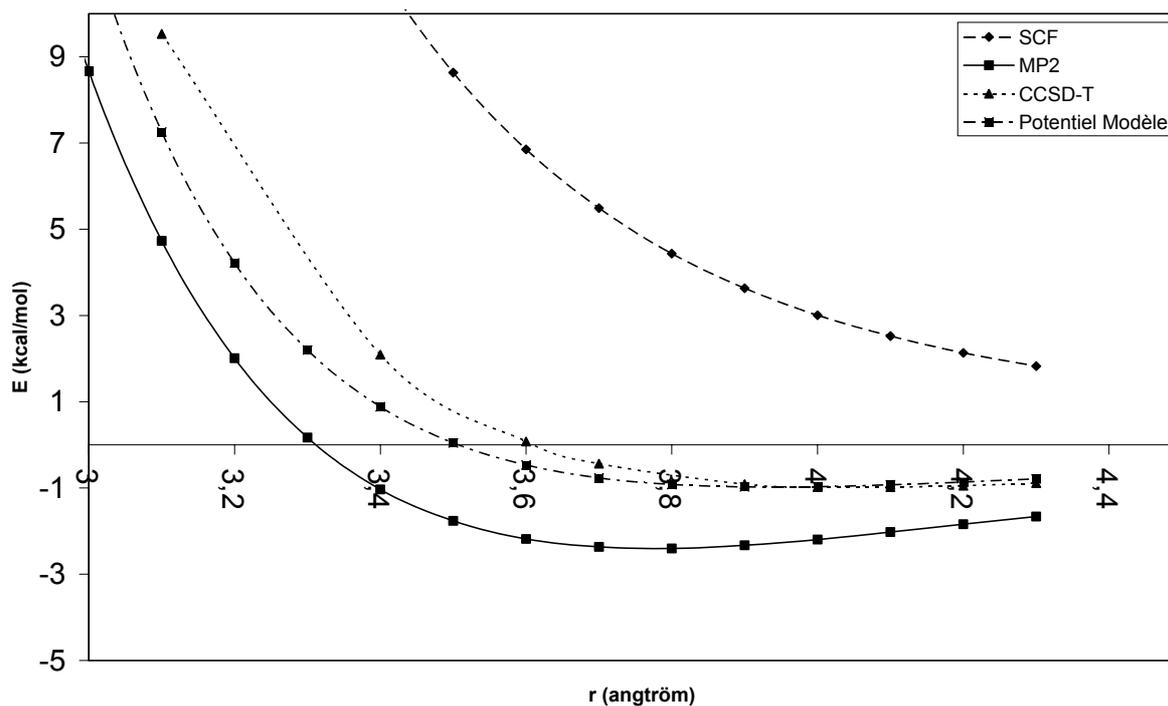
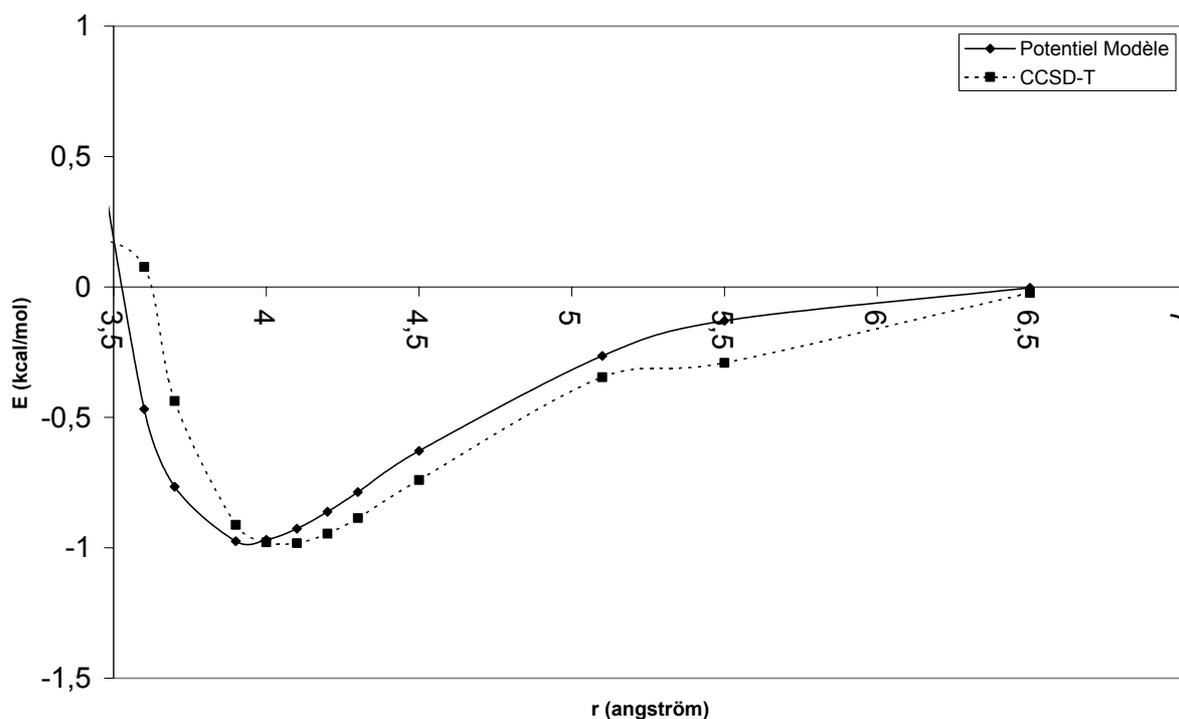


Figure 2.4

La figure 2.5 se restreint à la comparaison entre les niveaux CCSD-T et potentiels modèle, sur une plus large gamme de distances d'approche entre les deux molécules de benzène.



## 1. Description du Transfert de Charge

L'accord entre les deux méthodes est généralement bon. Il tend à se dégrader à courte distance, ce qui met en lumière trois limitations du modèle de dispersion-répulsion employé ici : d'une part, un même jeu de paramètre doit rendre compte simultanément des phénomènes de dispersion (moyennes distances d'approche) et de répulsion (courtes distances d'approche), ce qui est délicat et dépend beaucoup des configurations sur lesquelles l'ajustement desdits paramètres a eu lieu. D'autre part, la forme du 'mur répulsif' aux très courtes distances d'approche, telle qu'elle est obtenue *ab initio*, est assez mal reproduite par la partie monoexponentielle du terme de dispersion-répulsion. Enfin, l'isotropie spatiale du modèle est peu adaptée à des molécules telles que les aromatiques ; or, c'est aux courtes distances d'approche que cette anisotropie est la plus fortement ressentie.

Ceci dit, il est bien plus important de retranscrire fidèlement l'énergie d'interaction aux moyennes et longues distances intermoléculaires (au voisinage de la position d'équilibre notamment), qu'aux courtes. Or, pour de telles situations, le potentiel modèle se révèle de bonne qualité.

La paramétrisation du terme de dispersion-répulsion fournit donc un potentiel modèle précis et bien adapté aux systèmes à l'étude, malgré leur grande complexité. Le modèle de dispersion-répulsion gagnerait cependant à être étendu, notamment en prenant en compte une certaine forme d'anisotropie spatiale.

### 1.4. Calcul des éléments non diagonaux de l'hamiltonien effectif

#### *1.4.1. Valeur ab initio des éléments non diagonaux*

Il s'agit ici de déterminer, puis de calculer, l'expression des éléments non diagonaux de l'hamiltonien effectif  $\hat{H}$  défini plus haut, c'est-à-dire les termes couplant entre elles les formes diabatiques qui en forment la base. Ces formes, toutes des représentations d'un même agrégat ionisé, diffèrent par la localisation de la charge sur l'un ou l'autre des monomères constituant l'agrégat. Ceci peut être modélisé par l'arrachement d'un électron d'une orbitale moléculaire bi-occupée de l'agrégat neutre, qui serait localisée spatialement sur un des monomères.

La détermination de l'élément de couplage est explicitée dans ces lignes sur un cas simple. Soit un système neutre à couches électroniques fermées, de fonction d'onde

## Deuxième Partie – Le Transfert de Charge

monodéterminantale  $|\bar{a}\bar{b}\bar{c}\bar{c}|$ , où  $a$ ,  $b$  et  $c$  sont des orbitales moléculaires. L'ionisation de ce système peut s'effectuer de deux façons distinctes, qui ont pour conséquence l'arrachement d'un électron de l'orbitale  $a$  (soit le déterminant  $|ab\bar{c}\bar{c}|$ ) ou de l'orbitale  $b$  (c'est-à-dire le déterminant  $|\bar{a}b\bar{c}\bar{c}|$ ). Dans le contexte de ce travail, les orbitales  $a$ ,  $b$  et  $c$  seraient des orbitales moléculaires d'un agrégat neutre localisées sur des monomères distincts, de sorte que les deux configurations électroniques ci-dessus seraient celles de deux formes diabatiques à charge localisée.

L'élément de matrice correspondant au couplage entre ces deux formes s'écrit :

$$\begin{aligned} \langle \bar{a}\bar{b}\bar{c}\bar{c} | \hat{H} | ab\bar{c}\bar{c} \rangle &= (\bar{a} | \hat{h} | \bar{b}) \\ &+ (\bar{a}\bar{a} | \hat{g} | \bar{b}\bar{a}) - (\bar{a}\bar{a} | \hat{g} | a\bar{b}) \\ &+ (\bar{a}\bar{b} | \hat{g} | \bar{b}\bar{b}) - (\bar{a}\bar{b} | \hat{g} | b\bar{b}) \\ &+ (\bar{a}\bar{c} | \hat{g} | \bar{b}\bar{c}) - (\bar{a}\bar{c} | \hat{g} | c\bar{b}) \\ &+ (\bar{a}\bar{c} | \hat{g} | \bar{b}\bar{c}) - (\bar{a}\bar{c} | \hat{g} | c\bar{b}) \end{aligned}$$

où  $\hat{h}$  est la partie monoélectronique de l'hamiltonien effectif, et avec la notation  $(pq | \hat{g} | rs) = \int p^*(\vec{r}_1)r(\vec{r}_1) \frac{1}{r_{12}} q^*(\vec{r}_2)s(\vec{r}_2) d\vec{r}_1 d\vec{r}_2$ . Il convient de noter que, dans le cas où  $a$ ,  $b$  et  $c$  sont localisées sur des monomères distincts, les intégrales de type  $(ac | \hat{g} | bc)$  sont tricentriques et traduisent les effets d'environnement sur le couplage. Comme mentionné en 1.2, leur contribution est généralement faible et sera négligée dans ce travail.

L'opérateur hamiltonien peut également s'écrire, en employant les opérateurs coulombien et d'échange définis en première partie :

$$\hat{H} = \hat{h} + J_a + J_b + 2J_c - K_c$$

L'opérateur de Fock pour le système neutre s'écrit quant à lui :

$$\begin{aligned} \hat{F} &= \hat{h} \\ &+ 2J_a - K_a \\ &+ 2J_b - K_b \\ &+ 2J_c - K_c \end{aligned}$$

L'élément de matrice de Fock  $F_{ab}$  entre  $|a\rangle$  et  $|b\rangle$  a donc pour expression :

## 1. Description du Transfert de Charge

$$\begin{aligned} F_{ab} &= h_{ab} \\ &+ 2\langle aa|ba\rangle - \langle aa|ba\rangle \\ &+ 2\langle ab|bb\rangle - \langle ab|bb\rangle \\ &+ 2\langle ac|bc\rangle - \langle ac|cb\rangle \\ &= H_{ab} \end{aligned}$$

Ainsi, le couplage entre deux formes *valence bond* du système ionisé est égal à l'élément de matrice de Fock entre orbitales moléculaires du système neutre impliquées dans le processus de transfert d'électron. Cette identité n'est utile que dans la mesure où il est possible de calculer les éléments de matrice de Fock entre orbitales localisées. Or, les orbitales canoniques issues d'un traitement SCF de l'agrégat neutre n'ont aucune raison de présenter cette propriété. La localisation préalable des orbitales moléculaires de l'agrégat est donc nécessaire.

### 1.4.2. Localisation des orbitales moléculaires

La description des états électroniques d'un agrégat ionisé, dans la base de ses fonctions d'onde diabatiques à charge localisée, nécessite de disposer d'un jeu d'orbitales moléculaires spatialement localisées sur les monomères de l'agrégat, de sorte que l'arrachement d'un électron à l'une de ces orbitales suite à l'ionisation positionne la charge sur le monomère en question.

Les méthodes de localisation génèrent, à partir des orbitales moléculaires canoniques Hartee-Fock, un jeu d'orbitales moléculaires localisées, *via* un critère de localisation<sup>120,219,220</sup> propre à chacune d'entre elles. Il sera montré plus bas que ce processus, appliqué à un ensemble d'orbitales canoniques de même nombre d'occupation, ne modifie pas l'opérateur de Fock. En conséquence, les orbitales occupées et virtuelles sont généralement traitées de manière indépendante.

La méthode retenue pour cette étude est la localisation par projection, qui consiste à projeter les orbitales canoniques sur un jeu d'orbitales préalablement localisées. Elle a l'avantage de permettre un contrôle du processus de localisation, par le choix du jeu d'orbitales sur lesquelles la projection a lieu. En effet, dans un agrégat de molécules aromatiques, il faut à la fois obtenir une localisation des orbitales moléculaires sur chacun des monomères, tout en préservant la délocalisation intramoléculaire nécessaire à la retranscription de l'aromaticité. Comme dans toute méthode de localisation, une

## Deuxième Partie – Le Transfert de Charge

orthogonalisation post-projection des orbitales s'impose. Celle-ci remélange en partie les orbitales entre elles, dégradant au passage très légèrement leur localisation.

Reste à prouver que l'identité démontrée ci-dessus entre les éléments non diagonaux de l'hamiltonien effectif et les éléments de matrice de Fock reste vraie après localisation.

Soit à cet effet une fonction d'onde construite sur un unique déterminant, lui-même constitué d'un ensemble de spin-orbitales  $\{\phi_i\}$  de nombre d'occupation  $n_i$ ,  $0 < n_i < 1$ . La matrice densité à une particule s'exprime de la façon suivante :

$$\rho = \sum_i n_i \phi_i^* \phi_i$$

Si, de plus,  $n_i = n \forall i$ , c'est-à-dire si l'ensemble  $\{\phi_i\}$  est constitué d'orbitales canoniques occupées ou virtuelles :

$$\rho = n \sum_i \phi_i^* \phi_i$$

Soit  $U$  une transformation unitaire (vérifiant  $UU^\dagger = I$ ) qui projette les orbitales  $\{\phi_i\}$  sur un jeu d'orbitales  $\{\psi_i\}$  (qui, dans le contexte de ce travail, sont localisées):

$$\psi_i = \sum_j U_{ij} \phi_j$$

La matrice densité à une particule associée peut s'exprimer :

$$\begin{aligned} \rho &= \sum_i n_i \psi_i^* \psi_i \\ &= \sum_i n_i \sum_{j,k} \phi_j^* U_{ji}^* U_{ki} \phi_k \\ &= \sum_{j,k} \phi_j^* \phi_k \sum_i n_i U_{ji}^* U_{ki} \\ &= \sum_{j,k} \phi_j^* \phi_k \sum_i n_i U_{ki} U_{ij}^{-1} \end{aligned}$$

puisque  $U_{ji}^* = U_{ij}^\dagger = U_{ij}^{-1}$ . Si, de plus, la condition  $n_i = n \forall i$  est vérifiée :

$$\begin{aligned} \rho &= n \sum_{j,k} \phi_j^* \phi_k \underbrace{\sum_i U_{ki} U_{ij}^{-1}}_{\delta_{jk}} \\ &= n \sum_j \phi_j^* \phi_j \end{aligned}$$

Soit la même expression que pour la matrice densité associée aux orbitales  $\{\phi_i\}$ . Le processus de localisation n'invalide donc pas le raisonnement tenu en 1.4.1, à la condition que la projection se restreigne à un ensemble d'orbitales canoniques de même degré d'occupation (toutes occupées ou toutes virtuelles).

## 1. Description du Transfert de Charge

### 1.4.3. Corrélation et polarisation des orbitales

Le traitement ci-dessus laisse de côté des effets pourtant importants. En premier lieu, la description des fonctions d'onde des formes diabatique à charge localisée  $A^+B$  et  $AB^+$  dans la base des orbitales moléculaires du système neutre  $AB$  revient à négliger la polarisation de ces orbitales sous l'effet de la charge. Il s'agit, de surcroît, d'orbitales SCF qui ne prennent pas en compte la corrélation électronique.

Les effets de polarisation peuvent être inclus en étendant la description des formes *valence bond* à plusieurs déterminants, à l'aide d'une interaction de configurations monoexcitées sélectionnées. Soit par exemple la forme diabatique  $A^+B$ . Les déterminants pour lesquels l'excitation a lieu sur le monomère chargé traduisent la repolarisation des orbitales neutres de celui-ci, sous l'effet de sa propre charge. Parmi eux, ceux pour lesquels l'excitation comble la lacune de l'orbitale ionisée par un électron issu d'une orbitale bioccupée retranscrivent la délocalisation du 'trou' dans l'espace des orbitales occupées, et permettent de vérifier que l'état ionique considéré est correctement décrit, à l'ordre zéro, par un unique déterminant. La polarisation des orbitales du monomère neutre  $B$  par la charge en  $A$  est quant à elle retranscrite par l'ensemble des monoexcitations occupées  $\rightarrow$  virtuelles sur  $B$ . La prise en considération de la corrélation électronique par rapport à l'état fondamental de l'ion nécessiterait l'inclusion de toutes les diexcitations par rapport à cet état ; pour la présente étude, les effets de polarisation ont cependant été jugés plus importants que ceux issus de corrélation<sup>113</sup>.

L'élément de couplage effectif  $F_{ab}^{\text{eff}}$  issu de ce traitement d'interaction de configurations sera donc préféré à l'élément de matrice de Fock  $F_{ab}$  correspondant, pour l'évaluation du couplage entre formes *valence bond*.

### 1.4.4. Simplification du modèle

Si la connaissance des éléments de matrice de Fock entre orbitales moléculaires donne accès aux éléments de couplage entre formes diabatiques à charge localisée, il n'en reste pas moins que la construction d'un opérateur de Fock est fastidieuse, voire inenvisageable si elle doit être exécutée pour chaque configuration d'un calcul Monte Carlo ou à chaque pas de temps d'une simulation de dynamique moléculaire. L'inclusion des effets

de polarisation par interaction de configurations monoexcitées sélectionnées accroît encore le coût du processus. En pratique, la disponibilité d'une méthode simplifiée pour le calcul des éléments de couplage s'impose. Il a donc été choisi de mettre à profit la relation de proportionnalité bien connue entre les éléments de matrice non diagonaux de l'hamiltonien effectif et une intégrale de recouvrement entre orbitales moléculaires, fréquemment employée dans les méthodes semi-empiriques<sup>120</sup>. Les orbitales moléculaires en question sont les orbitales canoniques des monomères qui entrent en jeu dans le transfert d'électron, translatées et pivotées en fonction de la position de ceux-ci au sein du complexe ; par définition, ces orbitales sont spatialement localisées sur les monomères et ne sont pas orthogonales entre elles (autrement dit, leur recouvrement n'est pas nul). Avec les notations du paragraphe 1.4.1, l'élément de matrice non diagonal de l'hamiltonien effectif s'écrit :

$$\langle a\bar{a}b\bar{c}c\bar{c} | \hat{H} | ab\bar{b}c\bar{c} \rangle = K \langle \varphi_p | \varphi_q \rangle$$

La physique contenue dans le coefficient de proportionnalité  $K$ , les paramètres dont il dépend, et la procédure à l'aide de laquelle il est déterminé, seront détaillés par la suite. Il convient cependant de noter que l'intégrale de recouvrement concerne les orbitales  $\varphi_p$  et  $\varphi_q$  des monomères et non les orbitales correspondantes  $a$  et  $b$  du supersystème, orthogonales entre elles.

#### 1.4.5. Calcul des intégrales de recouvrement et de leurs dérivées

Le modèle simplifié adopté dans cette étude nécessite le calcul de l'intégrale de recouvrement entre orbitales moléculaires données. Dans l'optique d'applications telles que l'exploration locale de surfaces d'énergie potentielle ou la dynamique moléculaire, il faut pouvoir évaluer ces intégrales mais également leurs dérivées partielles par rapport à tous les degrés de liberté du système.

L'intégrale de recouvrement entre deux orbitales moléculaires  $\phi_1$  et  $\phi_2$  quelconques s'écrit :

$$\langle \phi_1 | \phi_2 \rangle = \sum_{A,B} c_{A1} c_{B2} \langle \chi(A) | \chi(B) \rangle$$

où  $\chi(A)$  et  $\chi(B)$  sont deux orbitales atomiques respectivement centrées aux points  $A$ , de coordonnées  $(A_x, A_y, A_z)$  et  $B$ , et  $c_{A1}$  et  $c_{B2}$  les coefficients correspondants du développement LCAO. Ces orbitales sont elles-mêmes des combinaisons de fonctions gaussiennes cartésiennes  $g$ :

## 1. Description du Transfert de Charge

$$|\chi(\mathbf{A})\rangle = \sum_i d_{iA} |g(\mathbf{A}, \alpha_i)\rangle$$

avec :

$$g(\mathbf{A}, \alpha_i, \vec{n}, \vec{r}) = (x - A_x)^{n_x} (y - A_y)^{n_y} (z - A_z)^{n_z} e^{-\alpha_i [(x-A_x)^2 + (y-A_y)^2 + (z-A_z)^2]}$$

L'intégrale de recouvrement de deux orbitales moléculaires est donc la combinaison d'intégrales de recouvrement de gaussiennes, qui se factorisent de la sorte :

$$\langle g(\mathbf{A}) | g(\mathbf{B}) \rangle = \langle g(A_x) | g(B_x) \rangle \langle g(A_y) | g(B_y) \rangle \langle g(A_z) | g(B_z) \rangle$$

Le calcul de la dérivée d'un recouvrement entre orbitales moléculaires, par rapport à  $A_x$  par exemple, se ramène donc au calcul de la combinaison de recouvrements de gaussiennes :

$$\frac{\partial}{\partial A_x} (\langle g(\mathbf{A}) | g(\mathbf{B}) \rangle) = \left\langle \frac{\partial g(A_x)}{\partial A_x} \middle| g(B_x) \right\rangle \langle g(A_y) | g(B_y) \rangle \langle g(A_z) | g(B_z) \rangle$$

Reste donc à calculer la dérivée :

$$\left\langle \frac{\partial g(A_x)}{\partial A_x} \middle| g(B_x) \right\rangle = \int \frac{\partial}{\partial A_x} \left[ (x - A_x)^{n_x} e^{-\alpha_i (x - A_x)^2} \right] (x - B_x)^{n_x} e^{-\alpha_i (x - B_x)^2} dx$$

avec :

$$\frac{\partial}{\partial A_x} \left[ (x - A_x)^{n_x} e^{-\alpha_i (x - A_x)^2} \right] = \left[ n_x (x - A_x)^{n_x - 1} - 2(x - A_x)^{n_x + 1} \right] e^{-\alpha_i (x - A_x)^2}$$

Le calcul de la dérivée du recouvrement de deux gaussiennes se ramène donc au calcul du recouvrement d'autres gaussiennes, dont les nombres quantiques  $n_x$ ,  $n_y$  ou  $n_z$  ont changé d'une unité. Il est classiquement déterminé à l'aide de la méthode des polynômes orthogonaux de Gauss-Hermite, implémentée par exemple dans le logiciel Hondo95<sup>221</sup>.



## 2. Phénomènes de Résonance de Charge dans les Agrégats Aromatiques Ionisés

### 2.1. Introduction

**L**a description du transfert de charge intermoléculaire dans un agrégat chargé à l'aide par la combinaison d'une méthode à hamiltonien effectif et de potentiels modèle adaptés, a été appliquée à l'étude structurale de petits agrégats, neutres et chargés, de benzène, de naphthalène et d'anthracène, donnant lieu à la publication d'un article retranscrit ci-dessous.

L'ajustement du modèle, par le biais de la constante de proportionnalité reliant couplages électroniques et recouvrements entre orbitales moléculaires, y est discuté en détail. La nature des calculs *ab initio* utilisés comme référence lors de ce processus de paramétrisation y est elle aussi présentée. Enfin, les limites du modèle sont mis en lumière dans des systèmes chimiques particuliers.

Les propriétés structurales et électroniques des agrégats aromatiques neutres et chargés sont déterminées puis comparées, entre elles et à leurs contreparties expérimentales. Il est montré que la compétition entre la résonance de la charge et la polarisation des monomères neutres par cette charge, qui évolue en fonction de la nature et de la taille des agrégats, régit nombre de ces propriétés. Une première approche théorique des mécanismes d'ionisation verticaux ou adiabatiques dans ces systèmes est également proposée.

### 2.2. Article

# A Model Potential Approach to Charge Resonance Phenomena in Aromatic Cluster Ions

Benjamin Bouvier,\* Valérie Brenner, Philippe Millié, and Jean-Maïk Soudan

CEA Saclay, DSM/DRECAM/SPAM, Laboratoire François Perrin, URA 2453, Bât. 522,  
91191 Gif-sur-Yvette, France

Received: March 20, 2002; In Final Form: July 30, 2002

This work proposes a simple yet accurate methodology to account for charge resonance in ionic clusters. The supersystem's model Hamiltonian is described via a basis set of valence bond structures for which the charge is localized on a given monomer, and whose intermolecular binding energies are computed using a polarizable model potential. The coupling elements between these structures are proportional to an overlap integral between relevant nonorthogonal monomer molecular orbitals. Ab initio calculations are employed to calibrate and validate the model, but also to define its limits. The methodology is then applied to the global exploration of potential energy surfaces for small homocluster ions of benzene, naphthalene, and anthracene. The structural and electronic properties of these systems are discussed, with emphasis on important trends such as the polarization vs charge-transfer competition or the difference between adiabatic and vertical ionization potentials. Extensions to stacked cluster ions of higher aggregation number ( $n = 15$ ) conclude this work.

## I. Introduction

Molecular clusters of aromatic compounds have been the subject of ongoing experimental and theoretical studies for a long while. They provide valuable insight for the comprehension of  $\pi$  interactions,<sup>1</sup> which play a predominant role in phenomena such as stacking or  $\pi$  hydrogen bonding.<sup>2,3</sup> Since aromatic moieties are present in many large organic molecules and biopolymers and often play an important part in noncovalent interactions, this knowledge can be fruitfully extended to the comprehension of local or group-wise phenomena. Although less effort has been invested in the study of charged aromatic clusters than in their neutral counterparts, it is now well-established that, for small aggregation numbers, their structural and electronic properties are very different from those of the neutral aggregates. This is evidenced by the important gap between the adiabatic and vertical ionization energies, and can be explained by the dominating influence of charge-transfer interactions. Hence, small aromatic clusters are ideally suited as model systems for the study of charge resonance phenomena.

However, the apparent chemical simplicity of small aromatic clusters conceals the theoretical difficulties which ab initio calculations run into when applied to these systems. First of all, typical near-neighbor interaction energies in neutral clusters amount to less than 2 kcal/mol, that is, require very accurate (and costly) ab initio calculations. Additionally, this interaction originates from subtle tradeoffs between dispersion, repulsion, and electrostatic energy contributions, whose relative importance may vary drastically depending on the cluster structure. Particularly delicate among these interactions are the dispersion effects, based on electron correlation and whose convergence with regard to basis set size is quite slow.<sup>4,5</sup>

The most striking example of these difficulties occurs in the case of the neutral benzene dimer, which has undergone extensive theoretical study. The Hartree–Fock (HF) level of

calculation yields a repulsive interaction energy.<sup>6</sup> At the Møller–Plesset second-order perturbation theory level (MP2), the attractive nature of the interaction is correctly rendered, but systematically overestimated.<sup>7,8</sup> It is common knowledge that density functional (DFT) calculations often provide unreliable dispersion energies; hence, applying DFT advisably to these systems requires an additional dispersion potential.<sup>9</sup> On the whole, only expensive coupled-cluster calculations including single, double and perturbative treatment of triple excitations from the ground state (CCSD-T) can be said to yield accurate results.<sup>10</sup>

A different type of problem is encountered when studying small charged aromatic clusters. Typical pairwise interaction energies are 1 order of magnitude greater than for neutral aggregates, ensuring much smaller relative errors. The nature of the interaction is also significantly different, consisting, to a large extent, of polarization and charge resonance terms. However, the description of charge resonance phenomena by molecular orbital-based theories (MO) is biased by the method itself: a homodimer bearing a global +1 charge will favor symmetric solutions at short range and dissymmetric ones at longer distances, the occurrence of the symmetry-breaking artifact depending on the system and the method; discontinuity and instability can take place at the border separating the two zones.<sup>11,12</sup> Clearly, monodeterminantal approaches such as HF or HF-based methods are frequently plagued by this problem, but virtually all methods involving a self-consistent field procedure, whether mono- or multireferential, can also be affected. As a rule of thumb, the artifact has to be considered as soon as the correct description of the system requires more than one Lewis structure. For such a case, the valence bond (VB) theory sounds very appealing indeed, but generalized valence bond approaches rapidly become cumbersome as the system grows larger. Finally, let us mention that DFT, although not strictly speaking a molecular orbital method, does not offer a better description of charge resonance interactions.<sup>13</sup> On the whole, the process of selecting a reliable ab initio method for a given system featuring charge-transfer cannot be systematized

\* Corresponding author. Phone: +33-1-6908-3788. Fax: +33-1-6908-8707. E-mail: bouvier@pandora.saclay.cea.fr.

and, more often than not, resorts to choosing elaborate methods that are unsuitable for large systems.

The aforementioned problems are by no means trivial; in this context, model potential methods are very attractive. For neutral aromatic clusters, their absolute error margins are comparable to those of *ab initio* calculations, at a tremendously lower computational cost. Combined with a VB method, they should provide a physically coherent approach to the study of charge resonance phenomena in charged aggregates. Possible applications of these relatively inexpensive methods vary from the global exploration of potential energy surfaces to molecular dynamics calculations.

Model potentials calculations have already been carried out on small neutral aromatic clusters,<sup>14–18</sup> but to our knowledge, no such potential has been able to tackle the case of ionized clusters. Furthermore, although methodologies ensuring description of charge-transfer phenomena have already been utilized in other systems (most relevant among which that of Jortner et al in DNA strands<sup>19–21</sup>), they are currently restricted to static geometries.

The present work proposes to address the question using a VB-based charge resonance-inclusive model potential supporting potential energy surface (PES) exploration techniques. In a first part, the methodology used to construct and calibrate the potential is explained. Then, the *ab initio* method selected as the calibration reference is discussed. Once correctly parametrized, the potential is applied to the thorough exploration of the PES for small charged and neutral homoclusters of benzene, naphthalene and anthracene, with aggregation numbers up to  $n = 4$ . Structural properties of the charged and neutral clusters are compared, with a focus on the implications of the charge resonance vs polarization competition, as well as on the adiabatic and vertical ionization potentials. The spatial delocalization of the charge in ionized clusters is also investigated. Finally, columnar aggregates of these compounds are investigated up to  $n = 15$ , providing valuable insight into charge delocalization processes in larger systems.

## II. Methodology

### II.1. Effective Hamiltonian and Valence Bond Framework.

The energy and wave function of an  $n$ -molecular system bearing a global +1 charge are defined as the eigenvalues and eigenfunctions of an effective Hamiltonian matrix, built using a straightforward VB formalism; this procedure has already been applied to the study of charge-transfer effects in small water-ion clusters.<sup>22</sup> We shall discuss the formalism in the simplest case where the charged cluster is denoted  $(M_1M_2\dots M_n)^+$ , and each individual molecule  $M_i$  is characterized by a neutral form  $M_i$  and an ionic form  $M_i^+$ . The functions  $\{(M_1M_2\dots M_i^+\dots M_n)\}$ ,  $i \in \{1..n\}$ , form a basis set which underlies the effective Hamiltonian matrix, i.e., the wave function of the globally charged aggregate is written out as a linear combination of states where the charge is localized on a given molecule, all other entities being in their respective neutral ground states. The diagonal elements of the effective Hamiltonian matrix represent the energies of the aforementioned states. The off-diagonal elements coupling two of these mesomer forms are the direct translation of the charge-transfer phenomenon, and can be viewed as an extension to molecular systems of the “hopping integral” formalism in wide use among solid state physicists. They act as a link between the electronic structure of the monomers, which is known, and that of the cluster, which is not. This effective Hamiltonian approach has the advantage of modeling charge transfer as a many-body term, which is

effectively the case in near-resonant cases for which two-body or perturbative treatments fall short of the mark.<sup>23</sup> Additionally, when evaluating diagonal elements, full knowledge of the state (neutral or charged) of each molecule is accessible, so that it should be possible to treat these two states at a different calculation level altogether (i.e., by maintaining different model potential parameter sets for each species).

**II.2. Energy of the Valence Bond Structures.** The energy of the VB structures, i.e., the diagonal elements of the effective Hamiltonian matrix, are evaluated using an enhanced<sup>24–26</sup> version of the intermolecular polarizable model potential of Claverie and Vigné-Maeder.<sup>27,28</sup> An extensive description of the potential being beyond the scope of this text, only a brief outline is given hereafter.

The potential assumes intramolecular geometries to be frozen and describes each molecule by six degrees of freedom (three translations and three Euler angles). On the basis of the exchange perturbation theory, it contains terms corresponding to the electrostatic, dipole polarization, dispersion, repulsion, and exchange dispersion energy contributions.

The electrostatic energy originates from the interaction between multipolar (up to the quadrupole) multicenter (atoms and bond centers) distributions on each molecule, which are derived from an *ab initio* one-electron density matrix calculation at a requisite level, followed by a site reduction procedure.

The dipole polarization energy results from a self-consistent iterative procedure during which additive site polarizabilities on each molecule respond to the electric field created by the (permanent and induced) moments on all other molecules. Thus, the polarization energy is not merely treated as a many-body term, but the influence of back-polarization is also taken into account. In the original work of Claverie, site polarizabilities are calculated using an additive bond-polarizability model developed by Lefevre,<sup>29</sup> in which each bond is affected a parallel and a perpendicular component depending on its type, with the notable difference that those two parameters are combined in a mean isotropic site polarizability. This is clearly unsuitable for aromatic moieties whose polarizability anisotropy is high (ca. 53% for benzene<sup>30</sup>). Restoring the full site polarizability tensor using Lefevre’s parameters does not suffice, since a single perpendicular bond component is insufficient to account for the out-of-plane polarizability of aromatics. Separation of the  $\pi$  and  $\sigma$  subsystems (each with its own set of three bond parameters: one parallel, one in-plane perpendicular, and one out-of-plane perpendicular) is a baseline requirement for such systems, as has been explained by Mulder.<sup>31</sup> An enhanced version of Mulder’s parametrization, designed to yield realistic site polarizabilities (the original scheme only aims at reproducing molecular tensors and anisotropies) was used for this study; it shall be the object of a forthcoming communication.<sup>32</sup> Polarization energy, albeit small in neutral aromatic clusters, acts as a counterbalance against delocalization in charged clusters, as shall be seen hereafter; correct description of this contribution is therefore essential.

The dispersion, repulsion, and exchange dispersion contributions are approximated as a sum of atom-atom terms using a Kitaigorodski<sup>33</sup> formalism, in which individual atom types are defined by two parameters. Three different distance zones, using distinct analytical formulas, are governed by these parameters. This ensures a realistic treatment of short-range terms (i.e., penetration effects), while still providing continuity and derivability at the border between the zones. However, the implied interdependence between the domains is a drawback when adjusting the potential parameters, as shall be seen shortly.

**II.3. Implementing Charge Transfer. II.3.1. General Framework.** As mentioned earlier, the off-diagonal matrix elements coupling two distinct VB structures are responsible for charge resonance inside the cluster. Assuming each of these VB structures to be fairly described by a zeroth-order wave function  $|\Psi_i\rangle$ ,  $i \in \{1..n\}$ , built on the same set of mutually orthogonal supersystem molecular orbitals that are localized on each of the monomers (vide infra), the coupling element between VB structures  $i$  and  $j$ ,  $i \neq j$ , may in all cases be expressed as a Fock operator matrix element defined by

$$\langle \Psi_i | H | \Psi_j \rangle = F_{kl} = \langle \varphi_k | F | \varphi_l \rangle \quad (1)$$

where  $\varphi_k$  and  $\varphi_l$  refer to the two localized molecular orbitals involved in the electron transfer, i.e.,  $|\Psi_j\rangle = a^+_{ka} |\Psi_i\rangle$ . However, the prohibitive cost of constructing a Fock operator for each configuration or time step in Monte Carlo or molecular dynamics simulations calls for a simpler way of evaluating the off-diagonal elements of the effective Hamiltonian. This is most easily achieved by utilizing the popular linear relationship between these elements and an overlap integral between a suitably chosen pair of (nonorthogonal) neutral monomer molecular orbitals  $\chi_p$  and  $\chi_q$ :

$$\langle \Psi_i | H | \Psi_j \rangle = K_{pq} \langle \chi_p | \chi_q \rangle \quad (2)$$

The proportionality factor  $K_{pq}$  depends a priori on the system's nature, its geometry, and the molecular orbitals involved in the overlap. However, inside sets of cluster geometries, the ratio of the coupling element over the overlap integral (eq 2) is nearly constant, so that the aforementioned geometry dependence may often be safely ignored. This additional level of approximation will be detailed further on.

It is a reasonable assumption to suppose that the coupling element  $H_{ij}$  depends on the nature of the monomers involved in the charge-transfer mechanism (i.e., the monomers bearing the charge in valence bond structures  $\Psi_i$  and  $\Psi_j$ ) but not on the other molecules in the aggregate. The coupling element is thus defined as a two-body term, while many-body effects are included in the diagonal elements of the Hamiltonian (the intermolecular interaction energies of the corresponding valence bond structures). Hence, for a given pair of monomers, the parameter  $K$  can be inferred from ab initio calculations on the dimer, and used as-is in model potential calculations on larger clusters featuring the same monomer pair, regardless of the nature of the other molecules.

Finally, molecular orbital dependence is not much of an issue if a single ionization involving a single MO (generally the highest occupied one) is considered per monomer unit. Extension to more complex cases will also be discussed hereafter.

We attempt to characterize charge-transfer interactions in a given system using the aforementioned proportionality factor  $K$  as the sole parameter. Its determination from ab initio calculations is explained in the following paragraph.

### II.3.2. Ab Initio Parameter Determination and Validation.

As seen earlier, correct inclusion of charge-transfer effects in an ab initio calculation is by no means a simple task. Furthermore, using such a calculation as a calibration reference for the model potential requires some degree of equivalence between the two levels, which should include or neglect the same effects. The procedure is detailed in the simple case of two monomers for which the ionization occurs on a single MO, and is extended to species where more than one MO has to be explicitly considered. Let us stress the fact that the procedure provides a value of the parameter  $K$  for a given cluster geometry

only, and needs to be iterated over a set of cluster structures before a mean value for  $K$  can be found.

A Hartree–Fock calculation is carried out for both neutral monomers (monomer geometries inside the complex are considered equivalent to those of the isolated molecules). The molecular orbital involved in the ionization is identified. Another Hartree–Fock calculation is then applied to the neutral dimer. The combined set of occupied monomer orbitals is projected onto the occupied set of dimer molecular orbitals, resulting, after orthonormalization, in a set that is localized on each of the molecules but preserves intramolecular delocalization. This is a requisite to transcription of the aromaticity of the molecules, which stems from  $\pi$  electron delocalization: it justifies our choice of the aforementioned projection method compared to other popular localization schemes.<sup>34</sup> The same procedure yields a set of localized virtual orbitals from monomer and dimer virtual MOs.

Description of a VB limit-structure using a single determinant constructed on these orbitals leaves out the major part of polarization effects. We chose to render these effects using a selected CI approach. The polarization of the neutral monomer by the one bearing the charge is modeled by including all single excitations from the former's occupied orbitals to its virtual orbitals. Inclusion of all possible single excitations on the charged monomer (doubly occupied  $\rightarrow$  virtual, doubly occupied  $\rightarrow$  singly occupied, and singly occupied  $\rightarrow$  virtual) transcribes the repolarization of the monomer's neutral orbitals under the effect of the charge. Charge-transfer interactions, which are taken care of in the off-diagonal terms, must not contaminate the diagonal elements; this is ensured by omitting any excitation involving orbitals located on different monomers.

Two such selected CIs are conducted, their respective first roots  $E_1$  and  $E_2$  yielding the energies of the two VB structures. Off-diagonal elements are written out as  $KS_{12}$ , where  $S_{12}$  is the overlap between the monomer MOs involved in the charge resonance:

$$H_{\text{eff}} = \begin{pmatrix} E_1 & KS_{12} \\ KS_{12} & E_2 \end{pmatrix} \quad (3)$$

A third CI, whose active space is obtained by concatenation of those of the two previous ones, provides the global system energy taking charge transfer into account. This calculation is used as a reference to calibrate the effective Hamiltonian formalism (eq 3). It is done by adjusting the parameter  $K$  so that the first root  $\lambda_1$  of the global CI matches the lowest eigenvalue of the effective Hamiltonian; when this is verified, the second CI root  $\lambda_2$  and the second eigenvalue of  $H_{\text{eff}}$  are generally very similar as well. This ensures that both the ground and the first excited state of the dimer ion (respectively associated with the first and second eigenvectors of the effective Hamiltonian) are correctly described using the current value of  $K$ . In the case of two strictly equivalent monomers ( $E_1 = E_2$ ), the condition is fulfilled if  $KS_{12} = 1/2(\lambda_1 - \lambda_2)$ , the first two CIs being rendered superfluous.

This procedure is iterated over a set of physically relevant geometries of the dimer cluster. For each of these, the coupling element is plotted versus the corresponding value of the overlap integral. A subsequent linear regression yields a mean value for the parameter  $K$  over the set. The choice of the dimer geometries forming the set is detailed in part III, on specific examples.

The procedure allows realistic treatment of charge-transfer and polarization effects, and is as comparable as can be with the model potential level of calculation that is used for the

diagonal energies. It should provide reliable estimates of the mean parameter  $K$ .

**II.3.3. Extension: Multiple Ions.** The aforementioned procedure may be extended to cases where the inclusion of more than one possible VB structure per charged monomer is necessary. This occurs for species that have several close-lying ionization potentials (IP), such as pyridine or toluene. The case of benzene, whose lowest ionized state involves two molecular orbitals, can be treated in the same manner by formally considering two degenerate states involving one singly occupied molecular orbital each. The effective Hamiltonian for a homodimer cluster with two IPs per monomer now takes the form

$$H_{\text{eff}} = \begin{pmatrix} E_1^1 & 0 & V_{11} & V_{12} \\ & E_1^2 & V_{21} & V_{22} \\ & & E_2^1 & 0 \\ & & & E_2^2 \end{pmatrix} \quad (4)$$

Successive ions on the same monomer are considered decoupled, while four off-diagonal elements  $V_{ij}$  bind ion  $i$  on the first monomer to ion  $j$  on the second. These are written in the form  $V_{ij} = KS_{ij}$ , where  $K$  is a mean parameter averaged on all possible  $(i,j)$  pairs.

Five CI calculations are now required. The global CI contains singly excited configurations built on four references (two different ions on two monomers). The diagonal element  $E_i^a$  is the first root of a CI whose reference is ion  $a$  on molecule  $i$ . Since some active space overlap can occur between CIs for  $E_i^a$  and  $E_i^b$ ,  $a \neq b$ , some configurations have to be deleted from one of these calculations. As a rule of thumb, the concatenation of the four active spaces of the diagonal element CIs should coincide with that of the global CI.

Diagonalization of the effective Hamiltonian (eq 4) yields four solutions, the lowest of which is made to match the first root of the global CI by optimizing the value of  $K$ . As before, this results in a close agreement between the next three roots of the CI and the three remaining eigenvalues of the effective Hamiltonian, with the consequence that the ground state as well as the first three excited states of the ion cluster are correctly rendered.

The extension of the model to multiple ions per monomer implies an additional level of approximation compared to previously described cases. As mentioned earlier, the factor  $K$  is supposed to be specific to the pair of valence bond structures which it couples. In the case of single-ion monomers, there are as many valence bond structures as there are molecules, so that it is equivalent to say that  $K$  is specific to a pair of monomers. When multiple-ion monomers are considered, this is no longer true since more than one valence bond structure may correspond to the localization of the charge on a given monomer. For the sake of simplicity we chose to retain the monomer-pair dependence, at the cost of averaging  $K$  as mentioned above. Successes and failures of this model, applied to specific examples, are discussed in part III.

**II.4. Potential Energy Surface Exploration.** Our approach to charge transfer interactions features a sufficiently good tradeoff between precision and computational cost so as to allow real-time applications to systems of nonstatic geometries. This work will focus on global potential energy surface exploration.

The model potential energy surface for a given system is explored using a combination of three different methods: Monte Carlo growth (MCGM),<sup>35–37</sup> simulated annealing (SA)<sup>38,39</sup> and quasi-Newton optimization (QNO).<sup>40</sup> MCGM constructs a

cluster monomer by monomer and generates a Boltzmann sample of configurations for each cluster size at a fictitious temperature  $T$ . Each monomer is contained inside a bounding sphere, which is tested for intersections against that of the newly added molecule, at each size step. Hence, the method reaches full efficiency for aggregates of molecules that are either small or grossly spherical.

Starting from random configurations, SA samples local attraction basins on the potential energy surface (PES) by using a Metropolis<sup>41</sup> algorithm associated with a controlled “temperature” parameter, which is decreased according to an annealing schedule. This algorithm gives a good global picture of the PES for all kinds of molecules, but the exponential rise of the number of local minima on complex or high-dimensional surfaces may render subsequent optimizations unfeasible.

QNO optimizes a given configuration by moving across the surface in a (downhill) direction given by an approximation of the inverse Hessian matrix, until a minimum is reached. This requires knowledge of the energy gradient at each step. The energy is the lowest eigenvalue of the  $H_{\text{eff}}$  matrix, i.e., the lowest element of diagonal matrix  $E$  defined by

$$E = {}^t A H_{\text{eff}} A \quad (5)$$

where  $A$  is the eigenvector matrix obtained by diagonalizing  $H_{\text{eff}}$ .  $H_{\text{eff}}$ , being a real symmetric array, gives a straightforward demonstration<sup>42</sup> that the derivative of  $E$  relative to variable  $X$  is

$$\frac{\partial E}{\partial X} = {}^t A \frac{\partial H_{\text{eff}}}{\partial X} A \quad (6)$$

The analytical derivation of the diagonal element of  $H_{\text{eff}}$  will not be discussed here. Derivation of the off-diagonal element  $KS$  requires the derivative of the overlap integral  $S$ ; it can be shown that it is itself an overlap integral between Cartesian Gaussians whose quantum numbers  $n_x$ ,  $n_y$ , and  $n_z$  differ from one unit.

Finally, the QNO subroutine numerically evaluates the Hessian after convergence, to ascertain the nature of the optimum (minimum or saddle point). Since QNO is a local method, it was used as a second step to global techniques such as MCGM and SA.

**II.5. Computational Details.** Monomer geometries were optimized at the B3LYP/cc-pVDZ<sup>43</sup> level using Jaguar 4.1<sup>44</sup> and refined at the MP2/6-311G(2d,2p)<sup>45</sup> level using Gaussian98.<sup>46</sup> Neutral benzene dimer calculations were performed at the counterpoise-corrected<sup>47</sup> MP2 and CCSD-T levels using the 6-311G(2d,2p) basis set in Gaussian98. Electrostatic multipole distributions were obtained by a one-electron MP2 density calculation in this basis set. Selected CIs were carried out using HONDO95<sup>48</sup> and CIPSI,<sup>49,50</sup> with a 6-31G\*<sup>51</sup> basis set. CASPT2 calculations on pyridine systems were done using Molcas 5.0<sup>52</sup> and basis set aug-cc-pVDZ.<sup>43</sup> Model potential calculations as well as PES explorations were performed using software designed in the laboratory. All calculations were run on a Compaq ES40 AlphaServer and two Silicon Graphics O200 workstations.

### III. Ab Initio Modeling of Charge Transfer

The model, which allows adequate values of the parameter  $K$  to be inferred from ab initio selected CI calculations, was applied to several homodimer clusters. Ethylene and naphthalene were described using the single-IP formalism, the former as a

**TABLE 1: Typical Center-of-mass Distances, Coupling Elements, and Mean  $K$  Values for Sets of Relative Geometries of an  $(\text{Ethylene}_2)^+$  Cluster<sup>a</sup>**

symmetry operation	initial disposition	typical distances (Å)	typical coupling element (au)	$K$ (au)
$C_2(z)$ translation	stacked	2.5–7.0	$7 \times 10^{-2}$ – $7 \times 10^{-6}$	0.308
$C_2(z)$ rotation	stacked	3.3	$10^{-2}$	0.351
	stacked	5.0	$10^{-3}$	0.483
$C_2(z)$ rotation	coplanar	5.0	$10^{-3}$ – $10^{-4}$	0.538
$C_2(x)$ rotation	stacked	5.0	$10^{-4}$ – $10^{-5}$	0.676

<sup>a</sup> The sets are generated by applying a simple symmetry operation (a stepwise translation along or rotation about an axis) to one of the monomers of an initial cluster, while the other molecule is kept fixed. The initial cluster geometry is either stacked (the second monomer is deduced from the first by a translation along the  $C_2(z)$  axis) or coplanar (the second monomer is deduced from the first by a translation along the  $C_2(y)$  axis); both are of  $D_{2h}$  symmetry. Axes are named according to the symmetry elements of point group  $D_{2h}$ .

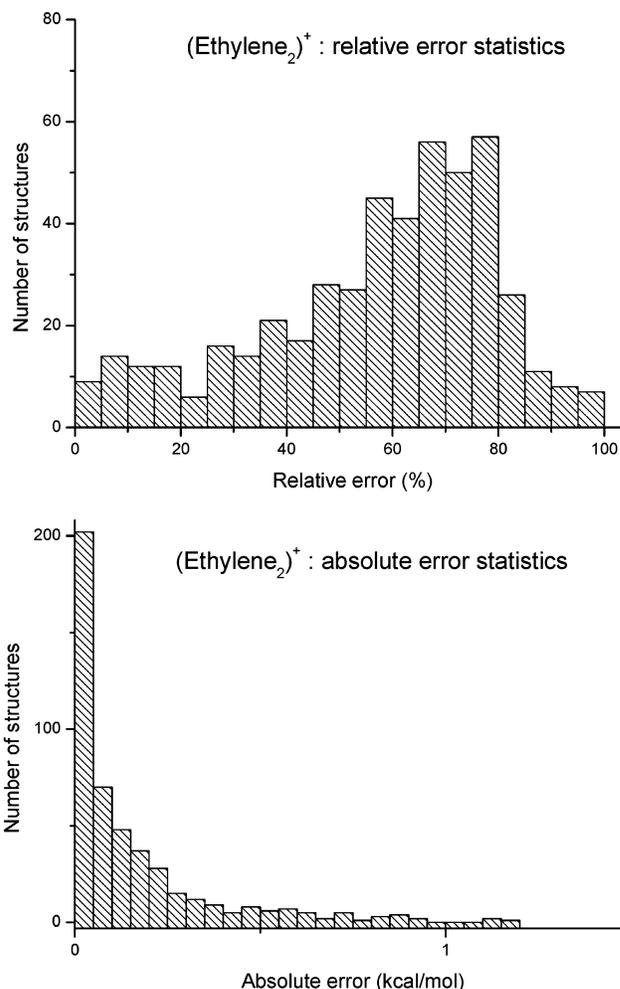
case study to validate the model, the latter as an extension to a larger species. For benzene and toluene, the multiple-IP scheme was employed and tested. Finally, the study of pyridine yielded valuable insight into the limits of the mean  $K$  approximation.

**III.1. A Case Study: Ethylene.** Ethylene is the simplest molecule containing a  $\pi$  subsystem; hence, it is well suited as a case study of the previously discussed method. The first ionized form of ethylene is a  $^2\Pi$  state. Coupling terms for charge transfer in an ethylene dimer will involve overlap between instances of these  $\pi$ -type monomer MOs.

From simple initial geometries of an  $(\text{ethylene})_2^+$  cluster (Table 1), sets of clusters were generated by allowing one molecule to undergo simple transformations (stepwise rotations along or translations about an axis), the other monomer being kept fixed. Inside each set, selected CI calculations were performed for each of the member clusters, and the resulting coupling element was plotted against the corresponding overlap value. A linear regression supplied the mean value of  $K$  over the set. Results are compiled in Table 1.

As can be seen, values of  $K$  vary much between cases. A general trend appears: when the coupling element  $H_{12}$  decreases (i.e., for long distances or near-orthogonal MOs), the value of  $K$  has to be increased to verify  $H_{12} = KS$ . This explains that the different sets, involving different magnitudes in overlap and/or coupling, feature different mean  $K$  values. We infer that the linear dependence of  $H_{12}$  on  $S$  is not rigorous, and may only be achieved inside given ranges of  $S$ . This problem could possibly be circumvented by maintaining different values of  $K$  for different overlap magnitudes, or by selecting the value of  $K$  that is typical of the structures under study. Neither solution is practical, however, when it comes to global exploration of a potential energy surface. Another issue concerns the  $\sigma$ - $\pi$  mixing which occurs for homodimer geometries where the planes of both monomers are nearly orthogonal. In such cases, the  $\pi$ - $\pi$  overlap (which monitors charge transfer in our model) vanishes, while the coupling element, albeit small, persists (because of limited  $\sigma$ - $\pi$  coupling which our model is unable to transcribe). This results in an altogether nonlinear relationship between  $H_{12}$  and  $S$ .

To ascertain the order of magnitude of the error introduced by the mean  $K$  approximation, we resorted to a more statistical approach: 500 distinct homodimers were randomly generated, with the constraints that (i) the center-to-center distance between monomers be smaller than 7.5 Å and (ii) the two monomers' bounding boxes (i.e., parallelepipeds enclosing the molecules' van der Waals spacefill) do not intersect. This provided a good



**Figure 1.** Absolute and relative error statistics of the determination of the off-diagonal coupling element using the mean  $K$  approximation, for 500 random  $(\text{ethylene}_2)^+$  geometries, with selected CI calculations as a reference.

sampling of the PES, with coupling values between  $10^{-1}$  and  $10^{-6}$  au. The mean value of  $K$  for this series ( $K_{\text{mean}} = 0.361$  au) was determined, and the error evaluated for each geometry by comparing  $K_{\text{mean}}S$  to the corresponding ab initio coupling. Absolute and relative error histograms appear on Figure 1. The difference between both graphs is striking: the relative error, albeit very small ( $<0.1\%$ ) for large values of  $S$ , can reach ca. 100% in some cases. On the contrary, the absolute error is always small and, with a mean value of 0.2 kcal/mol, is comparable to the precision of the model potential used for the diagonal terms of the Hamiltonian. This is of course due to the fact that high-overlap configurations contribute most significantly to the fit:  $K_{\text{mean}}$  is adequate to describe high-overlap configurations but fails for low-overlap ones. However, since these involve very small coupling elements, the absolute error remains acceptable in all cases.

Selected CI calculations do not scale favorably when the number of basis functions grows, preventing statistical studies on bigger systems. Considering the aforementioned trends as general, we will infer values of  $K$  from high-overlap homodimers in forthcoming systems (mostly stacked isomers in our case).

Finally, we investigated the basis set dependence on the value of  $K$ . From a reasonable number of basis functions on, basis set size has almost no influence on the coupling element  $H_{12}$ , in good agreement with Jortner's finding.<sup>20,21</sup> However, the

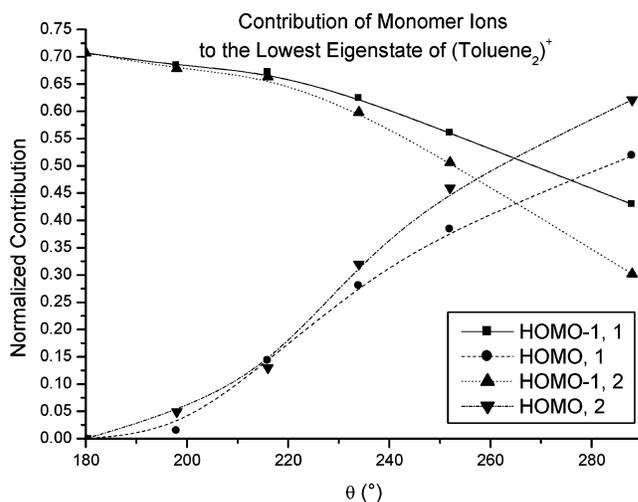
convergence of overlap integrals between molecular orbitals with basis set size generally occurs at larger basis dimensions. This apparent discrepancy is corrected by the choice of  $K$ : for a given cluster geometry, larger basis sets will generally give rise to larger overlap integrals but to lower values of  $K$ , while smaller basis sets will compensate for the underestimation of  $S$  by a higher value of  $K$ , in such a way that the coupling value (i.e., the product  $KS$ ) remains almost constant. This has two practical consequences: (i) the basis set used for the selected CI calculations from which  $K$  is inferred should be of sufficient quality to ensure correct transcription of the coupling element (most particularly the influence of polarization effects on the latter) and (ii) this same basis set is to be employed for the determination of overlap integrals in subsequent model potential computations, along with the corresponding value of  $K$ . The basis set 6-31G\* was found to verify these two conditions for the aromatic clusters under study, while still compatible with cost-effective selected CI calculations.

**III.2. Application: Naphthalene.** Naphthalene and ethylene both have in common a single low-lying ionization potential, and are similarly described in the framework of our model. Selected CIs on the naphthalene dimer are very costly, but can be alleviated by restricting the active space to the  $\pi$  subsystem. This amounts to neglecting the reorganization of the  $\sigma$  subsystem, which has negligible effect when no  $\sigma$ - $\pi$  mixing occurs, as is the case for stacked dimers (at least if the interplane distance is sufficiently large). Fitting sets were generated as for ethylene, by applying simple stepwise rotations and translations to high-overlap stacked homodimers. The mean value of  $K$  over the sets was found equal to 0.306 au, and was used for anthracene as well, since reasonably large selected CI calculations on anthracene are not feasible (due to high computational cost and basis size limitations inherent to the computer programs).

**III.3. Multiple Ionized Forms: Toluene and Benzene.** When two or more molecules interact, their respective states are split by a value ranging from twice the coupling (for two identical moieties) to four times the coupling (in the case of an infinite number of identical monomers giving birth to a band structure). Hence, the split can give rise to band mixing phenomena if two or more states are closer in energy than half the sum of their splitting values. In such a case, correct description of the system requires the inclusion of these states.

Toluene has two close-lying IPs, experimentally<sup>53</sup> located at 8.83 and 9.36 eV and involving the highest occupied molecular orbital (HOMO) and the one just below (HOMO-1), respectively. Both correspond to  $^2\Pi$  states. Ab initio selected CIs are carried out as mentioned earlier. The mean parameter  $K$ , which is averaged not only on different cluster geometries but on the four different types of overlap integrals, equals 0.281 au. For most clusters structures, this additional level of approximation does not deteriorate the quality of the fit. The necessity to include two IPs per monomer clearly appears on Figure 2, which represents the contribution of each monomer ion to the charged supersystem ground state for two stacked parallel molecules, one of which is rotated around an axis perpendicular to both molecular planes and joining the two entities' centers of mass (because the center of mass does not coincide with the center of the aromatic ring, the monomers are symmetry-inequivalent). Depending on the rotation angle, relative contributions of HOMO and HOMO-1 can vary drastically.

Although benzene is characterized by a single low-lying IP involving two degenerate MOs (HOMO and HOMO-1, both of



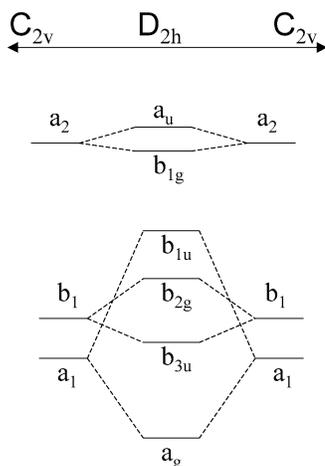
**Figure 2.** Normalized contributions of each monomer's two lowest ionized states to the ground-state electronic wave function of a stacked (toluene<sub>2</sub>)<sup>+</sup> cluster, as a function of the rotation angle  $\theta$  of the second monomer around the stack axis (which joins the two molecules' centers of mass). At  $\theta = 0^\circ$ , the second monomer is deduced from the first by a 3.5 Å translation along this axis. The contributions are the coefficients of the monomer states in the normalized supersystem eigenvector. Monomer states are denoted by a reference to the molecular orbital where the ionization takes place (HOMO or HOMO-1, cf. text) and by the index of the monomer (1 or 2).

$\pi$  type), it can formally be viewed as having two degenerate IPs involving one molecular orbital each, and is treated like toluene in the current model. The mean  $K$  value for the (benzene<sub>2</sub>)<sup>+</sup> system is 0.258 au.

**III.4. Limits of the Mean  $K$  Approximation: Pyridine.** The case of pyridine is a particularly difficult one. This molecule has three close-lying IPs inside a 1 eV range. Two  $\pi$ -character ions may result from an ionization inside the  $\pi$  subsystem, whereas the ionization of the nitrogen nonbonding orbital generates a  $\sigma$ -character ion. There is much experimental<sup>53-57</sup> as well as theoretical<sup>53,57</sup> controversy regarding the correct ordering of these three states<sup>53-57</sup> and the nature of the lowest one.

The study of the charged pyridine dimer in our model raises an additional difficulty compared to benzene derivatives. The effective Hamiltonian for the pyridine dimer is built upon six distinct valence bond structures and features nine coupling elements (as before, successive ions on the same molecule are considered decoupled). We will restrict our discussion to the case of two equivalent pyridine monomers of coplanar geometry, with the nitrogen atoms facing each other (point group  $D_{2h}$ ). In this configuration, the  $\sigma$ -nonbonding and  $\pi$  MOs on each monomer are orthogonal to each other, so that we need not preoccupy ourselves with  $\sigma$ - $\pi$  mixing phenomena. However, a single value of  $K$  must still be able to simultaneously account for the five remaining coupling elements governed by  $\sigma$ - $\sigma$  and  $\pi$ - $\pi$  overlaps, which may differ by several order of magnitude. We shall now investigate whether this is the case.

The pyridine monomer's molecular orbitals are representations of the  $C_{2v}$  point group. The three highest occupied MOs of neutral pyridine (two  $\pi$ -type and one  $\sigma$ -type, in decreasing energy order) are involved in the three ionic forms mentioned above. Their characters are  $a_2$ ,  $b_1$ , and  $a_1$ , respectively. In turn, the planar pyridine dimer belongs to point group  $D_{2h}$ . Its six highest occupied MOs are easily identified as pairwise combinations of each of the monomers' three highest occupied MOs, each of these yielding a *gerade/ungerade* pair; all of these dimer



**Figure 3.** Schematic molecular orbitals correlation diagram showing the relationship between the three highest MOs of the neutral  $C_{2v}$  pyridine molecule (left and right) and the six corresponding MOs of the neutral  $D_{2h}$  pyridine dimer (center). For the sake of lisibility, state energies and coupling values are not quantitatively represented. Refer to Tables 2 and 3 for exact values.

**TABLE 2: Ionization Potentials (eV) of the Pyridine Molecule at Two Different Levels of Theory<sup>a</sup>**

	$E(^2A_2)$	$E(^2B_1)$	$E(^2A_1)$
CASPT2 $\pi$	9.53	10.24	not available
ROMP2	9.39	10.05	9.38

<sup>a</sup> As a comparison, the first IP of pyridine is experimentally<sup>53</sup> located at 9.60 eV

orbitals belonging to a different symmetry class, no mixing occurs. The relationship between the corresponding ionic states are shown on Figure 3. The charge-transfer coupling element for a given pair of monomer states can be considered equal to half the difference between the resulting  $g/u$  dimer state energies. We may now check whether the three coupling elements are proportional to a relevant overlap integral, and if a mean proportionality factor is sufficient to account for all three.

The energy of the monomer  $\pi$ -type ions ( $^2A_2$  and  $^2B_1$ ) were obtained by a CASPT2 calculation using the  $\pi$  subsystem as the active space (five electrons in six orbitals). The same was done for the neutral molecule. However, such a method cannot be applied to the  $\sigma$ -type ( $^2A_1$ ) ion since it would require inclusion of the entire  $\sigma$  subsystem in the active space, which would de facto become prohibitively large. To be able to describe all ions at an equivalent level, we decided to check whether restricted open-shell MP2 calculations (ROMP2), modeled as a monodeterminantal CASPT2, would suffice. This approach yielded satisfactory results for the  $\pi$ -type ions, and was therefore applied to the  $\sigma$ -type one (Table 2). The experimentally postulated<sup>56</sup> accidental degeneracy between the  $^2A_1$  and  $^2A_2$  states is confirmed by these results.

The agreement between the CASPT2 and ROMP2 levels of theory was also noticed in the case of the  $D_{2h}$  pyridine dimer ion. ROMP2 calculations yielded the six possible ion energies. Coupling elements  $H_{12}$  were deduced as mentioned above; relevant MO overlap integrals between  $a_2$ ,  $b_1$ , and  $a_1$  MOs on each monomer were also evaluated. This procedure was iterated for N–N distances ranging from 2 to 4 Å; linear fit results of  $H_{12}$  vs  $S$  appear in Table 3. As predicted, the mean values for the three overlap integrals are very different and result in distinct  $K$  values. Even the two  $\pi$ -type overlaps differ of 1 order of magnitude (this is due to the fact that the  $b_1$  MO has a contribution on the nitrogen atom, while the  $a_2$  MO does not).

**TABLE 3: Typical Molecular Orbital Overlap Integrals and Corresponding Mean  $K$  Values (au) for the Charged  $D_{2h}$  Pyridine Dimer, Depending on the Nature of the Monomer Ion (cf. Text)**

	$a_2$	$b_1$	$a_1$
typical value of $S$ (au)	$10^{-3}$	$10^{-2}$	$10^{-1}$
mean $K$ (au)	0.349	0.289	0.402

The mean  $K$  approximation suffers from two different issues. First of all, one may infer from the case of pyridine that it is unable to simultaneously account for the coexistence of different charge-transfer interactions in a system, especially when different orbital types or distinct overlap ranges need be considered. Furthermore, the mean  $K$  value can differ significantly from species to species, even inside a given series of molecules displaying analogous properties (aromatics in our case). Hence, the method is expected to be much more reliable and easy to calibrate in homomolecular clusters than in heteromolecular ones.

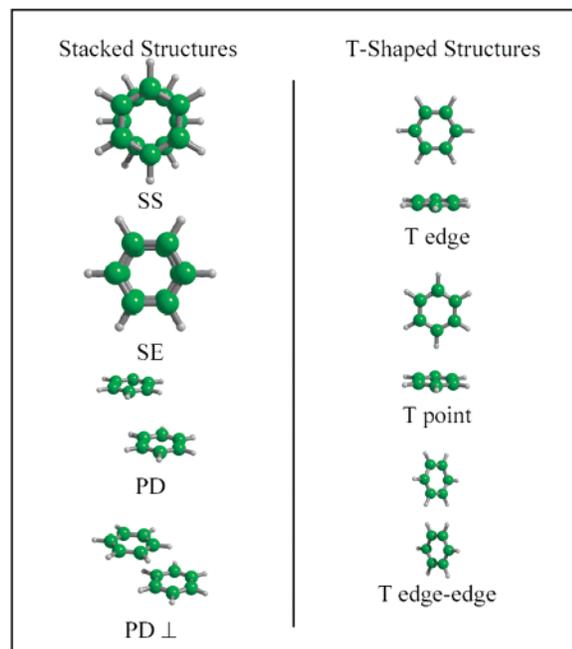
#### IV. Charge Transfer in Aromatic Clusters: A Model Potential Study

Once determined, the value of  $K_{\text{mean}}$  allows the calculation of the off-diagonal coupling terms of the model Hamiltonian. In the same way, adequate parametrization of the model potential is required to evaluate the diagonal energies of the valence bond structures.

**IV.1. Parametrization.** The electrostatic contribution to the interaction energy is described via interacting multipolar multicenter distributions on the monomers. These are derived from an ab initio calculation and do not need to be adjusted. Parameters for the polarization contribution include anisotropic bond polarizabilities; these were fitted so as to reproduce (i) the experimental and ab initio molecular dipole polarizability tensor of benzene,<sup>30</sup> (ii) its anisotropy, and (iii) the medium-range interaction energy of benzene-containing dimers, which is due exclusively to electrostatic and polarization terms. Once these contributions were correctly transcribed, the dispersion–repulsion–exchange repulsion terms could be adjusted. This was done on the neutral benzene dimer (vide infra), for which a wealth of theoretical and experimental studies have been published, by acting on the four parameters of the aromatic carbon and hydrogen atoms. We used Hobza’s counterpoise-corrected CCSD-T surface scans on stacked and T-shaped isomers,<sup>10</sup> which we complemented by several short-distance points using the same method and basis set.

This set of parameters was retained for all forthcoming studies on aromatic aggregates. Charged monomers were described using their own multipolar multicenter distribution, but with the same set of bond polarizabilities and dispersion–repulsion–exchange repulsion parameters than their neutral counterparts. Monomer ions were considered to share the same geometry as the corresponding neutral molecule, which is a good approximation in aromatics (despite a limited<sup>58</sup> Jahn–Teller effect in the benzene cation).

**IV.2. Global Potential Energy Surface Exploration for Small Aromatic Clusters.** Using the previously determined parameter set, a global potential energy surface exploration was carried out for small charged and neutral homoclusters of benzene, naphthalene and anthracene. While an extensive exploration is possible for dimer clusters, the number of isomers rapidly grows with the aggregation number in such a way that, for larger clusters, we mainly focused on the energetically low-lying conformations which we attempted to classify in structural families.

Bouvier *et al.*, Fig. 4

**Figure 4.** Generic structures of the stationary point of the  $(\text{benzene}_2)$  and  $(\text{benzene}_2)^+$  PES. The nature, the interaction energy and the geometric details of each structure are given in Tables 4 and 5 for neutral and charged clusters, respectively.

**TABLE 4: Interaction Energies ( $E_{\text{int}}$ ), Nature, and Typical Center-of-Mass Distances ( $R$ ) of the Stationary Points on the Neutral Benzene Dimer Potential Energy Surface (Figure 4)**

structure	$E_{\text{int}}$ (kcal/mol)	nature	$R$ (Å)
T-edge	-2.01	global minimum	4.93
T-point	-1.92	first-order saddle point	5.04
PD	-1.51	first-order saddle point	5.55
T-edge-edge	-1.49	minimum	6.01
SS	-0.98	second-order saddle point	3.94
SE	-0.98	third-order saddle point	3.94

**TABLE 5: Interaction Energies ( $E_{\text{int}}$ ), Nature, Typical Center-of-mass Distances ( $R$ ), and Delocalized Character of the Stationary Points on the  $(\text{Benzene}_2)^+$  Potential Energy Surface (Figure 4)**

structure	$E_{\text{int}}$ (kcal/mol)	nature	$R$ (Å)	delocalized character (%)
SS	-26.05	global minimum	3.02	50/50
SE	-25.99	first-order saddle point	3.03	50/50
PD⊥	-17.91	minimum	4.04	50/50
PD	-17.77	minimum	4.05	50/50
T-edge	-12.93	first-order saddle point	4.32	100/0
T-point	-12.41	second-order saddle point	4.27	100/0

Charged and neutral benzene dimer clusters were found to share common structure types, although the interaction energy, the nature (minimum or saddle point), and the center-of-mass distances of these structures may differ depending on the cluster charge. Figure 4 represents these generic structures, while Tables 4 and 5 give the nature, interaction energies and geometric details of each one for the neutral and charged benzene dimer cluster, respectively.

For the remaining clusters under study, the focus was set on minima rather than on saddle points. Since we usually found a clear dichotomy between neutral and charged structures, the corresponding figures (Figures 5–12) and tables (Tables 6–11) address the neutral and charged isomers separately.

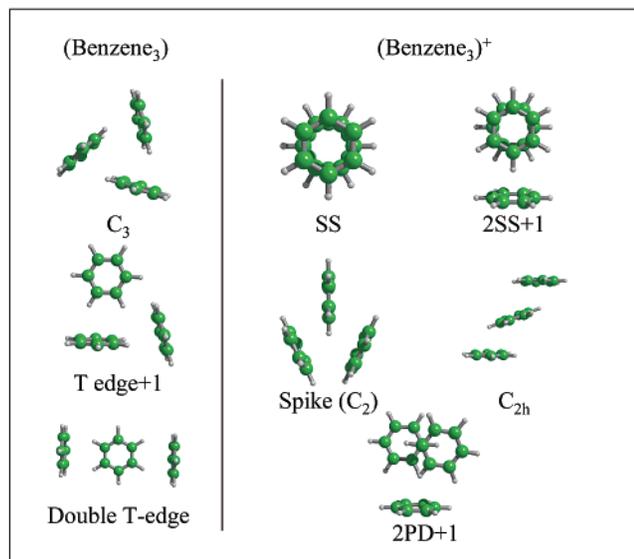
**IV.2.1. Neutral Benzene Dimer.** The optimal structures obtained for the neutral benzene PES are detailed in Table 4. They can be classified into two categories: stacked and T-shaped (Figure 4). The T-shaped isomers are due to attractive electrostatic and dispersion interactions, whereas the stacked ones are electrostatically unfavorable and rely exclusively on attractive dispersion. Polarization interactions are in all cases very small. Total interaction energies are lower than 2 kcal/mol, which stresses the necessity for very accurate calculations.

The global minimum is a T-shaped structure, denoted “T-edge” (the edge of the first ring is parallel to the second ring’s plane). Its center-to-center distance is very close to the experimental value of 4.96 Å.<sup>59</sup> A 30° rotation around the first molecule’s center of mass yields the “T-point” isomer (a hydrogen atom of the first ring points toward the second ring’s plane), which we identify as a first-order saddle point. In this respect, our calculation differs from other theoretical works<sup>10</sup> which find the T-point to be the global minimum and the T-edge to be a saddle point. This could result from subtle differences in the description of the quadrupole of benzene (necessarily less accurate than the dipole in our model, since the multipolar distribution is truncated at the quadrupolar term) and in the treatment of dispersion (especially concerning the hydrogen atoms). In any case, the two isomers are separated by a mere 0.09 kcal/mol (which is close to the precision of the potential), and unhindered rotation is expected to occur even at low temperatures, as stated by experimental intermolecular vibration studies.<sup>60</sup> Passim, we note the existence of another T-shaped structure, denoted “T-edge-edge”, which has often been overlooked in previous works.

The stacked structures (SS and SE), generally less stable than their T-shaped counterparts, are saddle points (up to third order) rather than minima. The “parallel displaced” ones (PD) are usually lower in energy than the “sandwich” ones, due to less unfavorable electrostatics.

**IV.2.2. Benzene Dimer Ion.** The benzene dimer ion’s optimal structures are detailed in Table 5. Although they belong to the same two classes as those of the neutral dimer (Figure 4), the stacked isomers are now much lower in energy than the T-shaped ones. For such stacked structures, the coupling elements of the model Hamiltonian may be up to twice as large as the diagonal interaction energies: interactions in the benzene dimer ion are governed by charge resonance. Hence, high-overlap configurations (e.g., stacked) will be favored. Analysis of the corresponding eigenvector show a perfect delocalization of the charge over the two monomers, whereas for T-shaped dimers (with overlap values close to nil), the charge is purely localized.

The “sandwich-staggered” (SS) structure is found to be the global minimum. The first-order saddle-point separating two SS structures, which we dub “sandwich-eclipsed” (SE), lies a mere 0.06 kcal/mol higher in energy (e.g., both conformations are degenerate to the precision of the model potential). This very weak energy difference is not due to charge resonance; during rotation around the axis joining the two monomers’ centers of mass, compensations between the four types of overlap are such that the global charge-transfer interaction remains exactly constant. It is mainly due to the dispersion and electrostatic contributions to the diagonal energies. As was the case for the neutral T-shaped dimer, rotational disorder should occur in sandwich structures; this has been postulated by SCF calculations.<sup>61</sup> As far as geometry is concerned, a 1-Å shortening of the interplane distance occurs relative to the neutral sandwich isomer.



**Figure 5.** Structures of the most relevant minima on the  $(\text{benzene}_3)$  and  $(\text{benzene}_3)^+$  PES (Tables 6 and 7).

The stacked nature of the global minimum, as well as the importance of charge resonance phenomena, was experimentally formulated in pioneering studies by Nishi et al.<sup>62,63</sup> and Hiraoka et al.<sup>64</sup> In these works, the binding enthalpies in the benzene dimer are estimated from  $-15$  to  $-20$  kcal/mol. Our calculation apparently overestimates this stabilization. The diagonal VB structures in stacked aggregates stem from an equilibrium between attractive polarization and (to a lesser extent) dispersion interactions on one hand, and repulsion on the other. The difference in interaction energy is probably due to an underestimation of the repulsive contribution: as already mentioned, repulsion and dispersion are fitted using the same set of atomic parameters, which in our case were calibrated on the neutral benzene dimer. Hence, our fit tends to favor the dispersive (long-distance) rather than the repulsive (short-distance) behavior. This problem is a recurrent one with our dispersion-repulsion model. It could possibly be solved by using a different set of parameters for the neutral and charged monomers, but for the sake of transferability, and for lack of convincing *ab initio* calculations on the benzene dimer ion, we chose to retain a unified version of the parameter set.

Finally, Nishi et al.<sup>63</sup> report a charge resonance band originating from the transition from the ground state of the dimer ion to the corresponding dissociative excited state. The corresponding photon energy is 1.3 eV (ca. 30 kcal/mol). Our model places the first dimer ion excited state at +13.295 kcal/mol, 39.3 kcal/mol higher than the ground state. The difference between experimental and theoretical figures is in fact probably not so large, because the experimentally detected excited state, being "hot", has a longer mean plane-to-plane distance than the ground state, resulting in a weaker coupling value, hence a smaller energy gap between the two states.

**IV.2.3. Neutral Benzene Trimer.** Three noteworthy minima were found on the neutral benzene trimer PES (Figure 5 and Table 6). Of these, only one is a typical three-molecule structure, whereas the others are derived from dimer geometries. This structure is actually the global minimum and features three equivalent benzene monomers; it belongs to point group  $C_3$ . Experimental evidence for such a global minimum has been provided by Raman studies.<sup>65</sup> The other two minima are built on the T-edge dimer structure and are differentiated by the position of the third benzene moiety (which interacts with both monomers in one case and with a single one in the other).

**TABLE 6: Interaction Energies ( $E_{\text{int}}$ ) of the Most Relevant Minima on the Neutral  $(\text{benzene}_n)$  PES,  $3 \leq n \leq 4$  (Figures 5 and 6)**

structure	$E_{\text{int}}$ (kcal/mol)
$C_3$	-5.89
T-edge+1	-5.38
double T-edge	-3.90
$C_2$	-10.84
$C_3$	-9.89
$C_1$	-9.01

**TABLE 7: Interaction Energies ( $E_{\text{int}}$ ) and Delocalized Character of the Most Relevant Minima on the  $(\text{benzene}_n)^+$  PES,  $3 \leq n \leq 4$  (Figures 5 and 6)**

structure	$E_{\text{int}}$ (kcal/mol)	delocalized character (%)
SS	-35.09	20/60/20
2SS+1	-34.39	50/50/0
spike ( $C_2$ )	-28.80	30/40/30
$C_{2h}$	-25.91	25/50/25
2PD+1	-25.03	50/50/0
3SS+1	-43.11	20/60/20/0
2SE+2T	-42.57	0/50/50/0
2SE+2V	-41.94	0/50/50/0
SS	-40.07	5/45/45/5
3 Spike+1	-36.45	30/40/30/0

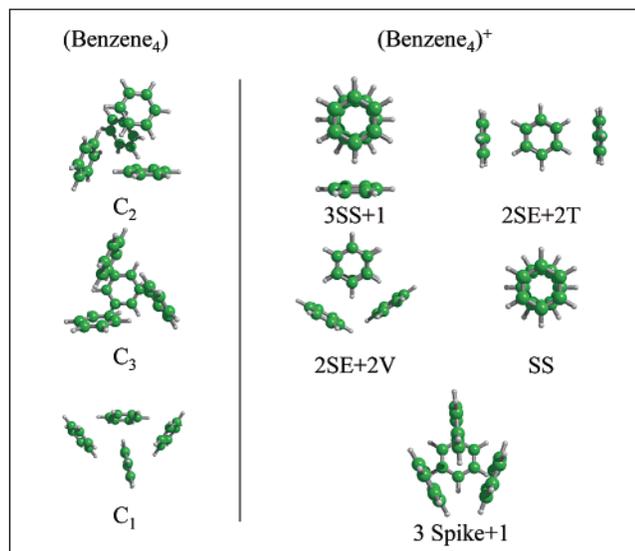
**IV.2.4. Benzene Trimer Ion.** The global minimum on the benzene trimer ion PES (Figure 5 and Table 7) is a sandwich-staggered (SS) structure. As already mentioned, such stacked structures maximize charge-transfer interactions through overlap integrals; the SS structure's nature as global minimum implies the predominance of charge resonance for the benzene trimer ion. The structure is symmetric, the central unit bearing the greater part of the charge. Such a sandwich structure for the benzene trimer ion has been experimentally inferred<sup>62</sup> from  $\pi$  orbitals stabilization.

However, the influence of polarization is greater than in the dimer ion, so that it competes with charge-transfer interactions. It is the driving force of isomers such as the 2SS+1, the third benzene unit's placement maximizing this effect rather than charge resonance (which is restricted to the two stacked monomers). The same is true for isomer 2PD+1. The remaining two minima feature intermediate behaviors. From the fact that the SS and 2SS+1 structures are close in energy (0.7 kcal/mol), we infer that there is a close tie between charge resonance and polarization effects in the benzene trimer ion, which thus constitutes a frontier case in the  $(\text{benzene}_n)^+$  series.

**IV.2.5. Neutral Benzene Tetramer.** Of the numerous stationary points on the neutral benzene tetramer PES (Figure 6, Table 6), only the global minimum, of point group  $C_2$ , features a characteristic four-coordinated structure. Higher minima are built on dimer and trimer motifs. Examples thereof are the  $C_3$  isomer, as well as the frequently occurring  $C_1$  structure.

There is experimental evidence for two types of symmetrically inequivalent sites<sup>65</sup> in the neutral benzene tetramer. An exp-6-1 model potential calculation by Van de Waal<sup>66</sup> proposes a tetrahedral cluster structure for which this is verified. However, our  $C_2$  isomer also fits the experimental bill.

**IV.2.6. Benzene Tetramer Ion.** The charge resonance-polarization competition, whose onset was witnessed for the trimer ion, amplifies in the case of the tetramer ion: polarization of neutral moieties by charged ones now becomes the predominant term. As a consequence, charge delocalization for most minima on the benzene tetramer ion PES does not involve more than three molecules; these clusters stem from the addition of one or two neutral monomers to a  $(\text{benzene}_2)^+$  or  $(\text{benzene}_3)^+$  motif. The



**Figure 6.** Structures of the most relevant minima on the  $(\text{benzene}_4)$  and  $(\text{benzene}_4)^+$  PES (Tables 6 and 7).

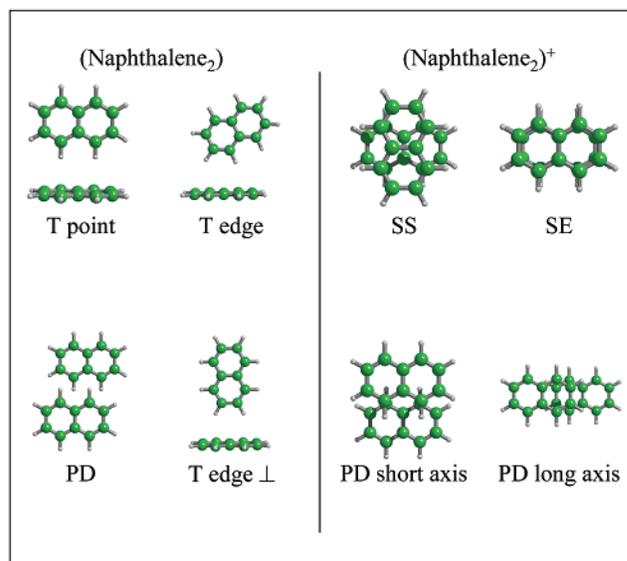
**TABLE 8: Interaction Energies ( $E_{\text{Int}}$ ) of the Most Relevant Minima or Saddle Points (SP) on the Neutral  $(\text{Naphthalene}_n)$  PES,  $2 \leq n \leq 4$  (Figures 7–9)**

structure	$E_{\text{int}}$ (kcal/mol)
T-point	-3.33
T-edge	-2.29
PD	-2.74
T-edge $\perp$	-2.74
SS (first-order SP)	-2.48
SE (third-order SP)	-2.25
$C_{3h}$	-10.17
T-point+1	-7.74
double T-point	-6.66
$C_1$	-15.74
$C_2$	-15.49
$C_{2h}$	-15.46

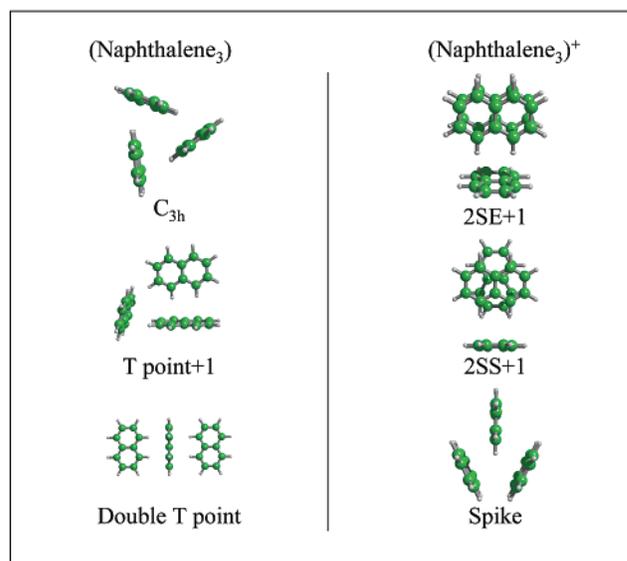
only exception is the SS structure, where the charge spreads on all four monomers; although still a minimum, it is relatively high in energy, unlike the corresponding structure for the trimer ion.

Of the numerous stationary points found on the surface, the most relevant families are detailed on Figure 6 and in Table 7. The global minimum (3SS+1) is the combination of a  $(\text{benzene}_3)^+$  sandwich-staggered structure and a neutral monomer, and features the best polarization–charge delocalization tradeoff.

**IV.2.7. Neutral Naphthalene Clusters: Dimer through Tetramer.** Global PES studies for neutral naphthalene dimer through tetramer aggregates were carried out. Results are compiled in Table 8, and in Figures 7–9. As for benzene, the naphthalene dimer global minimum is a T-shaped structure. Most other stationary points are reminiscent of those of  $(\text{benzene}_2)$ , with the exception that, due to the lower symmetry of naphthalene, T-point isomers are generally more stable than T-edge ones. This does not come as a surprise, since dispersion and repulsion are the driving forces in these clusters. The trimer PES, in agreement with Raman experiments,<sup>67</sup> features a triangle-shaped  $C_{3h}$  structure as the global minimum. A triangle-shaped global minimum was also found for the neutral benzene trimer; however, the energy separation between this structure and subsequent local minima is more clear in the present case (2.5 kcal/mol compared to 0.5 kcal/mol for benzene). The double T-point structure mentioned by Wessel et al.<sup>68</sup> is also a minimum, albeit high in energy. Experimental data on the



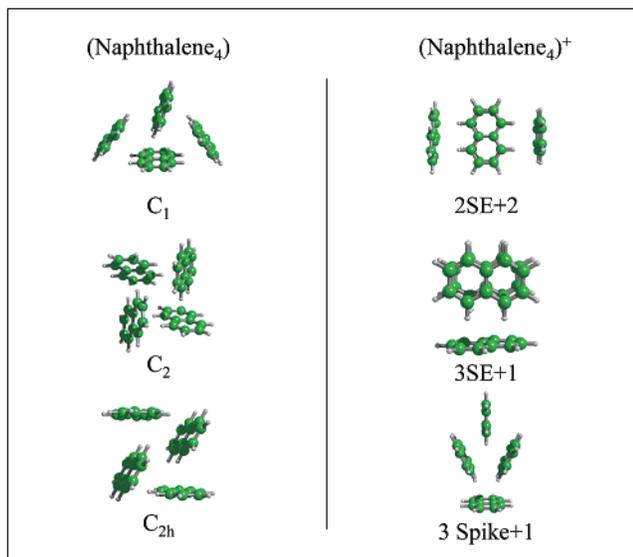
**Figure 7.** Structures of the most relevant stationary points on the  $(\text{naphthalene}_2)$  and  $(\text{naphthalene}_2)^+$  PES (Tables 8 and 9).



**Figure 8.** Structures of the most relevant minima on the  $(\text{naphthalene}_3)$  and  $(\text{naphthalene}_3)^+$  PES (Tables 8 and 9).

naphthalene tetramer<sup>67–70</sup> hints at four symmetry-inequivalent sites. Most authors acknowledge a low-symmetry herringbone structure (which occurs in crystals containing naphthalene-like molecules<sup>71</sup>), although they admit that it should result, strictly speaking, in only two different site types. From our point of view, we find the global minimum to be devoid of symmetry. The two following minima, respectively  $C_2$  and  $C_{2h}$ , feature two different symmetry-inequivalent sites. All higher minima stem from  $(\text{naphthalene}_3)$  or  $(\text{naphthalene}_2)$  building blocks.

**IV.2.8. Naphthalene Dimer through Tetramer Cations.** Stationary points (Table 9 and Figures 7–9) on the naphthalene dimer cation PES closely resemble that found for the benzene dimer cation. Sandwich structures are the most stable, which hints at the importance of charge resonance. However, both the staggered and eclipsed conformers are now minima (rotational barriers between these structures will be discussed hereafter). Charge delocalization occurs equally on both moieties. The predominance of polarization over charge resonance occurs earlier on than in  $(\text{benzene}_n)^+$ : the global minimum of the trimer PES is a 2SE+1 structure, where charge delocalization only



**Figure 9.** Structures of the most relevant minima on the (naphthalene<sub>4</sub>) and (naphthalene<sub>4</sub>)<sup>+</sup> PES (Tables 8 and 9).

**TABLE 9: Interaction Energies ( $E_{\text{int}}$ ) and Delocalized Character of the Most Relevant Minima on the (naphthalene<sub>*n*</sub>)<sup>+</sup> PES,  $2 \leq n \leq 4$  (Figures 7–9)**

structure	$E_{\text{int}}$ (kcal/mol)	delocalized character (%)
SS	-19.74	50/50
SE	-19.34	50/50
PD short axis	-14.30	50/50
PD long axis	-12.42	50/50
2SE+1	-30.42	50/50/0
2SS+1	-29.73	50/50/0
spike	-25.24	25/50/25
double T-point	-23.51	12/75/12
PD+1	-22.95	50/50/0
2SE+2	-40.71	50/50/0/0
3SE+1	-38.22	25/50/25/0
3spike+1	-35.51	25/50/25/0

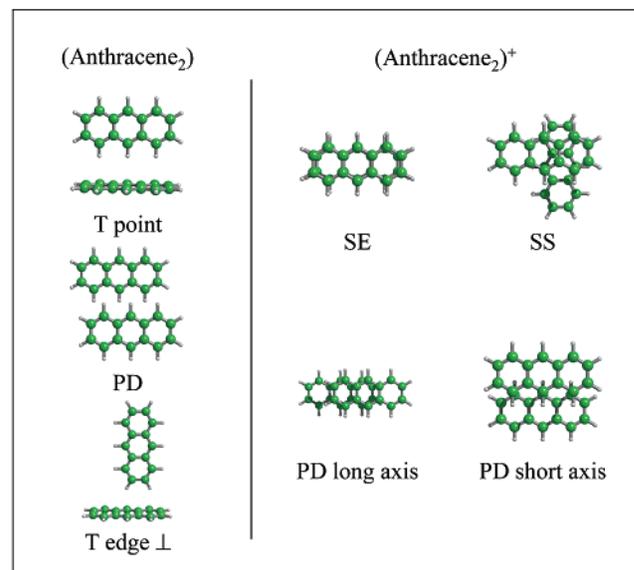
involves the two stacked monomers; the first real (naphthalene<sub>3</sub>)<sup>+</sup> isomer lies ca. 5 kcal/mol higher in energy. On the tetramer ion PES, no structure involving delocalization on the four monomers were found altogether; the global minimum features a sandwich-eclipsed charged dimer with two complementary neutral rings.

**IV.2.9. Neutral Anthracene Clusters: Dimer through Tetramer.** The PES of both charged and neutral anthracene clusters are remarkably complex, even for such small clusters as the ones we consider. In particular, a cluster of size  $n$  contains numerous minima due to the adjunction of a monomer to a size  $n - 1$  cluster. In the following two paragraphs, we will restrict ourselves to typical size  $n$  structures, which we classify in families (Table 10, Figures 10–12). Clearly, a point is reached where global PES exploration becomes cumbersome and should leave place to molecular dynamics at typical jet vibrational temperatures (fifty to a hundred kelvins).

The stationary points on the dimer PES closely resemble that of the naphthalene dimer. The global minimum is T-shaped, held together by attractive dispersion and electrostatic interactions. Most other minima are stacked, attractive dispersion interactions compensating unfavorable electrostatics. To date, the question of whether the global minimum is stacked or T-shaped is still unclear.<sup>72</sup> The trimer PES features a triangle-shaped C<sub>3</sub> global minimum, compatible with previous studies;<sup>73</sup> the gap with consecutive trimer structures widens (6.2 kcal/

**TABLE 10: Interaction Energies ( $E_{\text{int}}$ ) of the Most Relevant Minima on the Neutral (Anthracene<sub>*n*</sub>) PES,  $2 \leq n \leq 4$  (Figures 10–12)**

structure	$E_{\text{int}}$ (kcal/mol)
T-point	-4.59
PD	-4.05
SS	-3.86
T-edge $\perp$	-2.80
C <sub>3</sub>	-14.14
SS $D_{2h}$	-7.94
C <sub>2</sub>	-22.18



**Figure 10.** Structures of the most relevant minima on the (anthracene<sub>2</sub>) and (anthracene<sub>2</sub>)<sup>+</sup> PES (Tables 10 and 11).

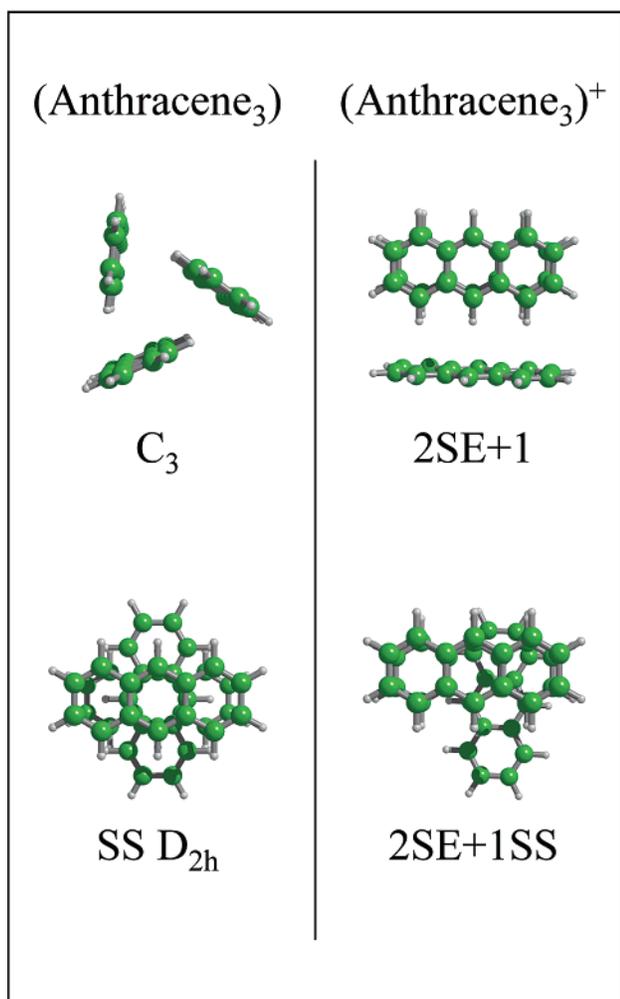
mol) compared to the benzene and naphthalene trimers. These forthcoming minima are mostly 2+1 T-shaped structures. The sandwich-staggered  $D_{2h}$  structure, identified as the global minimum in other works,<sup>15,74</sup> lies much higher in energy. The tetramer PES is characterized by an impressive number of possible structures inside a 0.1 kcal/mol range. Particularly relevant among these are different classes of C<sub>2</sub>-type isomers.

**IV.2.10. Anthracene Dimer through Tetramer Cations.** The anthracene dimer cation's most probable structure is a sandwich-eclipsed one, with other types of stacked minima forming local minima on the PES. Spatial delocalization over the two moieties is expected in each case.

As was the case for naphthalene, polarization effects rapidly dominate over charge resonance interactions, so that the global minimum of the trimer ion PES is a 2SE+1 structure where the charge extends over two monomers. Delocalization over the three monomers can occur, but the resulting isomers are more than 5 kcal/mol higher in energy. The same trend appears in the tetramer ion, characterized by a 2SE+2 global minimum (Table 11, Figures 10–12).

### IV.3. Global Trends in Small Charged Aromatic Clusters.

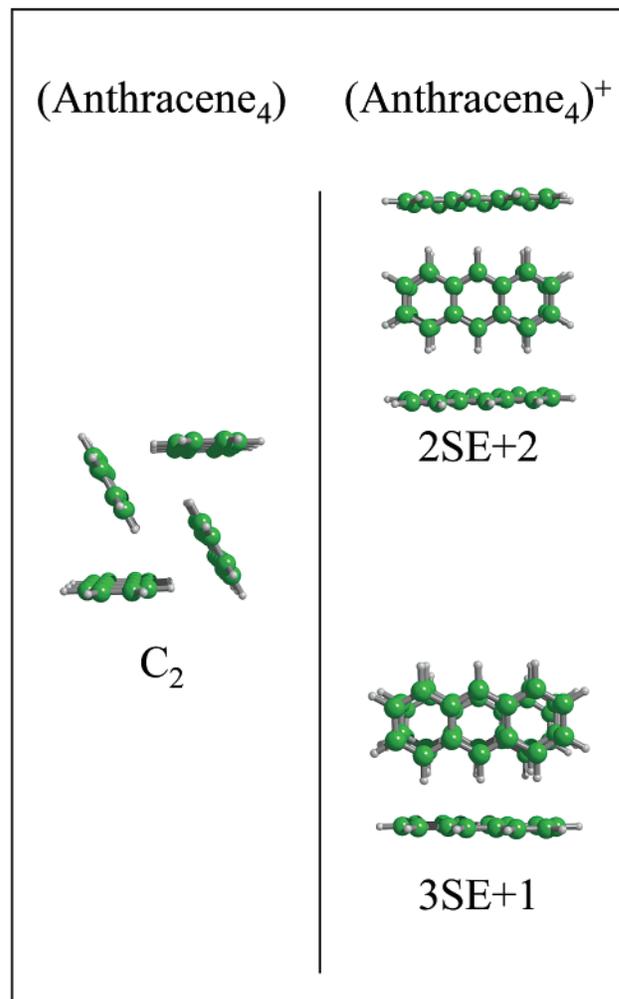
**IV.3.1. Rotation Barriers in Charged Dimer Stacks.** As explained in the previous paragraphs, sandwich structures for charged aromatic dimers feature two different conformers: eclipsed (i.e., parallel) and staggered (resulting from the rotation of one monomer around the ring plane normal by an angle of 30° for benzene and 90° for naphthalene and anthracene). The comparison of the interconversion barriers that separate these conformers hints at the degree of rotational disorder that one can expect to find in these systems. We performed surface scans



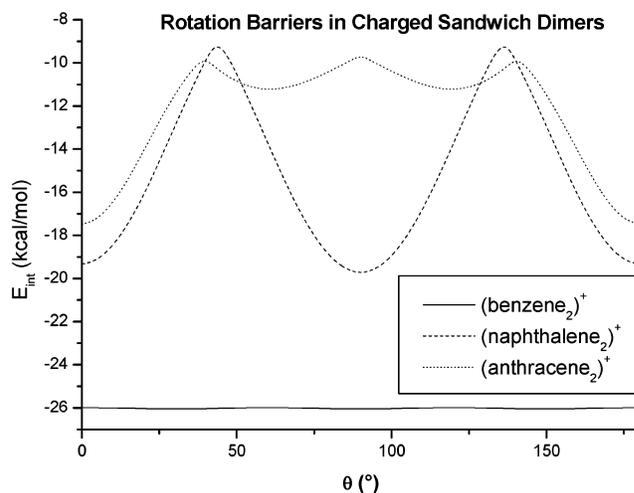
**Figure 11.** Structures of the most relevant minima on the (anthracene<sub>3</sub>) and (anthracene<sub>3</sub>)<sup>+</sup> PES (Tables 10 and 11).

along the rotation coordinate, reoptimizing the plane-to-plane distance at each step (variations of this distance with the rotation angle were found to be negligible); results are displayed on Figure 13. The barrier in the case of benzene is almost nonexistent since, as already noted, the charge-transfer interaction in these structures does not depend on the rotation angle. On the contrary, the absolute value of the overlap in the naphthalene dimer ion is maximal at 0° and 90°, and nil at 45°, and the resulting coupling element varies much more (15 kcal/mol) than the diagonal elements. Hence, the rotation barrier in this system is high and, although SS and SE structures are of similar energies, interconversion is expected to be difficult. Charge-transfer effects are also responsible for the barrier in anthracene, but the variation of the overlap is more complex, its absolute value reaching a maximum at 0° ( $H_{12} = 10$  kcal/mol) and at 60° ( $H_{12} = 2.3$  kcal/mol). This explains the two energy-inequivalent minima which may be seen on the plot. The barrier, albeit lower than that of the naphthalene dimer ion (ca. 9 kcal/mol), should not permit easy interconversion. Hence, the benzene dimer ion should be the only system to feature rotational disorder at typical jet temperatures.

**IV.3.2. Charge Resonance vs Polarization: Where Does the Border Lie?** In charged aromatic clusters, as a rule of thumb, charge-transfer effects dictate the structure of small aggregates, while for larger entities polarization effects take over. The former case results in the addition of a (generally stacked) monomer to the cluster of size  $n - 1$ , forming a  $n$ -molecular



**Figure 12.** Structures of the most relevant minima on the (anthracene<sub>4</sub>) and (anthracene<sub>4</sub>)<sup>+</sup> PES (Tables 10 and 11).



**Figure 13.** Variation of the interaction energy in the benzene, naphthalene and anthracene sandwich dimer ions, as a function of the rotation angle  $\theta$  of the second monomer around the stack axis. Optimized interplane distances show almost no variation with  $\theta$  (3.02, 3.30, and 3.54 Å for benzene, naphthalene, and anthracene, respectively).

cluster where the charge is delocalized on all monomers; the latter case is characterized by the addition of a neutral monomer to the  $(n-1)$ -sized cluster, with no alteration to this cluster's charge delocalization properties. As mentioned in the previous

**TABLE 11: Interaction Energies ( $E_{\text{int}}$ ) and Delocalized Character of the Most Relevant Minima on the (Anthracene) $_n^+$  PES,  $2 \leq n \leq 4$  (Figures 10–12)**

structure	$E_{\text{int}}$ (kcal/mol)	delocalized character (%)
SE	−17.48	50/50
SS	−14.73	50/50
PD long axis	−13.84	50/50
PD short axis	−13.45	50/50
2SE+1	−30.73	50/50/0
2SE+1SS	−25.23	36/47/17
2SE+2	−42.89	0/50/50/0
3SE+1	−39.34	48/26/26/0

paragraphs, polarization effects become dominant from  $n = 4$  in benzene, and  $n = 3$  in naphthalene and anthracene, polarizabilities increasing with molecular size. The additional neutral monomer may be placed at different near-equivalent sites with regard to the  $(n - 1)^+$  group, progressively generating a neutral shell around the charged cluster core; this explains the rise in the number of isomers as cluster dimensions grow. As a comparison, the charge resonance–polarization border was found to lie at  $n = 3$  or 4 in charged xenon clusters<sup>75</sup> and at  $n = 3$  in ionized argon aggregates.<sup>76</sup> In all these cases, the extent of charge delocalization is quite limited.

**IV.3.3. Evolution of the Adiabatic-Vertical Ionization Potential Difference.** For each of the aforementioned clusters, adiabatic and vertical ionization potentials were calculated. The vertical ion was obtained by a single-point calculation at the neutral global minimum geometry. Two limit cases were considered for the adiabatic ion: the “closest” ion (obtained by allowing the neutral structure to relax to the closest-lying attraction basin on the ion PES) and the “global” ion (i.e., the global minimum on this PES). Results are presented in Table 12.

For a given cluster size, adiabatic and vertical IP differences decrease in the series benzene–naphthalene–anthracene. This is due to the fact that, for naphthalene and anthracene, common structures exist for both the neutral and the ionic states, whereas in benzene there is a clear dichotomy between neutral (T-shaped) and charged (stacked) structures. The dispersive forces that bind neutral stacks increase from benzene to anthracene, up to a point where they are able to counteract the repulsive electrostatic interactions. On the contrary, for a given species, the IP difference increases as the aggregation number grows. This is not surprising for such small charged clusters, in which most monomers (2 or 3 out of 4) participate in the delocalization of the charge. Nevertheless, we mentioned that the importance of charge-transfer interactions rapidly decreases, so that for very large aggregates a negligible part of the structure will be involved in reorganization upon ionization. Hence, the IP difference is expected to converge toward a value typical of the reorganization energy of the ionic cluster core.

The “closest” and “global” adiabatic ions were found to differ in all cases. This does not come as a surprise. Small clusters feature such a structural difference between neutral and charged geometries that the two corresponding points on the ion PES lie very far from each other, the probability of encountering a local minimum between them thus being very high. Larger clusters are characterized by an important number of local minima, with the same consequence.

Finally, Table 12 compares calculated adiabatic IPs in benzene clusters to the experimental values reported by Krause et al.<sup>77</sup> and Nishi et al.<sup>78</sup> Our “closest” value for the dimer is in relatively good agreement with both experimental determinations. However, as  $n$  grows, Krause’s values tend to stagnate and the gap with our results widens. Krause et al. discuss that low ionization intensities, as well as selectivity issues for larger clusters, might mitigate their results; however, it would be unfair not to mention that the theoretical determination of adiabatic IPs has its own difficulties, since it requires a good description of a whole area of the PES, and not only its minima.

**IV.4. Charge Delocalization vs Polarization in Stacked Clusters.** Among the different trends discussed above, the polarization vs charge delocalization competition in ionic clusters was proved to have a dominant influence over cluster structures and properties. Its manifestation in larger clusters thus seemed of interest. Unfortunately, the incentive for global exploration of a potential energy surface diminishes as the aggregation number grows, because of the exponential rise of the number of near-degenerate minima. We decided to restrict our study to the stacked class of isomers, which brings down the number of degrees of freedom per monomer from six to two (namely, translation along and rotation about the stack axis). As previously noted, although stacked structures dominate in small charged aromatic aggregates, their occurrence and stability tend to decrease for larger clusters, up to the point where they might not even be minima. Nevertheless, such constrained geometries are not devoid of interest: variations upon this theme include liquid crystals, allotropic varieties in solids (i.e., graphite), DNA helices, molecular wires, etc.

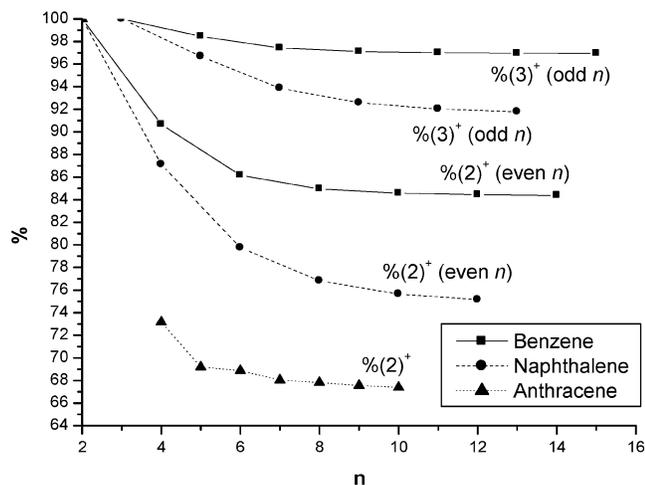
Stacked clusters of benzene (up to  $n = 15$ ), naphthalene (up to  $n = 13$ ) and anthracene (up to  $n = 10$ ) were constructed and submitted to constrained optimization.

**IV.4.1. Extent of the Charge Delocalization.** The model Hamiltonian approach, which describes the supersystem in the basis of its monomers’ states, bears close resemblance to the exciton theory. Stack eigenstates shall hence be studied using tools derived from exciton studies.<sup>79</sup>

The participation of a given monomer  $M_i$  to the charge delocalization phenomenon is most easily described, in the framework of our model, by the magnitude of the coefficient of the supersystem eigenvector associated with the valence bond form  $(M_1M_2\dots M_i^+ \dots M_n)$ . Alternately, the number of coherently

**TABLE 12: Computed and Experimental (exptl) Adiabatic and Vertical Ionization Potentials (kcal/mol) in Small Aromatic Homoclusters**

cluster	adiabatic IP (closest)	adiabatic IP (global)	vertical IP	difference (closest)	difference (global)	adiabatic IP (exptl)
(benzene) <sub>2</sub>	197.3	189.1	206.3	9.0	17.2	199.5 <sup>77</sup> 196.7 <sup>78</sup>
(benzene) <sub>3</sub>	190.2	183.9	203.7	13.4	19.7	197.9 <sup>77</sup>
(benzene) <sub>4</sub>	188.5	180.9	203.1	14.6	22.2	197.2 <sup>77</sup>
(naphthalene) <sub>2</sub>	176.7	171.3	179.9	3.2	8.6	
(naphthalene) <sub>3</sub>	172.6	167.5	178.7	6.1	11.3	
(naphthalene) <sub>4</sub>	167.5	162.7	174.2	6.7	11.5	
(anthracene) <sub>2</sub>	162.8	158.9	163.8	1.0	4.9	
(anthracene) <sub>3</sub>	160.1	155.2	163.4	3.2	8.1	
(anthracene) <sub>4</sub>	157.2	151.1	161.3	4.1	10.2	



**Figure 14.** Contribution of the two ( $2^+$ ) or three ( $3^+$ ) central monomers (bearing the major part of the charge) to the global ground-state wave function of benzene, naphthalene and anthracene stack ions, as a function of the number  $n$  of monomers in the stack.

coupled monomers in a supersystem eigenstate may be expressed using a simple “participation ratio” model:<sup>80</sup>

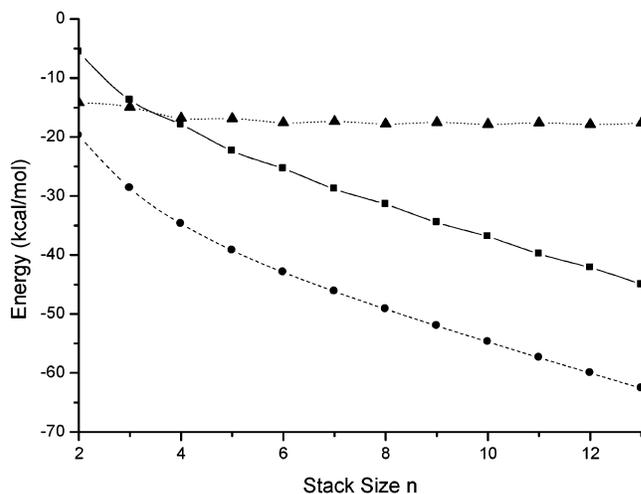
$$L_k = \sum_{\text{monomers } m} \left[ \sum_{\text{states } i} (C_{k,m}^i)^2 \right]^2 \quad (7)$$

where the participation ratio  $1/L_k$  (PR) is directly related to the number of monomers involved in the  $k$ th supersystem eigenstate, and  $C_{k,m}^i$  is the (normalized) eigenvector coefficient for monomer  $m$  in its  $i$ th state (as a reminder, two different ionic states have to be considered for the benzene monomer).

For all systems, the PR indicates that two to three monomers participate to the supersystem ion. However, this only reflects a mean trend. For benzene and naphthalene, delocalization occurs over two or three monomers for even or odd values of  $n$ , respectively, while for anthracene delocalization occurs mostly on two moieties. This can be seen by plotting the contribution of the ( $2^+$ ) or ( $3^+$ ) motif to the entire eigenstate (Figure 14). Albeit large, this contribution in benzene and naphthalene decreases monotonically with  $n$ , trading a slight delocalization on other monomers. Convergence is reached for  $n = 8$  in benzene, and  $n = 10$  in naphthalene. The case of anthracene is peculiar: as already mentioned, delocalization involves mostly 2 monomers but never goes beyond 4 units: the contribution of the ( $4^+$ ) motif equals 100% for every  $n \geq 4$ .

**IV.4.2. Stack Geometries.** The issue of stack geometries is closely related to that of charge delocalization, since it dictates the magnitude of the overlap integrals used as a model for charge-transfer interactions.

Benzene and naphthalene stacks feature a sandwich-staggered structure, with a rotation of  $30^\circ$  and  $90^\circ$  around the stack axis from one building block to the next, respectively. The distance between consecutive monomers vary from typical ion-stack distances in the ( $2^+$ )/( $3^+$ ) motifs to typical neutral-stack distances (ca. 1 Å longer) for outer monomers. This justifies the rapid but continuous decrease of off-diagonal elements and, hence, of charge delocalization. On the contrary, anthracene stacks feature a sandwich-eclipsed ( $4^+$ ) group surrounded by sandwich-staggered monomers. The overlap between HOMOs in staggered anthracene dimers are nil, which explains the sudden stop in charge delocalization outside the ( $4^+$ ) motif: polarization and charge-transfer interactions in anthracene stacks are mutually exclusive. This was not the case for benzene (for which the



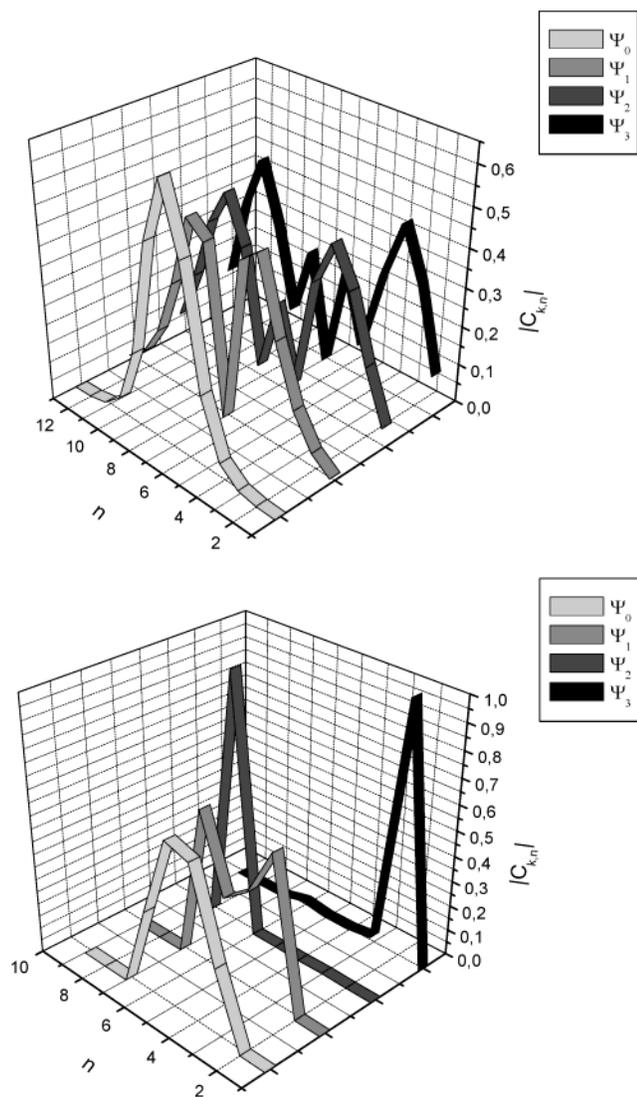
**Figure 15.** Global energy of stacked naphthalene ions, with charge-transfer interactions enabled (circles and dashes) or disabled (squares and continuous line), as a function of the aggregation number  $n$ . The difference between the two terms represents the stabilization due to charge resonance, and is also plotted (triangles and dots).

charge resonance effect depends only on the plane-to-plane distance) or naphthalene (for which a  $90^\circ$  rotation angle will accommodate both electrostatic and charge-transfer interactions).

**IV.4.3. Polarization vs Charge Resonance.** We come back to the polarization vs charge-transfer competition, whose importance was stressed throughout this work. A quantitative assessment of both terms may be computed. Single-point energies were determined for the optimized stack geometries, with the charge localized on the central monomer, and were compared with the corresponding charge-transfer-inclusive calculations. The difference between the two figures is the stabilization originating from charge delocalization. As can be seen on Figure 15, it does not vary much as  $n$  grows, with a mean value of  $-20$  kcal/mol for benzene,  $-16$  kcal/mol for naphthalene and  $-12$  kcal/mol for anthracene. Although dominant in small clusters, it is rapidly caught up by the other effects (among which polarization is paramount); this occurs as soon as  $n = 3$  and explains the aforementioned limited delocalization behavior.

**IV.4.4. Excited States of Ion Stacks.** The delocalized behavior of the ion stack excited states are compared with that of the ion ground state on Figure 16, via their participation ratio and the magnitude of the corresponding eigenvector coefficients. In the case of benzene and naphthalene, these states display a more delocalized character than the ground state: for higher excited states, the charge is able to reach the limits of the stack. The monomers bearing the greater part of the charge are different from those of the ground state and depend on the parity of the excited-state index. This “harmonic” behavior is due to the regularity of the stack structure. In anthracene the situation is far more complex, due to the presence of three distinct zones (a SE area surrounded by two SS areas). Low-index excited states tend to involve the central SE area, while higher excited states are located in either one of the SS zones and display strictly localized behaviors (owing to null coupling elements).

Excitation energies are represented on Figure 17. Beyond a sufficiently large value of  $n$ , additional monomers do not affect the properties of either the ground or the excited states, so that the corresponding excitation energy converges. Naturally, this process is slower for higher excited states whose delocalized characters are more extensive. In the case of naphthalene and anthracene, the first excitation energy converges toward a value



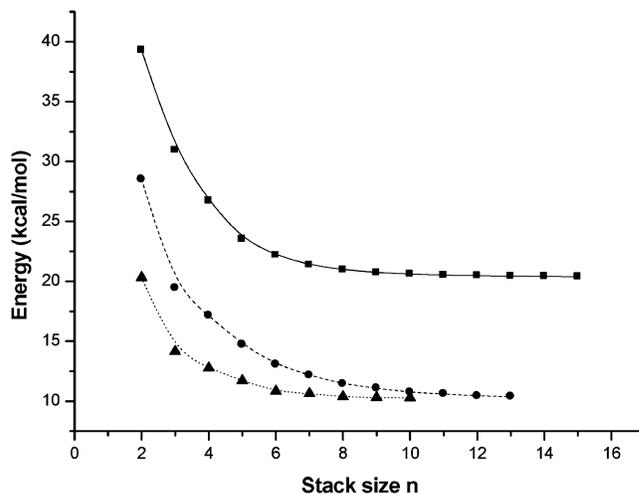
**Figure 16.** Topography of the ground state ( $\Psi_0$ ) and the first three excited states ( $\Psi_1$  to  $\Psi_3$ ) of typical naphthalene (above) and anthracene (below) stack ions. The coefficients  $C_{k,n}$  denote the contribution of monomer  $n$  to normalized eigenstate  $k$ .

of ca. 10 kcal/mol, which is typical of high-energy vibrational modes, and could lead to experimental mismatch between the electronic and vibrational excited state of the ion stacks.

## V. Conclusion and Perspectives

Notwithstanding the theoretical difficulties that arise when attempting to account for charge-transfer interactions, the present work aims at proving that they may be correctly described using a simple model, compatible with global potential surface exploration or molecular dynamics simulations. Its application to charged aromatic homoclusters reveals interesting trends on the structure and electronic properties of these systems, such as the charge resonance vs polarization competition. Remarkably few studies deal with charge-transfer phenomena in aromatic aggregates; we hope that this work will provide an incentive for future works in this domain.

The limits of global potential energy surface exploration at 0 K for multidimensional or complex cases has been highlighted throughout this study; it suggests, as the most obvious extension to the present work, to utilize molecular dynamics simulations to characterize the coupling between charge delocalization and intermolecular vibration as a function of temperature. This may



**Figure 17.** Excitation energies corresponding to the first excited state of benzene (squares and continuous line), naphthalene (circles and dashes) and anthracene (triangles and dots) stack ions, as a function of the aggregation number  $n$ .

be achieved by conducting the simulation on the PES corresponding to the lowest adiabatic state at each geometry. However, a more realistic description would have to take into account possible switches between adiabatic states in regions where several of these come in close vicinity (i.e., zones of avoided crossings), thus accounting for nonradiative transitions such as the ones that occur, for example, during the relaxation of an initially excited charged aggregate through internal conversion. To this effect, we are currently implementing a trajectory surface hopping formalism. Molecular dynamics simulations with electronic transitions should yield valuable insight into charge-transfer phenomena, and provide further proof of the interest of simple yet accurate potentials as tools for the study of complex systems.

**Acknowledgment.** The authors would like to express their gratitude to Jean-Pierre Flament for valuable information on overlap integral derivatives, as well as to Fernand Spiegelman, Tomas Baer, François Piuze, and Michel Mons for advice and orientation.

## References and Notes

- (1) Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112*, 5525.
- (2) Cabarcos, O. M.; Weinheimer, C. J.; Lisy, J. M. *J. Chem. Phys.* **1998**, *108*, 5151.
- (3) Kim, K. S.; Tarakeshwar, P.; Lee, J. Y. *Chem. Rev.* **2000**, *100*, 4145.
- (4) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1994**, *100*, 2975.
- (5) Dunning, T. H. *J. Phys. Chem. A* **2000**, *104*, 9062.
- (6) Tsuzuki, S.; Uchimaru, T.; Tanabe, K. *J. Mol. Struct.* **1994**, *307*, 107.
- (7) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Am. Chem. Soc.* **1994**, *116*, 3500.
- (8) Jaffe, R. L.; Smith, G. D. *J. Chem. Phys.* **1996**, *105*, 2780.
- (9) Meijer, E. J.; Sprik, M. *J. Chem. Phys.* **1996**, *105*, 8684.
- (10) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 18790.
- (11) Braidia, B.; Hiberty, P. C. *J. Phys. Chem. A* **2000**, *104*, 4618.
- (12) Braidia, B.; Lauvergnat, D.; Hiberty, P. C. *J. Chem. Phys.* **2001**, *115*, 90.
- (13) Braidia, B.; Hiberty, P. C.; Savin, A. *J. Phys. Chem. A* **1998**, *102*, 7872.
- (14) Smith, G. D.; Jaffe, R. L. *J. Phys. Chem.* **1996**, *100*, 9624.
- (15) White, R. P.; Niesse, J. A.; Mayne, H. R. *J. Chem. Phys.* **1998**, *108*, 2208.
- (16) Niesse, J. A.; Mayne, H. R. *J. Phys. Chem. B* **1997**, *101*, 9137.
- (17) Spirko, V.; Engvist, O.; Soldan, P.; Selzle, H. L.; Schlag, E. W.; Hobza, P. *J. Chem. Phys.* **1999**, *111*, 572.

- (18) Dang, L. X. *J. Chem. Phys.* **2000**, *113*, 266.
- (19) Jortner, J.; Bixon, M.; Langenbacher, T.; Michel-Beyerle, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12759.
- (20) Voityuk, A. A.; Rosch, N.; Bixon, M.; Jortner, J. *J. Phys. Chem. B* **2000**, *104*, 9740.
- (21) Voityuk, A. A.; Jortner, J.; Bixon, M.; Rosch, N. *J. Chem. Phys.* **2001**, *114*, 5614.
- (22) Derepas, A. L.; Granucci, G.; Brenner, V.; Dognon, J.-P.; Millie, P. *Phys. Chem. Chem. Phys.* Submitted for publication.
- (23) Gresh, N. J. *Comput. Chem.* **1995**, *16*, 856.
- (24) Brenner, V. Thesis, Université Paris XI, 1993.
- (25) Brenner, V.; Millie, P. *Z. Phys. D-At. Mol. Clusters* **1994**, *30*, 327.
- (26) Derepas, A. L.; Soudan, J. M.; Brenner, V.; Millie, P. *J. Comput. Chem.* **2002**, *23*, 1013.
- (27) Claverie, P. *Intermolecular Interactions: from Diatomics to Biopolymers*; Wiley: New York, 1978.
- (28) Vigné-Maeder, F.; Claverie, P. *J. Chem. Phys.* **1988**, *88*, 4934.
- (29) Lefevre, R. J. W. *Advances in Physical Organic Chemistry*; Academic Press: New York, 1965; Vol. 3.
- (30) Calaminici, P.; Jug, K.; Koster, A. M.; Ingamells, V. E.; Papadopoulos, M. G. *J. Chem. Phys.* **2000**, *112*, 6301 and references therein.
- (31) Mulder, F.; Van Dijk, G.; Huiszoon, C. *Mol. Phys.* **1979**, *38*, 577.
- (32) Bouvier, B.; Brenner, V.; Millié, P.; Soudan, J. M. In preparation.
- (33) Kitaigorodski, A. I. *Tetrahedron* **1961**, *14*, 975.
- (34) Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 296.
- (35) Garell, T.; Niel, J. C.; Orland, H.; Velikon, B. J. *J. Chem. Phys.* **1991**, *88*, 2473.
- (36) Poteau, R.; Spiegelman, F. *J. Chem. Phys.* **1993**, *98*, 6540.
- (37) Bertolus, M.; Brenner, V.; Millié, P.; Mailet, J.-B. *Z. Phys. D-At. Mol. Clusters* **1997**, *39*, 239.
- (38) Kirkpatrick, S.; Gellet, C. D.; Vecchi, M. P. *Science* **1983**, *220*, 671.
- (39) Kirkpatrick, S. *J. Stat. Phys.* **1984**, *34*, 975.
- (40) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes (FORTRAN Version)*; Cambridge University Press: Cambridge, 1989; p 307.
- (41) Metropolis, N.; Rosenbluth, A.; Rosenbluth, M.; Teller, A.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.
- (42) Kumanova, M. D. *Mol. Phys.* **1972**, *23*, 407.
- (43) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
- (44) *Jaguar 4.1*; Schrödinger, Inc.: Portland, OR, 1991–2001.
- (45) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.
- (46) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
- (47) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
- (48) Dupuis, M.; Marquez, A.; Davidson, E. R. HONDO 95.3 from CHEM-Station; IBM Corporation: Kingston, NY, 1995.
- (49) Huron, B.; Malrieu, J.-P.; Rancurel, P. *J. Chem. Phys.* **1973**, *58*, 5745.
- (50) Cimraglia, R. *Int. J. Quantum Chem.* **1996**, *60*, 167.
- (51) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
- (52) Andersson, K.; Barysz, M.; Bernhardsson, A.; Blomberg, M. R. A.; Fleig, T.; Fülischer, M. P.; de Graaf, C.; Hess, B. A.; Karlström, G.; Lindh, R.; Malmqvist, P.-A.; Neogrady, P.; Olsen, J.; Roos, B. O.; Sadlej, A. J.; Schütz, M.; Schimmelpfennig, B.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stalring, J.; Thorsteinsson, T.; Vervazov, V.; Widmark, P.-O. *Molcas*, version 5.0; Lund University: Lund, Sweden, 2000.
- (53) Kimura, K.; Katsumata, S.; Achiba, Y.; Yamazaki, T.; Iwata, S. In *Handbook of HeI Photoelectron Spectra of Fundamental Organic Molecules*; Press, J. S. S., Ed.; Halsted Press: New York, 1981; p 189.
- (54) Berg, J. O.; Parker, D. H.; El-Sayed, M. A. *Chem. Phys. Lett.* **1978**, *56*, 411.
- (55) Kuhn, W. F.; Lilly-Leister, D.; Kao, J.; Lilly, A. C. *J. Mol. Struct.* **1989**, *212*, 37.
- (56) Kao, J.; Eyerman, C.; Southwick, E.; Leister, D. *J. Am. Chem. Soc.* **1985**, *107*, 5323.
- (57) Almlöf, J.; Roos, B.; Wahlgren, U.; Johansen, H. *J. Electron. Spectrosc.* **1973**, *2*, 51.
- (58) Lindner, R.; Muller-Dethlefs, K.; Wedum, E.; Haber, K.; Grant, E. R. *Science* **1996**, *271*, 1698.
- (59) Arunan, E.; Gutowski, H. S. *J. Chem. Phys.* **1993**, *98*, 4294.
- (60) Schaeffer, M. W.; Maxton, P. M.; Felker, P. M. *Chem. Phys. Lett.* **1994**, *224*, 544.
- (61) Milosevich, S. A.; Saichek, K.; Hinchey, L.; England, W. B.; Kovacic, P. *J. Am. Chem. Soc.* **1983**, *105*, 1088.
- (62) Ohashi, K.; Nishi, N. *J. Chem. Phys.* **1991**, *95*, 4002.
- (63) Ohashi, K.; Nishi, N. *J. Chem. Phys.* **1993**, *98*, 390.
- (64) Hiraoka, K.; Fujimaki, S.; Aruga, K. *J. Chem. Phys.* **1991**, *95*, 8413.
- (65) Henson, B. F.; Venturo, V. A.; Hartland, G. V.; Felker, P. M. *J. Chem. Phys.* **1993**, *98*, 8361.
- (66) Van de Waal, B. W. *Chem. Phys. Lett.* **1986**, *123*, 69.
- (67) Schaeffer, M. W.; Kim, W.; Maxton, P. M.; Romascan, J.; Felker, P. M. *Chem. Phys. Lett.* **1995**, *242*, 632.
- (68) Wessel, J. E.; Syage, J. A. *J. Phys. Chem.* **1990**, *94*, 737.
- (69) Saigusa, H.; Sun, S.; Lim, E. C. *J. Chem. Phys.* **1992**, *97*, 9072.
- (70) Saigusa, H.; Sun, S.; Lim, E. C. *J. Phys. Chem.* **1992**, *96*, 2083.
- (71) Wyckoff, R. W. G. *Crystal Structures*; Interscience: New York, 1960.
- (72) Chakraborty, T.; Lim, E. C. *J. Phys. Chem.* **1993**, *97*, 11151.
- (73) Piuzzi, F.; Dimicoli, I.; Mons, M.; Millie, P.; Brenner, V.; Zhao, Q.; Soep, B.; Tramer, A. *Chem. Phys.* **2002**, *275*, 123.
- (74) Xiao, Y.; Williams, D. E. *Chem. Phys. Lett.* **1993**, *215*, 17.
- (75) Amarouche, M.; Durand, G.; Malrieu, J.-P. *J. Chem. Phys.* **1988**, *88*, 1010.
- (76) Grigorov, M.; Spiegelmann, F. *Surf. Rev. Lett.* **1996**, *3*, 211.
- (77) Krause, H.; Ernstberger, B.; Neusser, H. J. *Chem. Phys. Lett.* **1991**, *184*, 411.
- (78) Shinohara, H.; Nishi, N. *J. Chem. Phys.* **1989**, *91*, 6743.
- (79) Bouvier, B.; Gustavsson, T.; Markovitsi, D.; Millie, P. *Chem. Phys.* **2002**, *275*, 75.
- (80) Marguet, S.; Markovitsi, D.; Millie, P.; Sigal, H.; Kumar, S. *J. Phys. Chem. B* **1998**, *102*, 4697.



## 3. Mécanismes de l'Ionisation dans les Agrégats Aromatiques

### 3.1. Introduction

L'étude précédente a permis de mettre en lumière d'intéressantes propriétés structurales de petits agrégats, neutres et chargés, de benzène, de naphthalène et d'anthracène. A notamment été abordé le problème de la détermination des potentiels d'ionisation verticaux et adiabatiques des agrégats aromatiques neutres. Une controverse expérimentale subsiste encore à ce sujet, les différences topologiques entre surfaces d'énergie potentielle associées aux agrégats neutres et chargés rendant l'interprétation des résultats ambiguë. Les données issues de l'exploration de ces surfaces à 0 K ne sont pas directement comparables aux expériences d'ionisation, qui se produisent à une température non nulle (celle d'un jet moléculaire supersonique) et incluent des effets quantiques (recouvrement Franck-Condon des fonctions d'onde des états vibroniques *ad hoc* du neutre et de l'ion). L'étude qui suit tente de modéliser ces effets par des calculs de dynamique moléculaire sur les surfaces d'énergie potentielle des agrégats aromatiques neutres, à des énergies simulant les effets quantiques jugés importants. Une statistique sur les agrégats chargés obtenus par ionisation verticale de configurations issues de la dynamique est en effet susceptible de fournir une image plus réaliste des mécanismes d'ionisation dans ces espèces, donc de permettre une comparaison plus légitime avec les résultats expérimentaux. L'article qui suit a été soumis au *Journal of Physical Chemistry A*.

### 3.2. Article

# Investigation of the Photoionization Mechanism of Small Aromatic Homoclusters

Benjamin Bouvier<sup>‡</sup>, Philippe Millié and Michel Mons

*CEA Saclay, DSM/DRECAM/SPAM, Laboratoire Francis Perrin, URA 2453, Bât. 522*

*91191 Gif-sur-Yvette, France*

This work proposes to simulate the photoionization process of small aromatic clusters of benzene, naphthalene and anthracene, using the combination of a tailored model potential and a valence bond Hamiltonian. Classical molecular dynamics simulations are carried out on the potential energy surface of the neutral cluster, at energies meant to transcribe quantum mechanical effects such as intermolecular vibrational zero-point energy and the classically forbidden area of the vibrational ground state wavefunction. Statistics of the ionization potential values obtained by vertical ionization of neutral configurations originating from these simulations yield valuable information on the respective topologies of the neutral and ionic potential energy surfaces. Comparison of calculated values with experimental one-photon and two-photon ionization potential determinations and ion current curves is also performed.

---

<sup>‡</sup> Corresponding author. Phone : +33-1-6908-3788. Fax : +33-1-6908-8707. E-mail : [bouvier@pandora.saclay.cea.fr](mailto:bouvier@pandora.saclay.cea.fr)

## I. Introduction

Molecular clusters of aromatic compounds have been of interest to the theoretical and the experimental physical chemistry communities for a long while. They provide insight into the mechanisms of  $\pi$  interactions, which are as ubiquitous as they are complex. Aromatic and heteroaromatic moieties are present in many compound systems such as biopolymers and liquid crystals, and may often be held responsible for important properties of such systems. Molecular ions of aromatic assemblies also have attracted great attention. Albeit more complex to characterize than their neutral counterparts, they are compulsory to describe important phenomena such as charge transfer or charge transport in large systems. These phenomena are largely based on the interplay between the ionic and neutral varieties of local aromatic moieties, as well as their interactions as part of a larger system (DNA for example<sup>1-3</sup>).

It is widespread knowledge that aromatic clusters, whether neutral or charged, are complicated to characterize from the theoretical point of view<sup>4</sup>. Neutral clusters are most conveniently described using model potentials<sup>5-9</sup>, since *ab initio* calculations on such systems require high – and costly – levels of theory<sup>10</sup>; these calculations may however be mandatory if the model potential is to be correctly calibrated. Ionic clusters raise a different issue altogether. Their properties differ tremendously from these of their neutral counterparts, mainly because of charge resonance phenomena which are theoretically complicated to account for.

A previous communication<sup>9</sup> addresses these questions in much more detail. It combines a tailored model potential and an effective Hamiltonian approach to provide a correct description of both charged and neutral species, as well as charge resonance phenomena bridging the gap between both, for a set of aromatic homoclusters of benzene, naphthalene and anthracene. This methodology was applied to the extensive study of the topology of the potential energy surfaces (PES) of both neutral and ionized clusters, yielding global and local minima and their relative binding energies. Comparison to both experimental data and high-level *ab initio* calculations, whenever available, confirmed the validity of this approach. Using data gathered from PES topologies, theoretical ‘vertical’ and ‘adiabatic’ ionization potentials (IP) were determined. However, the experimentally measured ionization potential is more likely in-between these two values.

The present study focuses more acutely on this last issue. The first photoionization potentials of small aromatic homoclusters (benzene, naphthalene and anthracene, with coordination numbers up to four) were previously inferred from global 0 K PES explorations. They are to be tackled, in the present study, from the standpoint of classical molecular dynamics simulations on the PES associated with the electronic ground state of the neutral cluster, using the model potential whose reliability already has been demonstrated. The energies at which these simulations are led is chosen so as to model quantum effects such as the exploration of the classically forbidden part of the vibrational ground state wavefunction of the neutral cluster. Configurations originating from these simulations are assumed to be vertically ionized, and the IP statistics thus obtained are described and analyzed. A closer match with experimental conditions – and results – is expected of this model, since it is more realistic than the one previously employed<sup>9</sup>. Comparison with experiments is especially interesting, as the experimental issue of the photoionization threshold of small benzene clusters has been the object of much controversy<sup>11-18</sup>.

## II. Methodology

### *2.1. Model Potential*

The intermolecular model potential used in this work has been thoroughly described in the last instalment of this work<sup>9</sup>, and shall only be succinctly outlined in the present paper. Based on second-order perturbation theory, it contains individual terms corresponding to electrostatic, dipole polarization, dispersion and repulsion contributions to the intermolecular interaction energy. The electrostatic term originates from interactions between distributed multipoles (up to the quadrupole) fitted from a one-particle density matrix *ab initio* calculation. The polarization energy is inferred from a self-consistent approach using anisotropic site polarizabilities explicitly designed for aromatics. The dispersion and repulsion terms are two-body terms based on atom-specific parameters. A neutral aromatic moiety and its charged counterpart share the same dispersion and repulsion interactions but have distinct electrostatic and polarization parameters. Intramolecular rearrangement upon ionization was neglected, the neutral and the ionic forms of each monomer sharing the same geometry. This can be considered a good approximation for aromatics, despite a limited Jahn-Teller effect in the benzene cation<sup>19</sup>.

## *2.2. Effective Hamiltonian Framework*

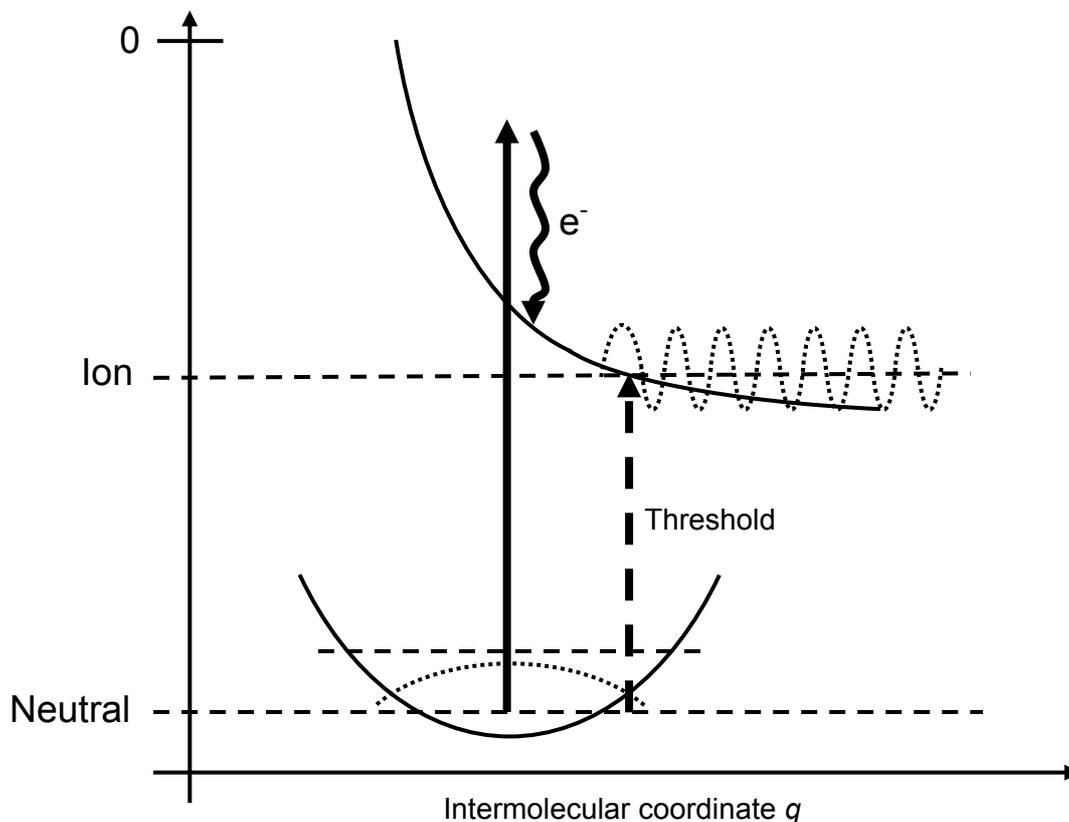
Charge transfer interactions in aromatic homoclusters are described *via* a valence bond treatment, using an effective Hamiltonian written out in the basis of all possible cluster configurations where the charge is localized on a given monomer, and whose energies are determined using the aforesaid model potential. Coupling elements between these basis wavefunctions are considered proportional to overlap integrals between the molecular orbitals of the charged monomers that are implicated in the electron transfer, on each of the two valence bond forms. The proportionality factor is a monomer-dependent environment-independent parameter which is computed from *ab initio* selected configuration interaction calculations on the corresponding dimer cluster. Diagonalization of the effective Hamiltonian yields ground and excited states wavefunctions for the supersystem. A single ionization potential per monomer is explicitly considered, higher IPs being neglected. This is a reasonable approximation for the description of neutral and charged clusters in their electronic ground state. Details on the parameterization of the potential, including reference *ab initio* calculations, may be found in reference <sup>9</sup>.

## *2.3. Quantum Description of the Photoionization Process*

Neutral and ionized clusters of aromatic molecules such as benzene, naphthalene and anthracene have been shown to possess very different structures. Therefore, adiabatic photoionization to the vibrational ground state of the ion is forbidden by the Franck-Condon principle. Ionization may only lead to high-lying excited intermolecular vibrational states of the ion, which enable the ion complex to explore conformations similar to those of the neutral ground state. An experimental signature of such a peculiarity is the slow rise of the photoionization current at the onset, found for example for the (benzene)<sub>2</sub> cluster.<sup>11,12,14,18</sup>

In molecular jet experiments, neutral aromatic clusters are mainly formed in their vibrational ground state, at temperatures typically lower than 10 K. This suggests that the threshold photoionization process of such clusters should be addressed in terms of spatial extent of the neutral vibrational ground state wavefunction. Due to the differences in the oscillating behavior of this wavefunction and that of the high-lying vibrationally excited states of the ion cluster, the only non-vanishing contribution to the Franck-Condon overlap originates from the classically forbidden parts of the neutral wavefunctions (Figure 1). This

work attempts to emulate this quantum mechanical behavior with classical molecular dynamics as a tool<sup>20,21</sup>.



**Figure 1:** Schematic representation of the potential energy surfaces of the neutral and ionic clusters, in the regions relevant to the photoionization process. The vibrational wavefunctions for the neutral ground state and a highly excited ionic state are plotted along an arbitrary intermolecular coordinate  $q$ . The dashed arrow materializes the position of the threshold. The full arrow denotes a vertical transition from the neutral potential energy surface, the ejected electron ( $e^-$ , curly arrow) taking away the excess energy.

#### *2.4. Molecular Dynamics Calculations*

Classical molecular dynamics was chosen as a means to sample the classically forbidden area of the vibrational ground state of the neutral cluster. It allows explicit consideration of all intermolecular degrees of freedom, which is particularly useful for cases where the relevant intermolecular coordinate  $q$  (Figure 1) is not an eigenmode of the system, but a combination of several eigenmodes. Rigid-body molecular dynamics calculations on neutral aromatic clusters are carried out on the PES associated with the model potential, in the

microcanonical (N, V, E) ensemble. As a consequence, intramolecular vibrations of the monomers were neglected. The configurations that we wish to sample are those belonging to the classically forbidden part of the vibrational ground state wavefunction. They can be accessed through a microcanonical molecular dynamics calculation if the total energy of the system (potential plus kinetic) is higher than the value of the zero-point energy (ZPE). As an example, the theoretical ZPE for the neutral benzene dimer was found to be of 130 K<sup>22</sup>.

The use of quaternions to describe molecular rotation bypasses gimbal lock issues. Forces are computed using analytical derivatives of the potential. Integration of translational movements resorts to the velocity Verlet algorithm, whereas the mid-step implicit leapfrog algorithm due to Svanberg<sup>23</sup> is employed where rotational motion is concerned. Time steps of 2 fs were found consistent with a good energy conservation (fluctuations under 0.1%) for all systems under study.

### *2.5. Sampling of the Potential Energy Surfaces*

Molecular dynamics calculations were performed at two different energies. Denoting the well-bottom energy of the neutral global minimum of the cluster as  $E_0$ , these energies were chosen to be  $E_0+100$  K and  $E_0+300$  K. The first simulation should give a reasonable description of the neutral cluster in its vibrational ground state (including the ZPE). The second one aims at simulating quantum effects, by giving access to configurations that are part of the classically forbidden region of the vibrational ground state wavefunction. The value of the ZPE is expected to vary depending on the cluster under study, so that using a single value (100 K) in all cases is an approximation. However, since this value is expected to increase with the aggregation number and to decrease in the series benzene-naphthalene-anthracene, employing a mean value has been deemed a good choice to compensate for the lack of adequate theoretical results.

Starting configurations for the molecular dynamics simulations were chosen so as to comply with the aforementioned energy criteria. Total simulation times were 2 ns in all cases.

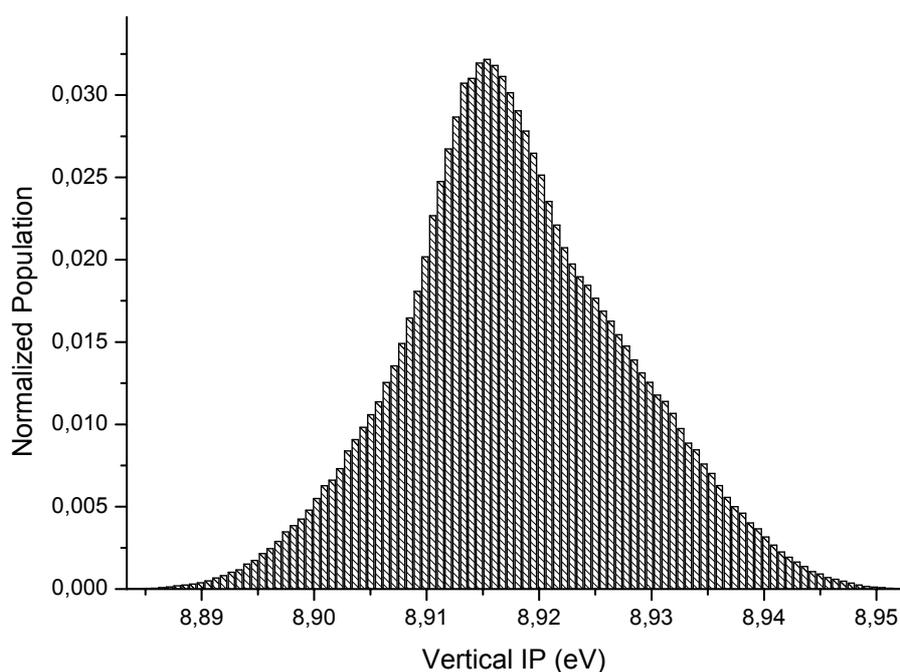
### *2.6. Photoionization Probability Cross-Section*

As already mentioned, the neutral molecular dynamics simulation aims at describing the cluster in its vibrational ground state, which is supposed to be the only one populated at typical molecular jet temperatures. Hence, the total energy of the neutral cluster was taken equal to the energy of the global minimum well-bottom on the neutral potential energy

surface ( $E_0$ ), plus 100 K to account for the ZPE, and was supposed constant throughout the dynamics. The higher energy of  $E_0+300$  K is designed to give access to configurations inside the classically forbidden region of the neutral ground state wavefunction, which cannot be sampled by a classical simulation at  $E_0+100$  K. These configurations, however, are part of the neutral vibrational ground state and share the same total energy of  $E_0+100$  K, regardless of the total energy of the dynamics from which they were obtained. In the framework of the model, photoionization occurs *via* a vertical transition from classical configurations, the ejected electron taking away the excess energy; strictly speaking, this model is only valid at threshold. Resulting ion clusters may show varying energies, depending on the associated neutral cluster configuration from which the vertical ionization process occurred.

The geometry of the neutral cluster was extracted at each time step of the molecular dynamics simulation. At this geometry, the energy associated with the electronic ground state wavefunction of the singly-charged cluster was computed using the valence bond Hamiltonian model. Subtraction of the energy of the vibrational ground state of the neutral cluster from the energy of the ion cluster yielded an instant vertical ionization potential. These were compiled into a histogram, featuring the normalized count of configurations versus the values of vertical IP. An example of such a graph, in the case of the benzene dimer at  $E_0+100$  K, may be found on Figure 2.

The ion current associated with a given IP value was computed as the area of the histogram taken between the threshold and the IP, the total area of the histogram (taken between the minimum and the maximum thresholds) being normalized to unity.



**Figure 2:** Histogram of ionization potential (IP) values encountered during a 2 ns molecular dynamics simulation on the neutral (benzene)<sub>2</sub> cluster at energy  $E_0+100$  K, where  $E_0$  is the potential energy of the global minimum of the cluster.

### III. Results and Discussion

#### *3.1. Energy Ranges of Neutral and Ionic Clusters*

Molecular dynamics simulations were run, at the two aforementioned energies, on neutral clusters of (benzene)<sub>2</sub> to (benzene)<sub>4</sub>, (naphthalene)<sub>2</sub> to (naphthalene)<sub>4</sub>, and (anthracene)<sub>2</sub> to (anthracene)<sub>3</sub>, under the previously mentioned conditions. The energies of the vibrational ground states of the neutral clusters ( $E_0+100$  K, see paragraph 2.6) are given in Table 1. The energy ranges of the ionic clusters encountered during the molecular dynamics simulations are compiled in Table 2, along with the energy values of the global minimum on the corresponding ion potential energy surface.

<b>Cluster</b>	<b>Binding energy (eV)</b>
(Benzene) <sub>2</sub>	-0.070
(Benzene) <sub>3</sub>	-0.238
(Benzene) <sub>4</sub>	-0.453
(Naphthalene) <sub>2</sub>	-0.127
(Naphthalene) <sub>3</sub>	-0.424
(Naphthalene) <sub>4</sub>	-0.665
(Anthracene) <sub>2</sub>	-0.182
(Anthracene) <sub>3</sub>	-0.596

**Table 1:** Intermolecular binding energy of the vibrational ground state of the neutral clusters under study. The energy reference is that of  $n$  non-interacting monomer moieties (at infinite distances), where  $n$  denotes the aggregation number.

Cluster	Neutral Energy	Ion Energy Range (eV)		Ion Global
		Minimum	Maximum	Minimum (eV)
(Benzene) <sub>2</sub>	E <sub>0</sub> +100 K	-0.432	-0.363	-1.130
	E <sub>0</sub> +300 K	-0.466	-0.225	
(Benzene) <sub>3</sub>	E <sub>0</sub> +100 K	-0.770	-0.555	-1.522
	E <sub>0</sub> +300 K	-0.820	-0.442	
(Benzene) <sub>4</sub>	E <sub>0</sub> +100 K	-0.994	-0.753	-1.869
	E <sub>0</sub> +300 K	-1.096	-0.522	
(Naphthalene) <sub>2</sub>	E <sub>0</sub> +100 K	-0.482	-0.450	-0.856
	E <sub>0</sub> +300 K	-0.488	-0.450	
(Naphthalene) <sub>3</sub>	E <sub>0</sub> +100 K	-0.883	-0.721	-1.319
	E <sub>0</sub> +300 K	-0.916	-0.575	
(Naphthalene) <sub>4</sub>	E <sub>0</sub> +100 K	-1.338	-1.133	-1.765
	E <sub>0</sub> +300 K	-1.343	-0.754	
(Anthracene) <sub>2</sub>	E <sub>0</sub> +100 K	-0.553	-0.531	-0.758
	E <sub>0</sub> +300 K	-0.562	-0.512	
(Anthracene) <sub>3</sub>	E <sub>0</sub> +100 K	-0.995	-0.920	-1.333
	E <sub>0</sub> +300 K	-1.002	-0.920	

**Table 2:** Intermolecular energy ranges (as minimum and maximum) of the ionic clusters encountered during the simulation. As a comparison, intermolecular binding energies of the global minimum of each ionic cluster potential energy surface are also given. The energy reference for a size  $n$  cluster ion is that of  $n-1$  neutral monomer moieties and one charged monomer moiety, at infinite distances.

As can be seen, in all cases the minimum and maximum values that bracket the ion energy range are much greater than the energy of the global minimum. This confirms the fundamental differences in the topologies of the potential energy surfaces of the ionic and neutral clusters, for which the location and nature of stationary points are very dissimilar. It is particularly remarkable in the clusters of benzene, where the difference may amount to more than 1 eV. This is consistent with the qualitative conclusions of our previous work<sup>9</sup>, and indicates that the true adiabatic ionization potential of these clusters should not be

experimentally available. Also noteworthy is the fact that, among the ion clusters encountered during this study, none were found to exhibit a dissociative character (*i.e.*, the corresponding binding energy was always negative). This is interesting evidence that the experimentally observed dissociation upon ionization<sup>15</sup> is in fact due to the inter- and intramolecular vibrational energy deposited inside the ion cluster, which the present model does not pretend to transcribe.

A rise in total energy allows the system to explore a vaster area of the neutral potential energy surface and, following vertical ionization, of the ionic potential energy surface. This translates into a spread of the ionic cluster energy range. In some cases, the energy values bracketing this range show little to no variation upon an energy increase; this is characteristic of the fact that a local minimum (if the minimum bound does not vary) or maximum (if the maximum bound is constant) is being explored at both energies. Examples of this include the maximum brackets of the (naphthalene)<sub>2</sub> and (anthracene)<sub>3</sub> clusters. In the case of (naphthalene)<sub>2</sub>, the local maximum was identified by local optimization as a first-order saddle point of T-shaped geometry, where the centers of mass of both monomers are separated by 4.99 Å (*versus* 5.15 Å for the neutral global minimum), and whose energy of  $E_0+87$  K makes it accessible to both molecular dynamics simulations.

In most cases however, both brackets do vary significantly: the difference amounts to *ca.* 0.4 eV for the upper limit of (naphthalene)<sub>4</sub>, for example. This is evidence that the area of the ion potential energy surface under exploration is quite steep and fully justifies the need to take the classically forbidden part of the neutral wavefunction into account, as was simulated here by an increase of the total energy.

### *3.2. Ionization Potential Statistics*

Histograms of vertical ionization potential values were generated, as previously explained, for all the clusters under study. An example of such a graph may be found on Figure 2. Table 3 presents the characteristic properties of these histograms.

Cluster	Energy	Maximum (eV)	Thresholds (eV)		Width (eV)
			Lower	Upper	
(Benzene) <sub>2</sub>	E <sub>0</sub> +100 K	8.92	8.88	8.95	0.07
	E <sub>0</sub> +300 K	8.97	8.85	9.09	0.24
(Benzene) <sub>3</sub>	E <sub>0</sub> +100 K	8.82	8.71	8.93	0.21
	E <sub>0</sub> +300 K	8.87	8.66	9.04	0.38
(Benzene) <sub>4</sub>	E <sub>0</sub> +100 K	8.82	8.70	8.95	0.24
	E <sub>0</sub> +300 K	8.93	8.60	9.18	0.57
(Naphthalene) <sub>2</sub>	E <sub>0</sub> +100 K	7.80	7.79	7.82	0.03
	E <sub>0</sub> +300 K	7.80	7.78	7.82	0.04
(Naphthalene) <sub>3</sub>	E <sub>0</sub> +100 K	7.75	7.68	7.84	0.16
	E <sub>0</sub> +300 K	7.85	7.65	7.99	0.34
(Naphthalene) <sub>4</sub>	E <sub>0</sub> +100 K	7.56	7.47	7.67	0.21
	E <sub>0</sub> +300 K	7.81	7.46	8.05	0.59
(Anthracene) <sub>2</sub>	E <sub>0</sub> +100 K	7.09	7.08	7.10	0.02
	E <sub>0</sub> +300 K	7.09	7.07	7.12	0.05
(Anthracene) <sub>3</sub>	E <sub>0</sub> +100 K	7.07	7.05	7.12	0.07
	E <sub>0</sub> +300 K	7.07	7.04	7.12	0.08

**Table 3:** Characteristics of the ionization potential (IP) histograms for all clusters under study: maximum (IP value corresponding to the most frequently occurring neutral cluster structures), lower and upper thresholds (lowest and highest IP value observed during the simulation), and width (highest IP minus lowest IP).

Since the energy of the neutral cluster is supposed constant (Table 1), the upper and lower thresholds of the histograms are directly related to the corresponding brackets for the ionic cluster energies (Table 2). Hence, the trends discussed above still apply here: a rise in the total energy will globally be accompanied by an increase in the width of the histogram, but a given threshold may remain fixed if the area of the ion potential energy surface being explored features stationary points.

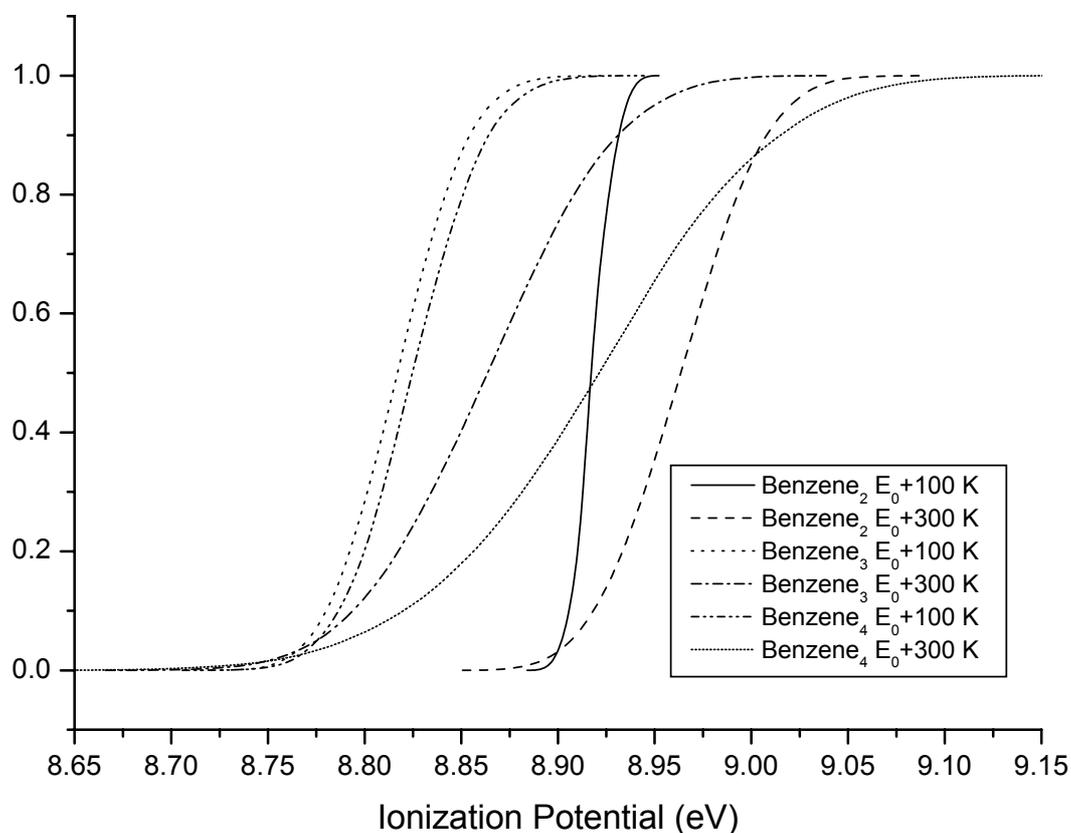
The maximum of the histogram is the centerpoint of the most frequently occurring vertical IP subdivision, and represents the vertical ionization potential, which may be theoretically computed<sup>9</sup> or inferred from experimental ion current curves, as shall be seen hereafter. However, let us mention that, since the maximum value is generally far from the threshold, direct comparison of theoretical and experimental vertical ionization potentials should be subject to caution. The lower bracket of the histogram corresponds to the onset of the ion current curve and is the theoretical counterpoint to the experimentally determined threshold IP.

Cluster configurations resulting in IP values that are close to the maximum are those that play the most important role in the ionization process. Some clusters, for example (naphthalene)<sub>3</sub> and (naphthalene)<sub>4</sub>, feature an increase of the maximum when the total energy rises: the configurations responsible for ionization tend to occur at higher energies only. It is hence essential, for such clusters, to take into account all possible configurations, including those inside the classically forbidden region (which are accessed at  $E_0+300$  K in the framework of our model). Moreover, these clusters are those for which the width of the histogram increases the most with the total energy. As previously mentioned, this is consistent with a very steep region on the ion potential energy surface, which probably corresponds to high-lying vibrationally excited states of the charged cluster; this stresses further the necessity to account for classically forbidden configurations, on which the ionization process depends.

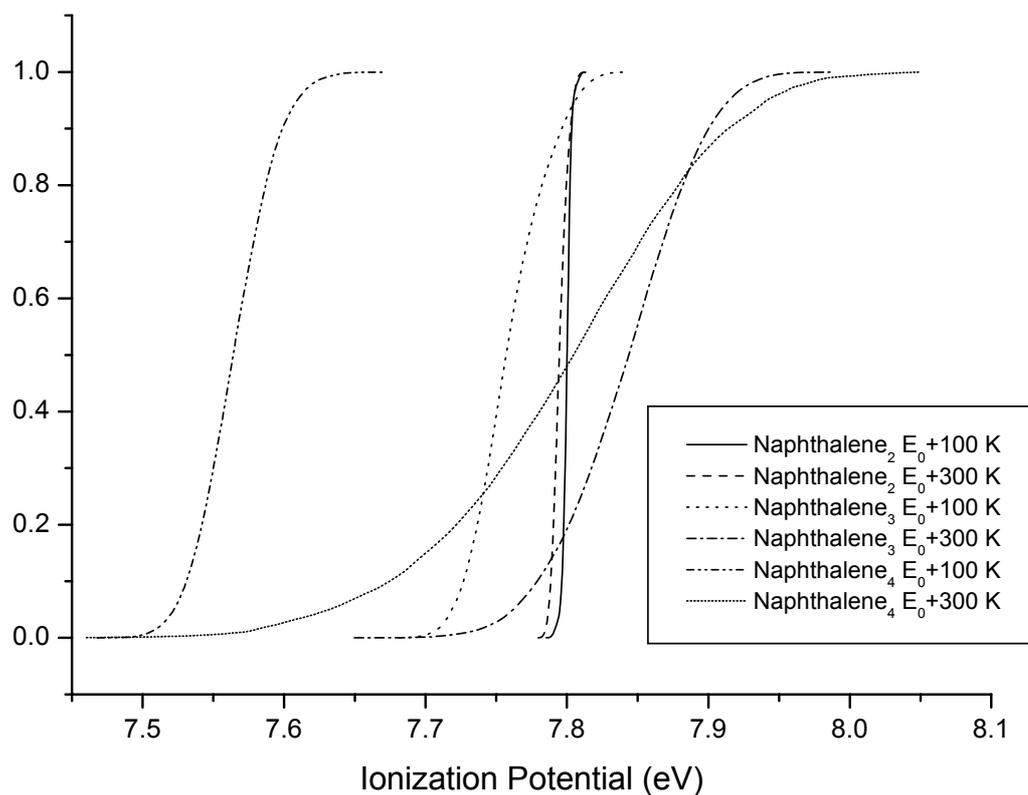
For other clusters such as (anthracene)<sub>2</sub> and (anthracene)<sub>3</sub>, on the contrary, the maximum does not evolve with the total energy. In these cases, the configurations responsible for ionization are accessible at lower energies. The correct vertical IP value may be found regardless of the energy at which the neutral dynamics was performed, *i.e.*, the configurations responsible for ionization are probably not located in the classically forbidden region of the neutral vibrational ground state wavefunction. This can be interpreted by the proximity of a local minimum on the ion potential energy surface and the global minimum on the neutral potential energy surface: the vertical transition from the neutral vibrational ground state leads to a rather low-lying vibrationally excited state of the ion, whose wavefunction overlaps that of the neutral vibrational ground state. This finding is in good agreement with our previous work<sup>9</sup>, where the large number of local minima on the potential energy surfaces of neutral and charged anthracene clusters had been mentioned.

### 3.3. Simulated Ion Current Curves

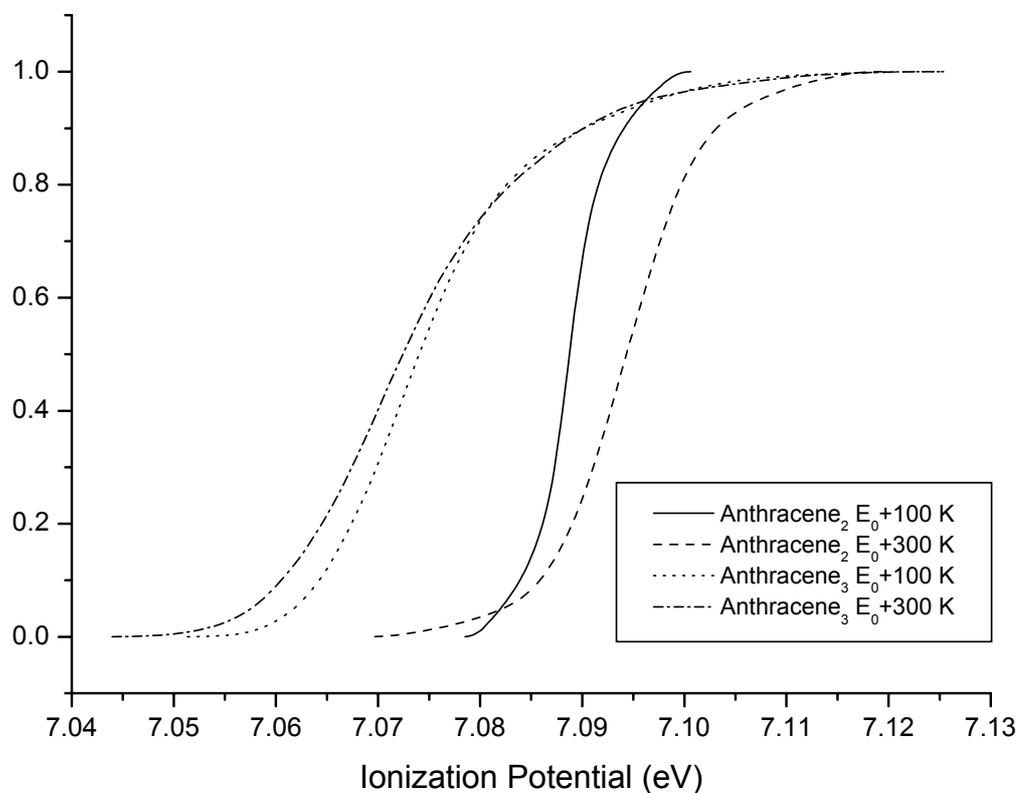
The aforementioned vertical IP histograms were integrated over their own width, to yield simulated ion current curves which will be compared to their experimental counterparts, whenever available, in the following paragraph. These curves may be found, for clusters of benzene, naphthalene and anthracene, on Figures 3, 4 and 5, respectively. We wish once again to stress the fact that, since our model neglects the intramolecular vibrationally excited states which, in reality, contribute to the experimental ion current curve, a rigorous comparison with our theoretical results is only feasible near the threshold. However, we have the feeling that interesting global trends may nevertheless be inferred from such a comparison.



**Figure 3:** Simulated ion current curves for the clusters (benzene)<sub>n</sub>, 2 ≤ n ≤ 4, for two different neutral molecular dynamics energies (refer to text).



**Figure 4:** Simulated ion current curves for the clusters (naphthalene)<sub>n</sub>, 2 ≤ n ≤ 4, for two different neutral molecular dynamics energies (refer to text).



**Figure 5:** Simulated ion current curves for the clusters (anthracene)<sub>n</sub>, 2 ≤ n ≤ 3, for two different neutral molecular dynamics energies (refer to text).

The onset of the ion current curves corresponds to the threshold ionization potentials appearing in Table 4. The steepness of the curves depends on two factors. Firstly, the area of the neutral potential energy surface that is explored during the molecular dynamics simulations dictates the area visited on the ion potential energy surface following vertical ionization. Unless the ion potential energy surface in this zone is flat, a more widespread exploration of this surface tends to generate a wider range of ion energies, hence of vertical IP values, hereby resulting in widespread and gently rising ion current curves. On the other hand, the thoroughness of the exploration depends on the nature of the neutral potential energy surface: a steep surface will limit the exploration, even at high energies, and result in a more pronounced slope for the ion current. Secondly, a steep ion potential energy surface in the region under exploration will produce a wider range of IP values, hence a gentler slope for the ion current

Cluster	Energy	Theoretical IP (eV)	Experimental threshold (eV)
(Benzene) <sub>2</sub>	E <sub>0</sub> +100 K	8.88	8.65 ± 0.01 <sup>12,15</sup> (2-photon)
	E <sub>0</sub> +300 K	8.85	< 8.865 ± 0.004 <sup>11,14</sup> (2-photon) 8.53 ± 0.03 <sup>16</sup> (2-photon) 8.69 ± 0.02 <sup>13,18</sup> (1-photon)
(Benzene) <sub>3</sub>	E <sub>0</sub> +100 K	8.71	8.58 ± 0.02 <sup>15</sup>
	E <sub>0</sub> +300 K	8.66	8.54 ± 0.03 <sup>16</sup>
(Benzene) <sub>4</sub>	E <sub>0</sub> +100 K	8.70	8.55 ± 0.02 <sup>15</sup>
	E <sub>0</sub> +300 K	8.60	
(Naphthalene) <sub>2</sub>	E <sub>0</sub> +100 K	7.79	7.58 ± 0.04 <sup>24</sup>
	E <sub>0</sub> +300 K	7.78	
(Naphthalene) <sub>3</sub>	E <sub>0</sub> +100 K	7.68	7.56 ± 0.03 <sup>24</sup>
	E <sub>0</sub> +300 K	7.65	
(Naphthalene) <sub>4</sub>	E <sub>0</sub> +100 K	7.47	7.49 ± 0.03 <sup>24</sup>
	E <sub>0</sub> +300 K	7.46	

**Table 4:** Comparison of the theoretical threshold values calculated for the clusters of benzene and naphthalene with experimental determinations from the literature.

When the total energy of the neutral molecular dynamics simulation rises, the system is allowed to explore vaster regions on both the neutral and the ionic potential energy surfaces: ion current curves at  $E_0+100$  K are generally steeper than their counterparts at  $E_0+300$  K. This is especially true for clusters such as (benzene)<sub>2</sub> or (naphthalene)<sub>2</sub>; in these cases, the structural difference between neutral and ionic clusters is such that configurations originating from the neutral molecular dynamics simulation correspond to high-lying ions located on steep areas far from any local minimum. Larger clusters are characterized by flatter potential energy surfaces featuring a large number of local minima; this results in more gentle slopes on the ion current curves, and a less marked dependence on the energy at which the neutral molecular dynamics calculation is carried out.

Finally, let us mention that the maximum value of the histogram (Table 3) corresponds to the inflexion point on the ion current curve. This far from the threshold, a comparison between inflexion points on theoretical and experimental ion current curves is probably not legitimate. However, as a tribute to our model, we note that the correspondence between the histogram maxima and the previously determined<sup>9</sup> 0 K vertical ionization potentials is quite good.

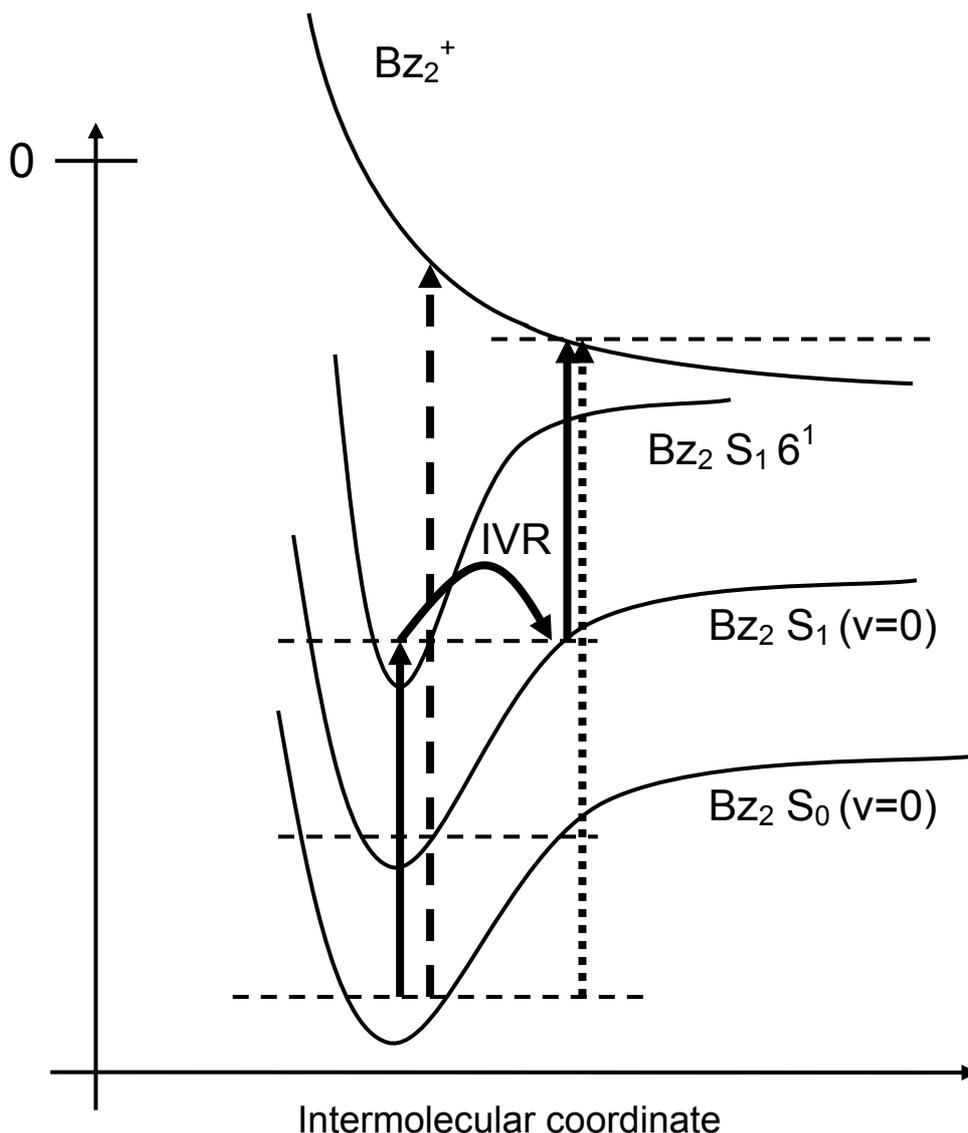
### *3.4. Comparison with Experiments*

Here, we compare the lower threshold values of the IP histograms with the corresponding experimentally determined ‘adiabatic’ ionization potentials. These are most readily available for small benzene clusters<sup>11-18</sup>; a recently published work by Fujiwara and Lim<sup>24</sup> presents IP values for naphthalene aggregates; to our knowledge, such values are not known for clusters of anthracene. Table 4 reproduces the theoretically determined threshold IPs alongside their experimental counterparts, as taken from various published works. Additionally, the simulated ion current curves on Figures 3 and 4 are qualitatively compared to their experimental counterparts.

As can be seen, a consensus has yet to be reached as far as the experimental determination of the IP of benzene<sub>2</sub> is concerned. The only real point of concordance between experimental studies is the slow rise of the ion current with photon energy, which stems from the structural difference between the neutral and ionic cluster geometries already observed in our previous study<sup>9</sup>. Depending on the experimental technique (one-photon or two-photon), different Franck-Condon factors may be invoked to explain the ionization mechanism, resulting in different extrapolated IP values. The sensitivity of the techniques also has a

determining influence on the extrapolated position of the threshold, especially in the case of two-color experiments that may suffer from a parasite 1-color signal. This explains the rather important differences in figures between references <sup>11,14</sup> on the one hand, and <sup>12,15</sup> on the other. The latter experiments, being much more sensitive since optimized for the determination of the threshold, are more reliable than the former, which should rather be considered as an upper boundary for the threshold (despite the reported sensitivity of  $4.10^{-3}$  eV). Two-photon experimental results themselves depend on the nature of the intermediate state: while most experimental setups employ the  $S_1 6^1$  state<sup>11,12,14,15</sup>, whose potential energy surface resembles that of the neutral benzene dimer, Shinohara and Nishi<sup>16</sup> use the  $S_2$  state which features a stacked geometry as in the charged dimer. Hence, their result is lower in energy, *i.e.*, closer to the theoretical adiabatic IP value (the energy of the ion global minimum minus the energy of the neutral global minimum).

From the aforementioned experimental results, we infer that two-color experiments result in lower threshold values than their one-color counterparts, and propose a qualitative explanation as to why this is the case. Intra-cluster vibrational redistribution in the intermediate  $S_1 6^1$  state produces excess intermolecular vibrational energy, which translates into a rise of the translational and rotational temperatures of the cluster. This results in a broader exploration of the potential energy surface of the  $S_1$  state (Figure 6). Assuming similarities in the respective potential energy surfaces of the  $S_0$  and  $S_1$  states, this phenomenon may be modelled by carrying out the molecular dynamics simulation on the neutral  $S_0$  cluster at a higher total energy ( $E_0+300$  K in our case) which, as previously mentioned, yields lower threshold values than at  $E_0+100$  K. The experimental difference in energy between the  $6^1$  and  $0^0$   $S_1$  states being  $750$  K<sup>25</sup>, vertical ionization statistics on the configuration subspace of the  $(\text{benzene})_2$  cluster sampled by a molecular dynamics simulation at  $E_0+750$  K could quantitatively confirm this picture. However, we were unable to carry out such a calculation because of frequent occurrences of monomer evaporation at this energy, which are due to the weakly bound character of the neutral benzene dimer.



**Figure 6:** Schematic representation of one-photon and two-photon ionization mechanisms of the benzene dimer. One-photon ionization is denoted by the dashed arrow. Two-photon ionization (full arrows) is a multistep process:  $6^1_0$  transition, intra-cluster vibrational redistribution (IVR), and ionization. The corresponding two-photon ionization potential is symbolized by the dotted arrow.

Selectivity issues appear for larger cluster sizes, limiting the number of experimental IP determinations. Our values for (benzene)<sub>3</sub> are much higher than the ones proposed by either Krause<sup>15</sup> or Nishi<sup>16</sup>; however, the theoretically determined IP shift to lower energies when adding a benzene molecule to the dimer cluster (*ca.* 0.17 eV) is in better agreement with the former work (0.07 eV) than with the latter (which finds the IPs of (benzene)<sub>2</sub> and (benzene)<sub>3</sub> to be nearly equal). Krause *et al.* find that the IP difference between (benzene)<sub>3</sub>

and (benzene)<sub>4</sub> is much smaller than between (benzene)<sub>2</sub> and (benzene)<sub>3</sub>. Our results concur, with differences of 0.01 eV and 0.06 eV at  $E_0+100$  K and  $E_0+300$  K, respectively. All values of ionization potentials are much larger than that of solid benzene, of 7.58 eV<sup>26</sup>. As previously mentioned, larger clusters are characterized by flatter potential energy surfaces, resulting in gentler slopes on the ion current curves. This trend, already remarked experimentally on clusters of benzene<sup>15</sup>, is confirmed by our simulations (Figure 3).

The ion intensity curve of the naphthalene dimer cluster is also characterized by a slow increase with photon energy near the threshold, which is experimentally attributed to a significant structural change between the sandwich-eclipsed ( $D_{2h}$ ) structure of the transient excimer<sup>27</sup> and the low-energy conformation of the dimer ion<sup>24</sup>. As remarked by Fujiwara and Lim, the sandwich-staggered structure found to be the global minimum on the ion potential energy surface in our previous study should fit the bill, resulting in an unfavorable Franck-Condon factor. The same slow increase in the photoionization yield accompanying the rise in photon energy was also noted for higher clusters of naphthalene. We agree with Lim's explanation of a charged dimer core surrounded by additional weakly-bound neutral monomers: polarization effects clearly dominate the charge resonance-polarization competition in the naphthalene trimer ion and in higher clusters<sup>9</sup>. The calculated values for the ionization potentials of the naphthalene clusters appearing in Table 4 are somewhat higher than those proposed by Fujiwara and Lim. However, the differences between figures tend to decrease as the cluster size grows: for (naphthalene)<sub>4</sub>, the agreement between theoretical and experimental results is quite good. We note that the experimental ion current curve for (naphthalene)<sub>2</sub> rises in a remarkably gentler fashion than those of (naphthalene)<sub>3</sub> or (naphthalene)<sub>4</sub><sup>24</sup>, as opposed to the theoretical trend explained above (paragraph 3.3 and Figure 4). This is evidence that the Franck-Condon overlap between the intermediate excimer and the ion cluster is more unfavorable in the dimer cluster than in either the trimer or the tetramer. In the framework of our model, both the trimer and the tetramer ion feature stacked-eclipsed dimer cores of similar geometries to that of the excimer, whereas the dimer ion global minimum is a stacked-eclipsed structure, corroborating the previous assumption. Our simple ion current model, consistent with one-photon experiments only, is of course unable to reproduce these results, since the neutral clusters explored by the molecular dynamics simulation do not feature stacked geometries.

## IV. Conclusion

Our tailored model potential has proved its efficiency for the determination of the minimum energy structures on the potential energy surfaces of small charged and neutral aromatic clusters<sup>9</sup>. In this work, it is applied to the study of the photoionization mechanisms in these species. The use of classical molecular dynamics simulations ensures a more realistic description of these phenomena, by providing a better sampling of the neutral potential energy surface at total energies chosen so as to model the quantum effects that are thought to play an important role in the process. We hope that these results will prove to be a useful addition to a subject that is still ripe with discussion and controversy, and provide incentive for experimental studies on clusters of anthracene.

The present approach is the first step toward a dynamic modeling of ionization and charge resonance phenomena in neutral and ionic aromatic clusters. The next logical step is to adopt a mixed classical–quantum mechanical molecular dynamics procedure such as a trajectory surface hopping method. In conjunction with the valence bond Hamiltonian framework, such an approach should provide a means to carry out molecular dynamics simulations on the potential energy surfaces of the ionic clusters and their electronic excited states, thus accounting for post-ionization relaxation and energy redistribution inside the aggregates. We are currently working to implement this methodology at reasonable CPU time costs.

## References

- (1) Voityuk, A. A.; Jortner, J.; Bixon, M.; Rosch, N. *J. Chem. Phys.* **2001**, *114*, 5614.
- (2) Voityuk, A. A.; Siriwong, K.; Rosch, N. *Phys. Chem. Chem. Phys.* **2001**, *3*, 5421.
- (3) Jortner, J.; Bixon, M.; Langenbacher, T.; Michel-Beyerle, M. E. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 12759.
- (4) Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112*, 5525.
- (5) Smith, G. D.; Jaffe, R. L. *J. Phys. Chem.* **1996**, *100*, 9624.
- (6) White, R. P.; Niesse, J. A.; Mayne, H. R. *J. Chem. Phys.* **1998**, *108*, 2208.
- (7) Spirko, V.; Engvist, O.; Soldan, P.; Selzle, H. L.; Schlag, E. W.; Hobza, P. *J. Chem. Phys.* **1999**, *111*, 572.
- (8) Dang, L. X. *J. Chem. Phys.* **2000**, *113*, 266.
- (9) Bouvier, B.; Brenner, V.; Millié, P.; Soudan, J. M. *J. Phys. Chem. A* **2002**, *106*, 10326.
- (10) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 18790.
- (11) Börnsen, K. O.; Selze, H. L.; Schlag, E. W. *J. Phys. Chem.* **1988**, *92*, 5482.
- (12) Ernstberger, B.; Krause, H.; Kiermeier, A.; Neusser, H. J. *J. Chem. Phys.* **1990**, *92*, 5285.
- (13) Grover, J. R.; Walters, E. A.; Hui, E. T. *J. Phys. Chem.* **1987**, *91*, 3233.
- (14) Selze, H. L.; Neusser, H. J.; Ernstberger, B.; Krause, H.; Schlag, E. W. *J. Phys. Chem.* **1989**, *93*, 7535.
- (15) Krause, H.; Ernstberger, B.; Neusser, H. J. *Chem. Phys. Lett.* **1991**, *184*, 411.
- (16) Shinohara, H.; Nishi, N. *J. Chem. Phys.* **1989**, *91*, 6743.
- (17) Rühl, E.; Bisling, P. G. F.; Brutschy, B.; Baumgärtel, H. *Chem. Phys. Lett.* **1986**, *126*, 232.
- (18) Grover, J. R.; Walters, E. A.; Baumgärtel, H. *J. Phys. Chem.* **1989**, *93*, 7534.
- (19) Lindner, R.; MullerDethlefs, K.; Wedum, E.; Haber, K.; Grant, E. R. *Science* **1996**, *271*, 1698.
- (20) Courty, A.; Mons, M.; Dimicoli, N.; Piuze, F.; Gageot, M. P.; Brenner, V.; de Pujo, P.; Millié, P. *J. Phys. Chem. A* **1998**, *102*, 6590.
- (21) Gregory, J. K.; Clary, D. C. *Mol. Phys.* **1996**, *88*, 33.
- (22) Hobza, P.; Selze, H. L.; Schlag, E. W. *Coll. Czech. Chem. Commun.* **1992**, *57*, 1186.
- (23) Svanberg, M. *Mol. Phys.* **1997**, *92*, 1085.
- (24) Fujiwara, T.; Lim, E. C. *J. Phys. Chem. A* **2003**, *107*, 4381.
- (25) Stephenson, T. A.; Radloff, P. L.; Rice, S. A. *J. Chem. Phys.* **1984**, *81*, 1060.
- (26) Asaf, U.; Steinberger, T. *Chem. Phys. Lett.* **1975**, *33*, 563.
- (27) Sadygov, R. G.; Lim, E. C. *Chem. Phys. Lett.* **1994**, *225*, 441.

### 3. Mécanismes de l'Ionisation dans les Agrégats Aromatiques



## 4. Conclusion

La modélisation du transfert de charge présentée dans cette partie s'est révélée très efficace dans le cas notoirement délicat des agrégats chargés de molécules aromatiques. La combinaison d'un potentiel modèle de bonne qualité et d'un hamiltonien effectif combine en effet plusieurs avantages. La contribution du transfert de charge à l'énergie d'interaction est traitée comme un terme à  $n$  corps, mais par le biais d'éléments de matrice couplant les formes *valence bond* deux à deux et paramétrés sur des dimères. Cette contribution est clairement séparée des autres termes de l'énergie d'interaction (ce qui facilite sa paramétrisation) ; pour autant, la compétition entre termes est correctement retranscrite et se montre essentielle à la bonne compréhension des propriétés structurales et électroniques des agrégats étudiés. En outre, le potentiel est suffisamment peu coûteux pour être asservi à des études de dynamiques moléculaires, plus réalistes qu'une simple exploration des surfaces d'énergie potentielle lorsque la comparaison aux résultats expérimentaux est envisagée. Une telle comparaison a d'ailleurs mis en évidence la bonne qualité du modèle, et l'intérêt qu'une telle étude peut revêtir pour l'interprétation de résultats expérimentaux qui, comme c'est le cas pour les potentiels d'ionisation d'agrégats aromatiques, est souvent délicate.

Un certain nombre d'extensions et de perfectionnements du modèle sont néanmoins envisageables. Il conviendrait, dans un premier temps, de tester le modèle dans le cas d'agrégats hétéroatomiques ; étant données les variations observées entre systèmes en ce qui concerne le coefficient de proportionnalité reliant éléments de couplage et intégrales de recouvrement, il est fort peu probable qu'un seul coefficient puisse rendre compte de manière précise du transfert de la charge entre les diverses combinaisons de monomères possibles dans un tel agrégat. Dans le même ordre d'idée, il a été montré sur le cas de la pyridine qu'un unique coefficient de proportionnalité est incapable de retranscrire la coexistence de plusieurs interactions à transfert de charge au sein d'un même système. Enfin, la non-linéarité du modèle aux faibles recouvrements, bien que de peu de conséquence, est conceptuellement peu plaisante. Une réflexion méthodologique incluant ces problèmes se révélerait donc très fructueuse.

L'extension de l'étude des aspects dynamiques du phénomène de transfert de charge pourrait s'effectuer à l'aide d'une méthode de dynamique moléculaire semi-classique,

## Deuxième Partie – Le Transfert de Charge

comme la méthode TSH présentée en première partie et appliquée en troisième partie au transfert d'excitation. Malheureusement, ce type de calcul se révèle rapidement coûteux lorsque la taille des systèmes ou les temps de simulation augmentent ; cet aspect a été vérifié en troisième partie dans le cas du transfert d'excitation, avec un potentiel pourtant considérablement plus simple que celui employé ici. Il serait donc sans doute utile de simplifier le calcul des éléments de la matrice hamiltonienne. La simplification du potentiel modèle employé pour les énergies des formes *valence bond* est actuellement à l'étude au laboratoire, dans le cas d'agrégats ion-molécules d'eau. La simplification du calcul des termes non diagonaux pourrait, quant à elle, passer par le précalcul et l'interpolation des valeurs des intégrales de recouvrement. L'investissement que représente un tel travail devrait être largement compensé par la variété des études qu'il devrait rendre possible, sur des systèmes aussi divers que les polymères d'intérêt biologique ou la solvation des métaux lourds.

## 4. Conclusion



*Troisième Partie*

LE TRANSFERT  
D'EXCITATION



*De cela découle une extraordinaire question. Les Égyptiens connaissaient-ils  
l'électricité ?*

Peter Kolosimo, *Terra senza tempo* (1964)



## 1. Les Acides Nucléiques

## 1. Les Acides Nucléiques

### 1.1. Présentation des acides nucléiques

#### *1.1.1. Introduction*

**L**es acides nucléiques constituent une classe de macromolécules biologiques dont la fonction est d'enchâsser l'information génétique, sous la forme d'un enchaînement d'unités répétitives constituant les lettres d'un alphabet. La simplicité et la robustesse de ce code, qui s'appuie sur un nombre très restreint de combinaisons structurales et fonctionnelles, est d'autant plus merveilleuse que les acides nucléiques, vecteurs universels de l'information génétique, expriment à travers chaque être végétal ou animal une incroyable diversité.

Les éléments structuraux des acides nucléiques, ainsi que leurs implications sur le stockage de l'information génétique, sont détaillés dans cette partie.

#### *1.1.2. Les bases azotées*

Les bases azotées standards qui constituent l'alphabet génétique sont au nombre de cinq. L'adénine (A) et la guanine (G) ont en commun d'être construites à partir d'un résidu indole et sont dénommées purines. La cytosine (C), l'uracile (U) et la thymine (T) sont des diazines substituées et sont en conséquence regroupées sous l'appellation de pyrimidines. La figure 1.1 présente les formules développées de ces molécules, ainsi que la numérotation universelle de leurs atomes constitutifs.

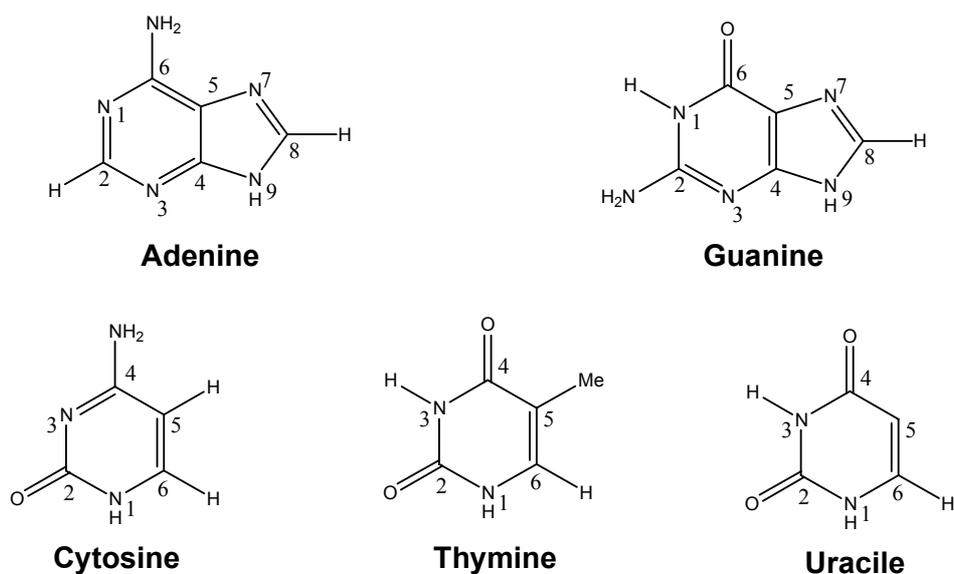


Figure 1.1

L'uracile, présente dans les acides ribonucléiques, est remplacée par la thymine dans les acides désoxyribonucléiques, dont elle ne diffère que par un substituant méthyle sur le carbone 5.

### 1.1.3. Nucléosides, nucléotides et brins

Au sein des acides nucléiques, les bases azotées sont assemblées au moyen d'un squelette composé de sucres pentacycliques et de substituants phosphate. Le sucre est le ribose dans l'ARN et le 2-désoxyribose dans l'ADN. Il substitue l'azote 9 des purines et l'azote 1 des pyrimidines ; l'ensemble constitué par la base et le sucre est dénommé nucléoside. La position 5' du sucre peut être substituée par un phosphate organique. L'édifice ainsi obtenu porte le nom de nucléotide (Figure 1.2).

## 1. Les Acides Nucléiques

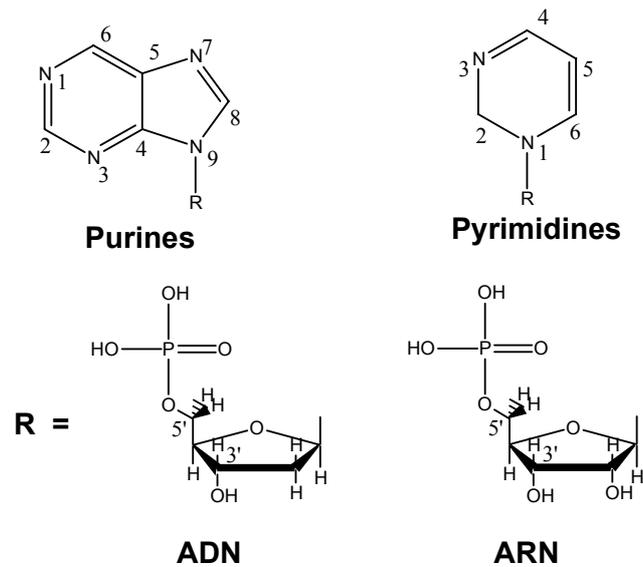


Figure 1.2

L'assemblage polymérique de nucléotides par le biais d'un pont ester phosphate entre le carbone 5' d'un nucléotide et le carbone 3' d'un autre génère un brin. La séquence de ce brin, vecteur de l'information génétique, est constituée de l'ensemble ordonné des bases le constituant, lues dans l'ordre 5' vers 3' (Figure 1.3).

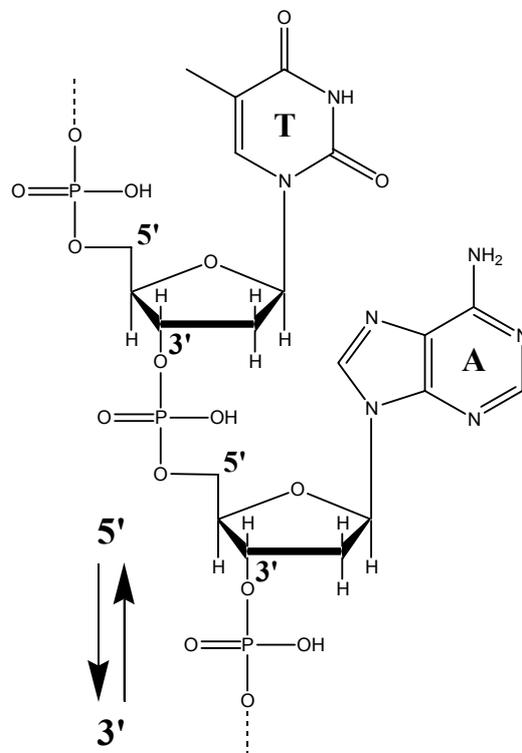


Figure 1.3

### 1.1.4. Appariement Watson & Crick

Les acides nucléiques peuvent former des chaînes extraordinairement longues et adoptent, tout comme les protéines, des structures secondaires et tertiaires. En 1953, Watson et Crick<sup>30</sup> proposèrent comme structure secondaire de l'ADN une double hélice formée de deux brins de séquences complémentaires. S'appuyant sur le fait que les rapports adénine/thymine et guanine/cytosine expérimentalement constatés valent un, ils conjecturèrent une complémentarité de ces couples de bases dans l'ADN (Figure 1.4). Les bases complémentaires sont assujetties par liaisons hydrogène (deux pour le couple AT, trois pour le couple GC). L'arrangement optimal d'un double brin, compte tenu des liaisons hydrogène et des autres contraintes structurales, est la double hélice.

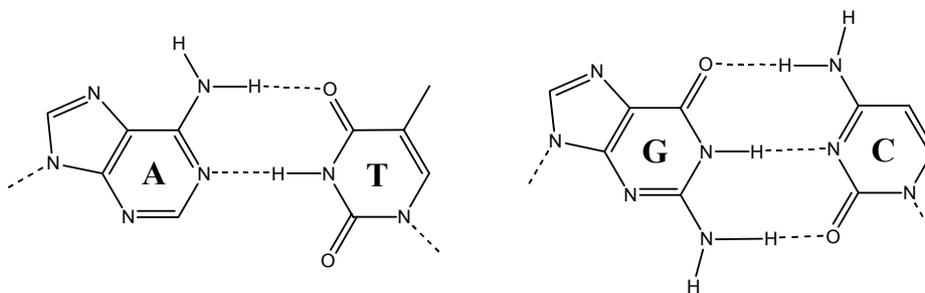
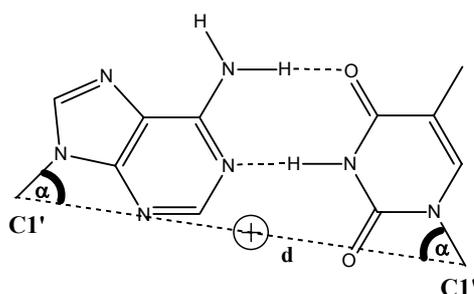


Figure 1.4

Chacun des deux appariements Watson & Crick possibles peut être construit en positionnant les bases de telle sorte que<sup>222,223</sup> (Figure 1.5):

- La distance entre atomes C1' inter-brins soit de 10,6 Å ;
- Il y ait un angle de 56° entre la liaison glycosyle C1'-N et la droite joignant deux atomes C1' opposés ;
- L'axe de rotation pseudo-C<sub>2</sub> de la double hélice mettant en relation les deux bases d'une même paire soit perpendiculaire au segment C1'-C1' en son milieu.



$$\alpha = 56^\circ$$

$$d = 10,6 \text{ \AA}$$

⊕ : axe pseudo-C<sub>2</sub>

Figure 1.5

## 1. Les Acides Nucléiques

### 1.1.5. La double hélice

La double hélice est la structure préférentielle des acides désoxyribonucléiques. D'autres types existent cependant : simples, triples et quadruples hélices. Les paramètres définissant la géométrie de la double hélice sont<sup>224</sup> :

- Le nombre  $n$  de résidus (paires de bases) par tour d'hélice ;
- La translation  $h$  par résidu le long de l'axe de l'hélice ;
- L'angle de rotation  $t$  par résidu autour de l'axe de l'hélice,  $t = 2\pi/n$  ;
- La périodicité  $P$ ,  $P = n.h$ .

Les types de chaînes les plus courants, qui diffèrent par l'orientation de l'hélice et les valeurs des paramètres ci-dessus, sont présentés dans le tableau 1.1 :

	<b>ADN A</b>	<b>ADN B</b>	<b>ADN B'</b>	<b>ADN Z</b>
<b>Sens de l'hélice</b>	Droite	Droite	Droite	Gauche
<b>n</b>	11	10	10	12
<b>t (°)</b>	32,7	36,0	34,1 / 36,8	-10 / -50
<b>h (Å)</b>	2,9	3,4	3,51 / 3,30	3,7
<b>P (Å)</b>	32	34	34	45

Tableau 1.1

De faibles déviations à ces structures d'hélice dites canoniques sont susceptibles d'apparaître en fonction de la séquence. Parmi les quatre hélices présentées, trois sont droites (A, B, B') et une gauche (Z). L'ADN B est le plus représenté dans la nature, suivi par l'ADN A. L'ADN B' est dit alternant ; ses valeurs de  $t$  et de  $h$  diffèrent suivant la parité du résidu au sein de la chaîne ; il en va de même pour l'ADN Z. La figure 1.6 présente une aperçu global de structures A, B et Z canoniques de brins d'égales longueur (quarante paires de bases).

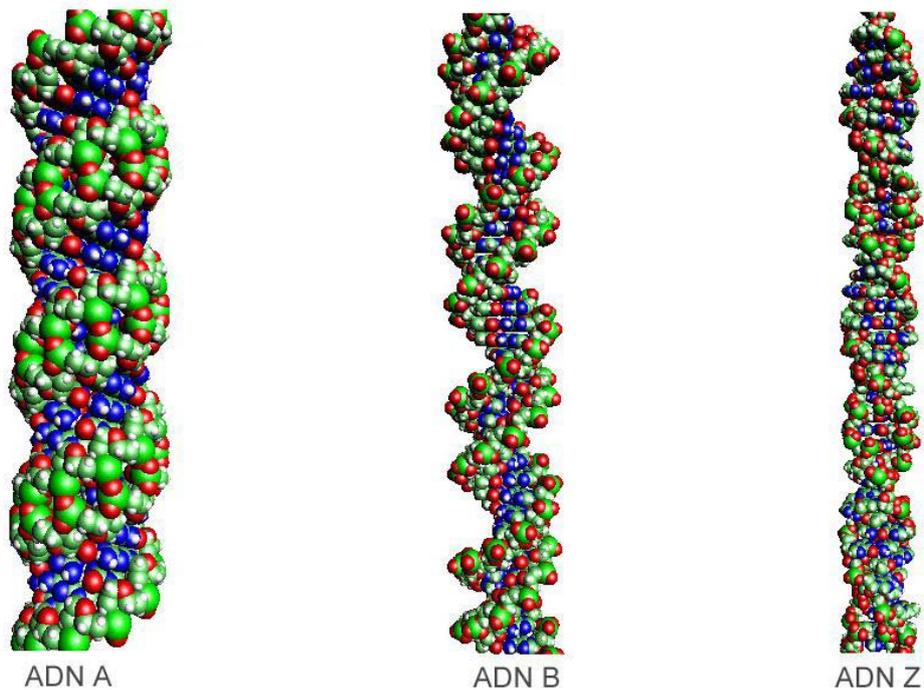


Figure 1.6

## 1.2. Les doubles brins poly-dA et poly-(dAdT)

Le présent travail est focalisé sur l'étude de deux fragments d'ADN : les fragments poly-dA et poly-(dAdT), dont les caractéristiques géométriques sont détaillées ci-dessous.

### *1.2.1. Le double brin poly-dA*

Le double brin poly-dA porte la séquence 5'-dA-dA-...-dA-dA-3', le brin complémentaire portant quant à lui la séquence 3'-dT-dT-...-dT-dT-5'. Sa topologie fut déterminée par diffraction des rayons X<sup>225</sup>. Il convient de noter qu'une telle séquence peut, dans certaines conditions, engendrer des structures à triple voire quadruple brins.

La structure de la double hélice est proche de celle de l'ADN B canonique : dix paires de bases par tour d'hélice ( $n = 10$  soit  $t = 36^\circ$ ) et une translation axiale  $h = 3,29 \text{ \AA}$  (soit  $P = 32,9 \text{ \AA}$ , ce qui représente une compression de 4 % par rapport à la structure canonique de

## 1. Les Acides Nucléiques

l'ADN B). Le double brin  $(dA)_{10}(dT)_{10}$ , qui comprend dix paires de bases, est représenté en figure 1.7.

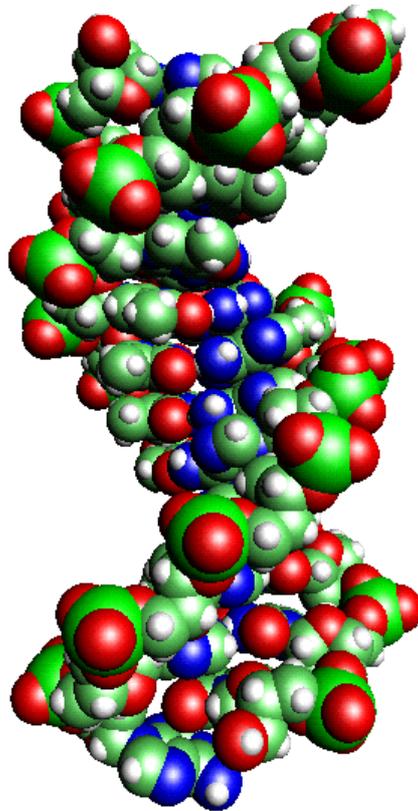


Figure 1.7

### 1.2.2. Le double brin poly-(dAdT)

Le double brin poly-(dAdT) porte la séquence  $5' \dots dA-dT \dots 3'$ , le brin complémentaire la séquence  $3' \dots dT-dA \dots 5'$ . La structure du brin fut extrapolée à partir de celle d'un oligomère comportant quatre paires de bases, obtenue par diffraction des rayons X<sup>226</sup>.

La structure est proche de celle de l'ADN B' canonique. Chaque tour d'hélice accomode dix paires de bases ( $n = 10$ ). L'angle de rotation entre paires de bases consécutives  $t$  vaut  $31^\circ$  entre A et T et  $41^\circ$  entre T et A, soit une valeur moyenne de  $36^\circ$  compatible avec le nombre de paires de bases par tour. La translation  $h$  vaut quant à elle  $3,34 \text{ \AA}$  pour tous les

résidus. Le double brin  $(dAdT)_5(dTdA)_5$ , comprenant dix paires de bases, est représenté en figure 1.8.

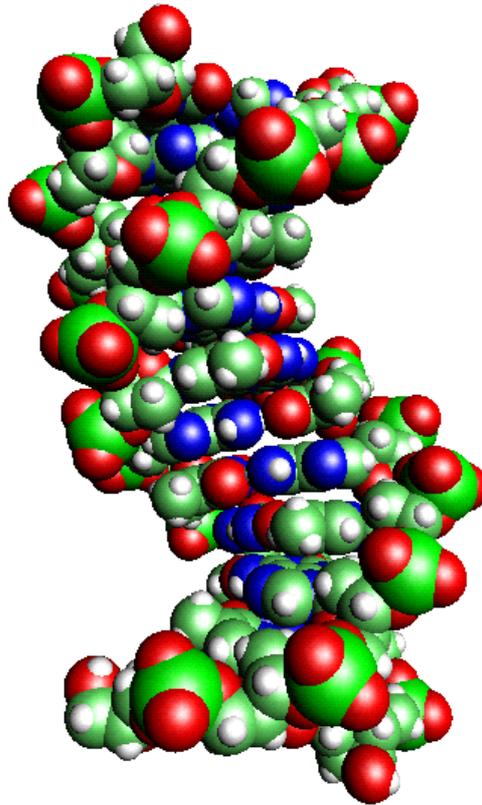


Figure 1.8

### 1.3. Tautomérie des bases azotées

Les bases azotées, bien que classiquement représentées au moyen d'une formule développée unique et bien définie, présentent en réalité un nombre de formes dites tautomères en équilibre thermodynamique. Leur existence est liée aux diverses formes que peuvent adopter leurs groupements fonctionnels (formes cétoniques ou énoliques, amino ou imino...) ainsi qu'à l'aromaticité des cycles les constituant. Le poids statistique des diverses formes tautomères est susceptible de varier en fonction de l'environnement des bases (solvant, pH...). Ils sont également fonction du degré de substitution des bases (base isolée, base dans l'ADN...).

## 1. Les Acides Nucléiques

La population des divers tautomères est asservie à une statistique de Boltzmann qui dépend de l'énergie de chaque forme. Cette énergie est elle-même reliée à l'énergie libre associée à la réaction de tautomérisation :

$$\Delta G = \Delta H - T\Delta S$$

De manière générale et en phase gazeuse,  $\Delta S$  est faible devant  $\Delta H$ , qui est lui-même voisin de la différence entre les énergies électroniques des tautomères considérés. En conséquence, la difficulté de déterminer les proportions relatives des différentes formes tient au fait qu'il est malaisé de prédire la stabilité relative des formes tautomères isolées. Expérimentalement, l'étude d'acides nucléiques en phase gazeuse est rendue délicate par le point de fusion élevé, la pression de vapeur saturante faible et la susceptibilité à la décomposition thermique de ces molécules. Elles ne peuvent pas non plus être étudiées en solvant inerte, n'étant pas solubles dans ce type de solvant. Ces molécules sont donc le plus souvent étudiées en matrice inerte (Ar, N<sub>2</sub>) ou en solvant polaire. La chimie théorique, qui est à même de traiter les formes tautomères de manière indépendante, se révèle ici un précieux outil d'analyse.

L'étude de la stabilité relative des formes tautomères des bases de l'ADN étant un problème expérimental et théorique complexe, le présent travail se limitera à détailler celles de l'adénine et de la thymine, qui composent les fragments d'ADN étudiés (*cf.* 2.). Il est néanmoins indispensable de quantifier l'erreur induite, dans cette étude, par la description des bases de l'ADN à l'aide d'une unique structure chimique.

### 1.3.1. Formes tautomères de la thymine

La thymine comporte cinq formes tautomères principales en équilibre thermodynamique (Fig. 1.6):

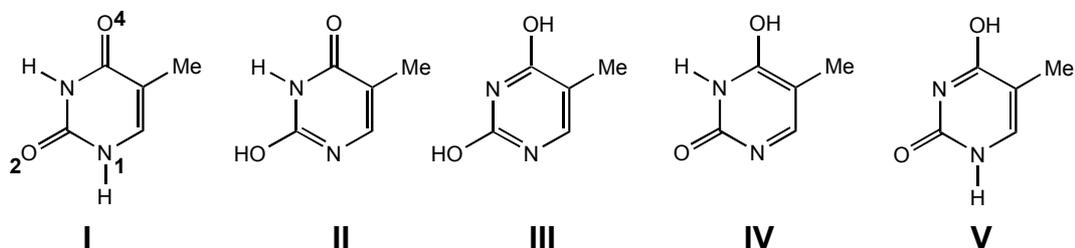


Figure 1.9

Ces formes tautomères ont deux origines : l'équilibre céto-énolique des deux groupes carbonyle en position 2 et 4 d'une part, la labilité de l'atome d'hydrogène en position 1 d'autre part. Elles sont classifiées en trois catégories : dicéto (I), céto-énol (II, IV et V) et diénol (III). Leurs énergies, calculées à différents niveaux de théorie, sont résumées dans le tableau 1.2, avec pour référence des énergies celle de la forme I.

<b>Energie (<math>\Delta_f H</math>) (kcal/mol)</b>	<b>I</b>	<b>II</b>	<b>III</b>	<b>IV</b>	<b>V</b>
<b>T isolée</b> <b>AM1</b> <sup>227</sup>	0	9,13	14,65	/	12,73
<b>T isolée</b> <b>AM1</b> <sup>228</sup>	0	9,12	14,88	/	13,47
<b>T phase aqueuse</b> <b>AM1</b> <sup>228</sup>	0	11,08	17,10	13,05	/
<b>T isolée</b> <b>MP2, 6-31G**/3-21G</b> <sup>229</sup>	0	/	/	14,81	/
<b>T isolée</b> <b>AM1</b> <sup>230</sup>	0	8,98	18,89	13,13	17,51
<b>T isolée</b> <b>DFT/MRCI, BHLYP</b> <sup>231</sup>	0	10,83	11,98	21,89	12,44

Tableau 1.2

Les diverses méthodes employées s'accordent pour situer la seconde forme tautomère (II) environ 10 kcal/mol plus haut en énergie que la forme I. Ceci correspond, à 298 K, à une population relative II/I de  $5 \cdot 10^{-8}$ . Il est donc raisonnable de considérer que la seule forme tautomère de la thymine dans le vide est la forme I. La phase aqueuse ne déplace que peu l'équilibre ; les proportions relatives expérimentales (normalisées à un pour la forme I) sont présentées dans le tableau 1.3 ; les formes II et V n'ont pas été mises en évidence dans ces études.

<b>Proportions relatives</b>	<b>I</b>	<b>III</b>	<b>IV</b>
<b>Phase aqueuse</b> <sup>232</sup>	1	$2,6 \cdot 10^{-4}$	$1,0 \cdot 10^{-5}$
<b>Phase aqueuse</b> <sup>233</sup>	1	/	$< 1,0 \cdot 10^{-2}$

Tableau 1.3

## 1. Les Acides Nucléiques

La thymine sous forme de désoxythymidine, telle qu'elle intervient dans un double brin d'ADN, est substituée en position 1 par le désoxyribose de la chaîne latérale. Les formes II, III et IV n'existent en conséquence pas dans ce système. L'approximation consistant à substituer à la base sa forme tautomère majoritaire, si elle est valide pour la base *per se*, l'est encore davantage dans l'ADN.

### 1.3.2. Formes tautomères de l'adénine

L'adénine présente quatre formes tautomères principales en équilibre thermodynamique (Figure 1.7):

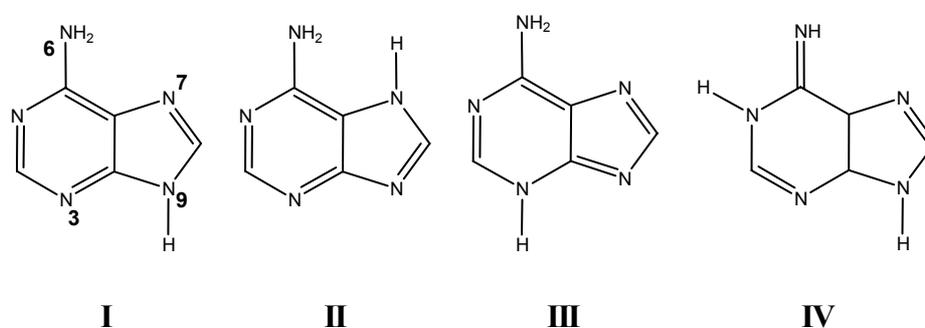


Figure 1.10

L'atome d'hydrogène labile porté par l'azote 9 dans la forme I peut se retrouver en position 7 (tautomère II) ou 3 (tautomère III). L'équilibre amino-imino du groupement NH<sub>2</sub> en position 6 engendre la forme IV. Il convient de noter que d'autres tautomères de type imino sont envisageables : par exemple, les tautomères 7H-imino ou 3H-imino. Ceux-ci sont ultraminoritaires et ont été délibérément omis. Les énergies des formes tautomères, calculées à divers niveaux de théorie, sont compilées dans le tableau 1.4, en référence à l'énergie de la forme majoritaire I.

<b>Energie (<math>\Delta_f H</math>) (kcal/mol)</b>	<b>I</b>	<b>II</b>	<b>III</b>	<b>IV</b>
<b>A isolée</b> <b>AM1<sup>228</sup></b>	0	7,01	11,58	16,00
<b>A phase aqueuse</b> <b>AM1<sup>228</sup></b>	0	2,73	9,96	14,47
<b>A isolée</b> <b>AM1<sup>230</sup></b>	0	6,68	11,52	13,59
<b>A isolée</b> <b>MP4SDQ / 6-31G*<sup>234</sup></b>	0	8,20	/	/
<b>A phase aqueuse</b> <b>MP4SDQ / 6-31G*<sup>234</sup></b>	0	1,91	/	/
<b>A isolée</b> <b>MCQDPT2/MCSCF<sup>235</sup></b>	0	8,0	8,7	/

Tableau 1.4

La seule forme tautomère dont l'énergie est suffisamment proche de celle de la forme majoritaire I pour ne pouvoir être négligée est la 7H-adénine (II). Distant de d'environ 7 kcal/mol *in vacuo*, elle est nettement plus stabilisée que la forme I en solution aqueuse de par la valeur plus importante de son moment dipolaire permanent. Ceci a pour effet de restreindre l'écart énergétique entre les deux formes. Les proportions relatives expérimentales fournies dans le tableau 1.5 confirment la coexistence des deux formes en solution aqueuse.

<b>Proportions relatives</b>	<b>I</b>	<b>II</b>
<b>Phase aqueuse<sup>236</sup></b>	1	0,28

Tableau 1.5

Cependant, la substitution de l'atome 9 par le désoxyribose dans l'adénosine rend les formes tautomères II et III inexistantes. La forme imino IV, bien que plausible, est minoritaire (le rapport IV/I vaut  $2,5 \cdot 10^{-5}$ <sup>237</sup>). Il est donc raisonnable de représenter l'adénine par sa forme tautomère majoritaire I à la condition que la position N<sub>9</sub> soit substituée, ce qui

## 1. Les Acides Nucléiques

est le cas dans l'ADN. Dans le cas contraire, deux formes tautomères (9H et 7H, I et II) devront être considérées, particulièrement en solution.



## 2. Description des États Électroniques Excités de Fragments d'ADN

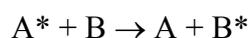
**L**a description des états électroniques excités de fragments d'ADN par des méthodes *ab initio* est, aujourd'hui encore, quasiment inenvisageable. En effet, un double brin de dix paires de bases compte déjà non loin de 800 atomes. De surcroît, la plasticité de la double hélice lui fait adopter, aux températures usuelles, une assez large gamme de configurations spatiales au voisinage de sa géométrie canonique. Le désordre ainsi introduit pouvant affecter le caractère des états excités du brin, il est prudent de le prendre en considération au niveau théorique, ce qui a pour effet de multiplier le nombre de calculs à réaliser. La prise en compte du désordre est d'autant plus importante que la comparaison à des résultats expérimentaux est envisagée. Il apparaît donc essentiel de pouvoir simplifier la description des états électroniques excités des acides nucléiques. La présente étude propose de recourir à cet effet à une méthode à hamiltonien effectif connue sous le nom de théorie excitonique, qui est présentée dans cette partie. Les états excités d'un brin y sont décrits par le biais des états électroniques des bases le constituant, plus facilement accessibles aux calculs de chimie quantique. Leur caractérisation est également abordée dans ces pages.

### 2.1. Théorie excitonique

#### *2.1.1. La matrice excitonique*

La théorie excitonique fut initialement introduite, dans le domaine de la physique des solides, afin de rendre compte de la délocalisation de l'excitation dans des cristaux par un processus de saut entre sites atomiques ou moléculaires<sup>146,238,239</sup>.

Le transfert de l'excitation électronique entre deux molécules A et B s'écrit conceptuellement :



L'excitation peut être représentée quantiquement par un ou plusieurs électrons excités vers des orbitales virtuelles et un nombre égal de trous dans des orbitales qui étaient

occupées à l'état fondamental, l'ensemble étant dénommé exciton. L'exciton de Frenkel, employé dans cette étude, correspond au cas limite où les électrons et les trous sont localisés sur la même entité chromophorique et revient à négliger l'influence des états excités à transfert de charge intermoléculaire.

L'hamiltonien électronique exact d'un système comportant  $N$  chromophores peut s'écrire sous la forme d'une somme d'hamiltoniens électroniques exacts de chacun des monomères et d'un hamiltonien d'interaction, qui couple le mouvement des électrons du monomère  $m$  à ceux des électrons et noyaux du monomère  $n$ :

$$\hat{H} = \sum_{m=1}^N \hat{H}_m + \sum_{n < m} \hat{V}_{mn}$$

L'hamiltonien d'interaction est responsable du transfert de l'excitation entre les monomères de l'agrégat. Soit  $|\psi_m^a\rangle$  la fonction d'onde de l'état adiabatique  $a$  de la molécule  $m$ , telle qu'elle est définie par l'équation de Schrödinger stationnaire pour cette molécule :

$$\hat{H}_m \psi_m^a = \varepsilon_m^a \psi_m^a$$

De manière générale, les états appartenant à des molécules distinctes ne sont pas mutuellement orthogonaux. De plus, l'hamiltonien électronique  $\hat{H}_m$  et ses fonctions et valeurs propres dépendent *a priori* de l'ensemble des coordonnées nucléaires.

Il faut à présent se doter d'une base pour la description des états électroniques de l'agrégat. Elle est constituée des fonctions d'onde antisymétriques construites sur le produit des fonctions d'onde adiabatiques des monomères :

$$\Phi_{\{A\}} = \eta \sum_p (-1)^p \mathfrak{S} \left( \prod_{m=1}^N \psi_m^{a_m} \right)$$

où  $\{A\} \equiv \{a_1, a_2, \dots, a_N\}$  est l'ensemble des nombres quantiques électroniques des monomères qui définit la configuration électronique de l'agrégat,  $\eta$  est un facteur de normalisation et  $\mathfrak{S}$  une permutation des coordonnées électroniques des molécules de l'agrégat. La matrice de l'hamiltonien  $\hat{H}$  dans la base des états  $\Phi_{\{A\}}$  est appelée matrice excitonique.

Soit un agrégat constitué de deux monomères. Soient également deux fonctions de base quelconques des états électroniques de cet agrégat,  $\Phi_{\{A\}}$  et  $\Phi_{\{B\}}$ , définies par  $\{A\} = \{a_1, a_2\}$  et  $\{B\} = \{b_1, b_2\}$ . L'élément de matrice de l'hamiltonien d'interaction correspondant s'écrit alors comme la différence d'un terme coulombien  $J_{12}$ , décrivant

## 2. Description des États Électroniques Excités de Fragments d'ADN

l'interaction classique entre densités de charge associées aux configurations électroniques des monomères, et d'un terme d'échange  $K_{12}$ , sans équivalent classique :

$$\left\langle \Phi_{\{a_1 a_2\}} \left| \hat{V}_{12} \right| \Phi_{\{b_1 b_2\}} \right\rangle = J_{12}(a_1, a_2, b_2, b_1) - K_{12}(a_1, a_2, b_2, b_1)$$

Le terme d'échange, qui contient en général plusieurs intégrales électroniques, fait intervenir des recouvrements entre les orbitales des monomères qui échangent des électrons et décroît donc exponentiellement avec la distance. Compte tenu des séparations spatiales typiques entre molécules, il sera en général bien plus petit que, et donc négligeable devant, le terme coulombien. Ceci n'est cependant vrai que dans le cas d'états excités singulets, le terme d'échange jouant un rôle important dans le transfert entre états triplets. Dans ce type de cas, les intégrales coulombiennes intervenant dans les éléments non diagonaux de la matrice sont nulles pour des raisons de spin, le couplage étant en conséquence gouverné par les intégrales d'échange. Le transfert d'excitation à l'état singulet (pour peu qu'il soit permis) est donc correctement décrit par un couplage dipolaire fort et à longue distance, tandis que le transfert d'excitation à l'état triplet est caractérisé par un couplage faible et à courte distance. Le reste de ce propos traitera exclusivement du transfert d'excitation entre états singulets dans le cadre de l'approximation formulée ci-dessus (voir par exemple la référence<sup>240</sup> pour une discussion des termes de couplage à courte portée).

Le terme de Coulomb est donc la seule contribution à l'élément de matrice de l'hamiltonien d'interaction dans le cadre de cette étude. Il peut s'exprimer, de manière très générale :

$$J_{mn}(a, b, c, d) = \int \psi_m^a(\mathbf{r}_m) \psi_n^b(\mathbf{r}_n) \hat{V}_{mn}(\mathbf{r}_m, \mathbf{r}_n) \psi_m^d(\mathbf{r}_m) \psi_n^c(\mathbf{r}_n) d\mathbf{r}_m d\mathbf{r}_n$$

Supposant une nouvelle fois le cas d'un agrégat bimoléculaire où chaque molécule peut adopter deux états quantiques électroniques (fondamental ou excité, 0 ou 1), les différents types de termes coulombiens et leur signification physique sont explicités dans le tableau 2.1 :

Type	Élément de matrice	Interaction
<b>I</b>	$J_{mn}(0,0,0,0)$	Entre densités permanentes sur les molécules $m$ et $n$
	$J_{mn}(0,1,1,0)$	
	$J_{mn}(1,0,0,1)$	
	$J_{mn}(1,1,1,1)$	
<b>II</b>	$J_{mn}(0,0,0,1)$	Entre densités de transition sur $m$ et densités permanentes sur $n$
	$J_{mn}(1,0,0,0)$	
	$J_{mn}(0,1,1,1)$	
	$J_{mn}(1,1,1,0)$	
<b>III</b>	$J_{mn}(0,1,0,1)$	Entre densités de transition de sens opposés sur $m$ et $n$
	$J_{mn}(1,0,1,0)$	
<b>IV</b>	$J_{mn}(0,0,1,1)$	Entre densités de transition de même sens sur $m$ et $n$
	$J_{mn}(1,1,0,0)$	

Tableau 2.1

Les termes de type (I) représentent les interactions entre densités de charge permanentes sur chacune des deux molécules, à l'état fondamental ou excité. Les termes de type (II) décrivent l'interaction entre une densité de charge permanente sur un monomère et une densité de charge de transition sur l'autre. Les termes de type (III) et (IV) correspondent à l'interaction de densités de charge de transition sur chacune des deux molécules. Les contributions (III) sont celles qui assurent à proprement parler la délocalisation de l'exciton de Frenkel, puisqu'elles correspondent à l'excitation d'un des monomères et à la désexcitation simultanée de l'autre. Au contraire, les interactions de type (IV) font apparaître des transitions simultanées et de même sens (excitation ou désexcitation) sur chacun des deux chromophores. Il convient de noter que les termes de type (II) et (IV), dits *off-resonance*, ne participent que peu lors d'un traitement perturbatif et seront négligés dans la suite de ce propos.

La présente étude ajoute trois approximations au modèle. En premier lieu, les molécules constituant l'agrégat sont supposées partager la même géométrie figée, qu'elles soient isolées ou au sein de l'agrégat. En conséquence, la dépendance formelle de l'hamiltonien  $\hat{H}_m$  vis-à-vis des coordonnées des noyaux au sein du monomère  $m$  sera négligée. En second lieu, l'étude adopte le formalisme de l'exciton simple, qui implique que

## 2. Description des États Électroniques Excités de Fragments d'ADN

les fonctions de base  $\Phi_{\{a\}}$  décrivent les états électroniques d'agrégats dans lesquels seul un monomère est excité, les autres se trouvant dans leur état fondamental. Ceci a pour effet d'annuler les termes faisant intervenir des excitations simultanées sur plus d'une molécule, comme  $J_{mn}(1,1,1,1)$  (les termes de type (II) et (IV) ayant déjà été négligés). Enfin, la base des fonctions d'onde  $\Phi_{\{a\}}$  est supposée orthonormée, ce qui revient à négliger l'interaction d'échange et est raisonnable dans le cas d'états singulets.

Dans ce cadre, les états électroniques de base de l'agrégat monoexcité sont construits à partir des états adiabatiques des monomères à l'aide d'un produit de Hartree:

$$\Phi_{\{0_1, 0_2, \dots, a_m, \dots, 0_N\}} = \Phi_m^a = \Psi_m^a \prod_{\substack{n=1 \\ n \neq m}}^N \Psi_n^0$$

Le calcul des éléments diagonaux et non diagonaux de la matrice excitonique peut maintenant être présenté.

### 2.1.2. Calcul des éléments diagonaux

L'élément diagonal de la matrice excitonique associé à la fonction de base  $\Phi_m^a$  s'exprime :

$$\begin{aligned} \langle \Phi_m^a | \hat{H} | \Phi_m^a \rangle &= \langle \Phi_m^a | \hat{H}_m | \Phi_m^a \rangle + \sum_{n \neq m} \langle \Phi_m^a | \hat{H}_n | \Phi_m^a \rangle + \langle \Phi_m^a | \hat{V} | \Phi_m^a \rangle \\ &= \varepsilon_m^a + \sum_{n \neq m} \varepsilon_n^0 + \sum_{n \neq m} E_{m,n}^{a,0} + \sum_{\substack{n \neq m \\ p < n \\ p \neq m}} \sum_{p \neq m} E_{n,p}^{0,0} \end{aligned}$$

où  $E_{i,j}^{a,b}$  est l'énergie d'interaction entre la molécule  $i$  dans son état  $a$  et la molécule  $j$  dans son état  $b$ . Soit  $E_0$  l'énergie du système constitué de tous les chromophores dans leur

état fondamental, de fonction d'onde associée  $\Phi_0 = \prod_{n=1}^N \Psi_n^0$  :

$$E_0 = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle = \sum_n \varepsilon_n^0 + \sum_n \sum_{\substack{p < n \\ p \neq n}} E_{n,p}^{0,0}$$

L'élément diagonal  $\langle \Phi_m^a | \hat{H} | \Phi_m^a \rangle$  peut alors s'écrire :

$$\langle \Phi_m^a | \hat{H} | \Phi_m^a \rangle = E_0 + (\varepsilon_m^a - \varepsilon_m^0) + \sum_{n \neq m} (E_{m,n}^{a,0} - E_{m,n}^{0,0})$$

Le second terme du membre de droite correspond à l'énergie associée à l'excitation du chromophore  $m$  isolé, depuis son état fondamental vers son état électronique  $a$ .

Elle peut être calculée à divers niveaux, ou déterminée expérimentalement. Le troisième terme traduit la différence des interactions de la molécule  $m$  avec le reste de l'agrégat selon que cette molécule est dans son état fondamental ou excité, et constitue en ce sens un effet d'environnement différentiel (Figure 2.1). Il est généralement bien plus faible que le second terme ; son influence sur les états excités d'un multichromophore (en l'occurrence un double brin d'ADN) sera mise en évidence au cours de cette étude. Son calcul requiert un potentiel intermoléculaire adéquat.

Il faut noter, enfin, que tous les éléments diagonaux de la matrice excitonique contiennent l'énergie  $E_0$ , qui est indépendante du chromophore excité, donc de l'élément de matrice calculé. Il est possible de la soustraire à chaque élément diagonal afin de s'abstenir de son calcul. Les valeurs propres de la matrice excitonique représenteront alors les énergies associées aux transitions du multichromophore plutôt que les énergies absolues de ses états excités.

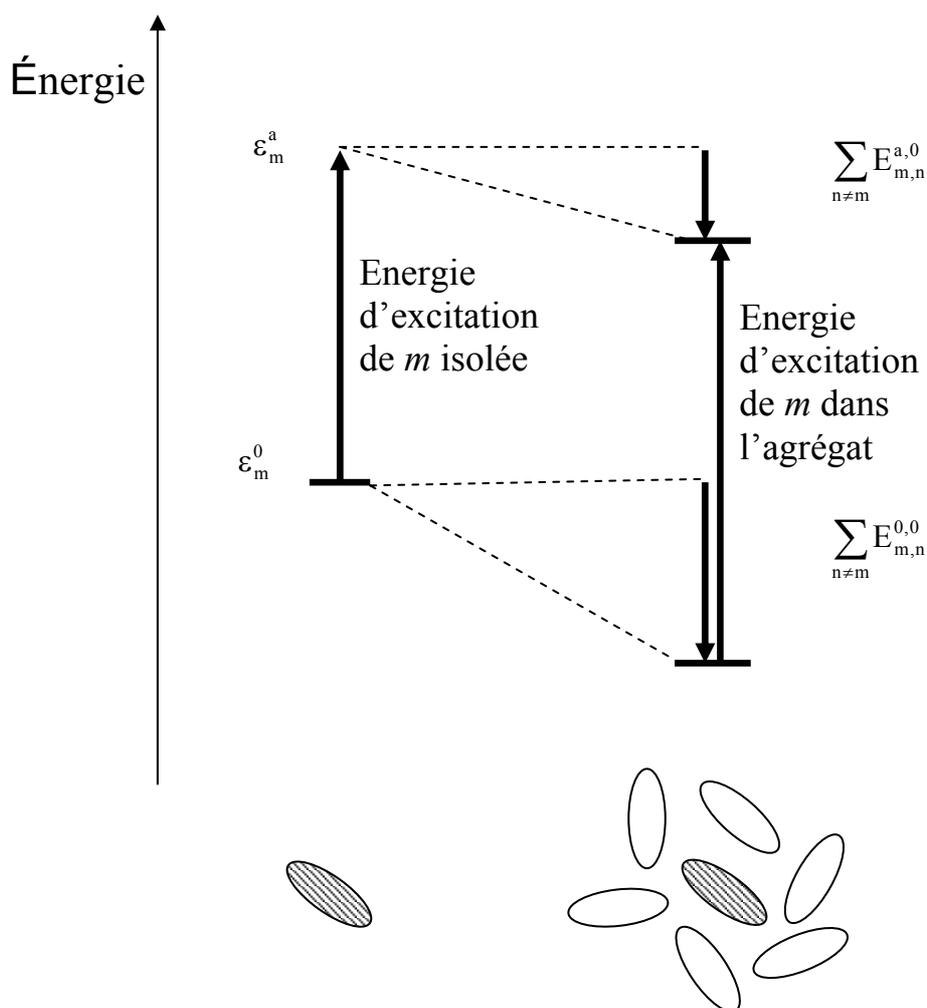


Figure 2.1

## 2. Description des États Électroniques Excités de Fragments d'ADN

### 2.1.3. Calcul des éléments non diagonaux

Dans le cadre des approximations précédemment formulées, les éléments non diagonaux de la matrice excitoniques qui couplent entre elles les formes à excitation localisée sont constitués de termes coulombiens correspondant à l'excitation d'un monomère et à la désexcitation simultanée de l'autre (termes de type (III) du tableau 2.1).

Soit  $J_{mn}(0,1,0,1)$  l'un de ces termes. La densité de transition associée sur la molécule  $m$  est définie par :

$$\rho_m^{0 \rightarrow 1}(\mathbf{r}_m) = (\psi_m^0)^*(\mathbf{r}_m) \psi_m^1(\mathbf{r}_m)$$

Elle traduit le degré de recouvrement entre les fonctions d'onde des états fondamental et excité de la molécule  $m$ . Le couplage excitonique  $J_{mn}(0,1,0,1)$  s'écrit alors :

$$J_{mn}(0,1,0,1) = \int (\rho_n^{0 \rightarrow 1})^* V_{mn} \rho_m^{0 \rightarrow 1} d\mathbf{r}_m d\mathbf{r}_n$$

Le couplage excitonique apparaît donc comme l'interaction électrostatique entre les distributions de charge des deux moments dipolaires de transition associés. Différents niveaux d'approximation peuvent être invoqués pour son calcul.

Si les distances entre molécules sont grandes par rapport à l'étendue spatiale des densités de transition, l'approximation dipôle-dipôle peut être invoquée. A cet effet,  $V_{mn}$  est retranscrit en fonction du vecteur  $\vec{R}_{mn}$  reliant les centres de masse des molécules  $m$  et  $n$  :

$$V_{mn} = \frac{1}{2} \sum_{i,j} \frac{e^2}{|\vec{R}_{mn} + \vec{r}_i(m) - \vec{r}_j(n)|}$$

un développement en puissances de  $|\vec{r}_i(m) - \vec{r}_j(n)|/|\vec{R}_{mn}|$  au second ordre fournit, pour  $V_{mn}$ , l'expression approchée suivante<sup>241</sup> :

$$V_{mn} \approx \frac{\vec{\mu}_m \cdot \vec{\mu}_n}{|\vec{R}_{mn}|^3} - 3 \frac{(\vec{R}_{mn} \cdot \vec{\mu}_m)(\vec{R}_{mn} \cdot \vec{\mu}_n)}{|\vec{R}_{mn}|^5}$$

avec  $\vec{\mu}_m = \sum_i e \vec{r}_i(m)$  (le moment dipolaire électronique). Soit de plus le dipôle de transition  $\vec{d}_m$  :

$$\vec{d}_m = \langle \psi_m^0 | \hat{\mu}_m | \psi_m^1 \rangle$$

Remplaçant dans l'expression du terme  $J_{mn}(0,1,0,1)$ , il vient :

$$J_{mn}(0,1,0,1) = \int (\rho_n^{0 \rightarrow 1})^* V_{mn} \rho_m^{0 \rightarrow 1} dr_m dr_n$$

$$= [\vec{n}_m \cdot \vec{n}_n - 3(\vec{e}_{mn} \cdot \vec{n}_m)(\vec{e}_{mn} \cdot \vec{n}_n)] \frac{|\vec{d}_m| |\vec{d}_n^*|}{|\vec{R}_{mn}|^3}$$

$$\text{avec } \vec{n}_m = \frac{\vec{d}_m}{\|\vec{d}_m\|}, \vec{n}_n = \frac{\vec{d}_n}{\|\vec{d}_n\|} \text{ et } \vec{e}_{mn} = \frac{\vec{R}_{mn}}{\|\vec{R}_{mn}\|}$$

Le calcul du couplage excitonique se résume donc à l'interaction de deux dipôles ponctuels de transition. Comme mentionné précédemment, cette approximation n'est valide que lorsque la séparation spatiale entre les molécules est grande (typiquement supérieure à 10 Å). Il a été montré<sup>242</sup> qu'à courte distance, le couplage est surestimé d'un facteur trois à six.

Le modèle du dipôle étendu<sup>243</sup> a été développé pour mieux évaluer le terme de couplage aux courtes et moyennes distances. Le dipôle étendu est le couple  $(q, \vec{l})$  tel que  $\vec{d}_m = q\vec{l}$ , les paramètres  $q$  et  $\vec{l}$  pouvant être tirés d'un niveau de calcul plus élevé tel le modèle de charges atomiques de transition décrit plus bas. La validité du couplage aux longues distances est ainsi assurée, tandis qu'aux courtes et moyennes distances la représentation du moment dipolaire de transition par deux charges ponctuelles séparées se révèle supérieure à celle du dipôle ponctuel. Ce gain en précision reste cependant relatif, notamment aux courtes distances. De plus, le placement du dipôle étendu sur la molécule est problématique si celle-ci ne dispose pas d'un centre d'inversion.

L'étape suivante vers une bonne description du couplage aux courtes distances d'approche est le modèle de la distribution de charges de transition. Il consiste à décomposer le moment dipolaire de transition en charges atomiques<sup>132,244</sup> à l'aide du formalisme INDO<sup>125</sup>. Outre un jeu de charges atomiques, cette méthode fournit des dipôles locaux proportionnels aux recouvrements de type s-p entre orbitales atomiques. Ceux-ci étant très faibles pour des états excités de type  $\pi \rightarrow \pi^*$ , la contribution des dipôles locaux au moment dipolaire de transition global est généralement très faible (inférieure à 5 %) <sup>116</sup>. Ils seront négligés dans cette étude. Il convient de noter que cette approximation n'est pas valide lors de la décomposition, par la même méthode, d'une densité électronique correspondant à un moment dipolaire permanent ; dans ce cas, la contribution des dipôles locaux est bien plus grande (environ 30 %) et ne peut être négligée. Il en serait de même pour la décomposition d'une densité de transition  $n \rightarrow \pi^*$ .

## 2. Description des États Électroniques Excités de Fragments d'ADN

Dans le cas de transitions singulet  $\pi \rightarrow \pi^*$  permises, il est donc raisonnable d'assimiler le couplage excitonique à l'interaction entre jeux de charges de transition sur les molécules concernées. Parce qu'elle retranscrit fidèlement l'anisotropie des interactions, cette approche est bien supérieure à ses rivales aux courtes distances d'approche et a été retenue pour cette étude. Cette supériorité sera mise en évidence dans le cas d'un double brin d'ADN, en comparaison avec le modèle du dipôle étendu.

## 2.2. Etats excités des bases de l'ADN

### *2.2.1. Préliminaires*

La mise en pratique de la théorie excitonique en base de fonctions d'onde diabatiques à excitation localisée nécessite la connaissance des états excités adéquats des monomères constituant le multichromophore à l'étude.

Le présent travail a pour finalité l'étude des états excités collectifs singulets de fragments d'ADN. Il est à présent bien établi que, dans ces systèmes, seules les bases azotées sont concernées par l'excitation, le squelette sucre-phosphate n'y jouant aucun rôle<sup>245</sup>. Du point de vue de la théorie excitonique, un fragment d'ADN est donc considéré comme un agrégat de bases, positionnées telles qu'elles le seraient dans la double hélice. En revanche, comme mentionné plus haut, la substitution des bases par le squelette sucre-phosphate altère les propriétés de tautomérie de celles-ci : certains tautomères, de poids statistiques importants dans le cas des bases isolées, sont inexistantes dans l'ADN suite à cette substitution. Le blocage de ces positions a été modélisé, pour les besoins de ce travail, par un groupement méthyle, dont il a été montré qu'il n'altère que peu les propriétés des états excités de la base<sup>246</sup>.

L'étude des états excités des bases de l'ADN et de leurs dérivés méthylés fait donc l'objet de ce chapitre. Les calculs menés à bien dans le cadre de ce travail seront présentés et comparés avec les données disponibles dans la littérature. Les fragments d'ADN étudiés étant composés exclusivement de résidus adénosine et thymidine, seuls les états excités de l'adénine et de la thymine seront présentés ici. Des calculs semi-empiriques des états excités de la guanine et de la cytosine ont également été exécutés ; le lecteur est renvoyé à la publication du paragraphe 3.2 de cette partie pour les résultats correspondants.

### *2.2.2. Problématique*

L'étude théorique des états excités des bases de l'ADN n'est pas récente. Malgré la panoplie impressionnante de méthodes mises en œuvre à cet égard, la caractérisation correcte de ces états excités reste délicate, ce qui est bien résumé par les propos de Callis : *'a kind of theoretical anarchy reigns with regard to excited state calculations on the DNA bases*

## 2. Description des États Électroniques Excités de Fragments d'ADN

*[...] depending on one's criteria and focus, the general agreement [with experiments] may be considered quite good or quite poor*<sup>245</sup>.

En premier lieu, toutes les bases de l'ADN ont plus d'électrons  $\pi$  que de centres, ce qui les rend isoélectroniques avec des di- et trianions inconnus, empêchant tout raisonnement par analogie avec des systèmes plus simples ou mieux élucidés. De plus, la présence d'un nombre important de doublets non liants et de protons labiles est responsable d'équilibres complexes entre de nombreuses formes tautomères distinctes. Elle laisse également présager l'existence d'états excités  $n \rightarrow \pi^*$  de basses énergies. Enfin, l'absence d'éléments de symétrie d'ordre plus élevés qu'un simple plan de réflexion ne facilite pas l'approche théorique. Les données expérimentales existantes ne suffisent souvent pas à lever les ambiguïtés ; les spectres d'absorption haute-résolution en phase gaz, lorsqu'ils sont disponibles, sont dominés par des bandes larges et mal résolues, tandis que la très faible durée de vie de fluorescence de ces espèces complique l'obtention de spectres de fluorescence exploitables.

### 2.2.3. État des lieux

La présente étude a été menée conjointement avec un groupe d'expérimentateurs du Laboratoire Francis Perrin, qui ont procédé à l'enregistrement des spectres d'absorption des bases de l'ADN en solution (voir la publication ci-incluse pour davantage de détails). L'analyse de ces spectres a montré que la bande d'absorption la plus basse était constituée de deux transitions  $\pi \rightarrow \pi^*$  dans le cas de l'adénine contre une seule dans la thymine. Conformément au modèle de bandes présenté en première partie, paragraphe 3, l'étude théorique sera donc restreinte à ces transitions. Des transitions  $n \rightarrow \pi^*$  de basse énergie, dues à l'excitation des électrons non liants des atomes d'azote intracycliques, peuvent également être envisagées. Leur présence a été postulée expérimentalement dans la même zone que les transitions  $\pi \rightarrow \pi^*$  de plus basse énergie, de sorte que l'ordre relatif des transitions, de même que la nature du premier état excité, reste sujet à débat<sup>245,247-249</sup>. Ces transitions sont en revanche quasiment interdites, puisque caractérisées par des forces d'oscillateur très faibles ; de plus, aucune preuve expérimentale irréfutable de l'existence de telles transitions dans cette zone n'a pu être apportée. Il apparaît donc légitime de s'affranchir de ce type de transition dans le cadre de cette étude.

Le lecteur est renvoyé à la troisième partie (*Individual chromophores*) de la publication ci-incluse pour un état des lieux des études théoriques et expérimentales sur les états excités  $\pi \rightarrow \pi^*$  des bases isolées.

### 2.2.4. Choix des méthodes

La caractérisation des états excités des bases de l'ADN comme préliminaire à l'usage de la théorie excitonique impose des conditions spécifiques. Il faut tout d'abord assurer une description correcte des fonctions d'onde des états excités des bases, afin de pouvoir calculer avec une précision suffisante les énergies des formes à excitation localisée qui constituent les éléments diagonaux de la matrice excitonique. De surcroît, le calcul des éléments non diagonaux est rendu possible par la connaissance des énergies et densités de transition des bases, ce qui impose de savoir décrire l'état fondamental et les états excités de ces bases à un niveau équivalent ; il s'agit notamment de pouvoir rendre compte de la variation de la corrélation dynamique entre les états fondamental et excités.

#### a) Cas des éléments non diagonaux

Le choix de la méthode de calcul des propriétés de transition s'est porté sur la méthode semi-empirique CS-INDO-CIPSI. Cette méthode fait se succéder un calcul Hartee-Fock dans l'approximation INDO et une interaction entre configurations triées par perturbation. L'interaction de configurations est un choix classique lors d'études d'états excités électroniques. L'interaction de configuration limitée aux monoexcitations dans le système  $\pi$  est l'approche la moins coûteuse ; elle revient cependant à supposer que le système  $\sigma$  reste inchangé lors de l'excitation dans le système  $\pi$ . Cette approximation peut se révéler légitime dans le cas d'états localement excités, c'est-à-dire pour lesquels les orbitales moléculaires  $\pi$  et  $\pi^*$  impliquées sont localisées au même endroit sur la molécule. En revanche, bon nombre d'états excités ont un caractère à transfert de charge intramoléculaire : la probabilité de présence associée aux orbitales  $\pi$  et  $\pi^*$  ne culmine pas dans les mêmes zones, ce qui provoque une séparation de charge au sein de la molécule. Les charges partielles qui en résultent polarisent alors le système  $\sigma$  de manière non négligeable. Ce cas de figure se produit fréquemment, par exemple si l'excitation concerne des orbitales non liantes tels que des doublets libres. Les états excités des bases de l'ADN sont également concernés<sup>250</sup>.

## 2. Description des États Électroniques Excités de Fragments d'ADN

La polarisation du système  $\sigma$  peut être prise en considération en rajoutant à l'espace actif de l'interaction de configuration les monoexcitations  $\sigma \rightarrow \sigma^*$  par rapport aux configurations  $\pi \rightarrow \pi^*$  précédentes, en d'autres termes les diexcitations  $\pi\sigma \rightarrow \pi^*\sigma^*$  par rapport à l'état fondamental. Une partie de la corrélation par rapport à l'état fondamental est ainsi ajoutée. Le reste est contenu dans les diexcitations  $\pi\pi \rightarrow \pi^*\pi^*$ , qui traduisent l'autorepolarisation du système  $\pi$  après excitation. La description de l'état fondamental est alors très correcte. En revanche, les états excités ne sont pas traités avec le même degré de précision : la prise en compte de la corrélation par rapport à ces états nécessite l'inclusion de configurations triexcitées par rapport à l'état fondamental. Celles-ci sont en nombre tel que le calcul d'interaction de configurations devient pratiquement irréalisable (taille de la matrice, stockage des intégrales biélectroniques...). Il est pourtant indispensable, pour calculer des propriétés de transition avec rigueur, de traiter de la même façon les états fondamental et excités.

Il est possible de restreindre le nombre de configurations significatives en travaillant en orbitales  $\sigma$  localisées, tout en conservant le caractère délocalisé des orbitales  $\pi$  inhérent à l'aromaticité. De la sorte, les excitations impliquant deux orbitales  $\sigma$  localisées en différents endroits de la molécule ne contribueront que peu au phénomène de corrélation et pourront être négligées. La localisation du système  $\sigma$  s'effectue par liaison ; elle est facilitée par l'emploi, dans le cadre de la méthode INDO, d'orbitales atomiques hybrides. Un critère topologique simple (localisation des orbitales  $\sigma$  et  $\sigma^*$  sur la même liaison) peut alors être invoqué pour retenir ou écarter les différentes configurations<sup>251</sup>. Une fois l'espace des déterminants actifs constitué, l'interaction de configuration à proprement parler est exécutée à l'aide de l'algorithme CIPSI, qui effectue un second tri de l'espace actif<sup>252</sup>.

Cette méthodologie cumule un certain nombre d'avantages. En premier lieu, les bonnes performances de la méthode CS-INDO quant au calcul des états excités de molécules aromatiques ont été vérifiées à plusieurs reprises<sup>126,253,254</sup>. De surcroît, le double tri des configurations (par la localisation des orbitales  $\sigma$  suivie de l'algorithme CIPSI) optimise le coût en temps de calcul tout en conservant les éléments essentiels à la bonne description des états excités. Enfin, tout au long du processus, la plus grande attention est portée à ce que les fonctions d'onde des états fondamental comme excités soient traités avec la même qualité.

### b) Cas des éléments diagonaux

L'élément diagonal de la matrice excitonique représente l'énergie du système constitué par le fragment d'ADN à une géométrie donnée, l'une des bases étant dans un état excité et les autres à l'état fondamental. Comme démontré précédemment, si la référence des énergies est celle du système de même géométrie mais dont tous les chromophores se trouvent dans leurs états fondamentaux respectifs, l'élément diagonal est la somme de l'énergie d'excitation du chromophore excité et d'un terme d'environnement différentiel traduisant la différence entre les énergies d'interaction au sein de l'agrégat induite par l'excitation du chromophore concerné, aussi dénommé 'désordre diagonal'.

Dans la première partie de cette étude, il a été choisi de négliger ce terme au profit de la seule énergie d'excitation du chromophore, déterminée expérimentalement (voir les parties 2.1. et 3. de l'article ci-inclus). Ainsi, l'énergie d'une forme à excitation localisée ne dépend que de la nature et de l'état électronique de la base excitée, quelles que soient la géométrie du brin d'ADN et la position de la base en son sein. Cette approximation est comparable à celle effectuée par Orlandi et Troisi dans leur étude dynamique du phénomène de transfert de charge dans l'ADN<sup>201</sup> : ces auteurs assimilent effectivement les éléments diagonaux de l'hamiltonien effectif aux potentiels d'ionisation expérimentaux des bases isolées concernées.

L'influence du désordre diagonal a été évaluée dans la deuxième partie de cette étude. Elle nécessite de disposer d'un potentiel intermoléculaire capable de retranscrire les interactions au sein de l'agrégat, pour toute configuration électronique possible des chromophores constitutifs. L'élaboration d'un tel potentiel sur la base du champ de force AMBER<sup>255,256</sup> est décrite au paragraphe 4.2.

### 3. États Excités Collectifs de Fragments d'ADN : Approche Statique

## 3. États Excités Collectifs de Fragments d'ADN : Approche Statique

### 3.1. Introduction

**L**a première étape vers la description des états excités collectifs de brins d'ADN de type poly-dA et poly-(dAdT) est une étude statique menée sur les géométries canoniques des brins (voir 1.2.), en négligeant l'influence du désordre diagonal. La caractérisation des états propres de ce système modèle, au moyen des outils de la théorie excitonique, a fait l'objet d'une publication qui est présentée ci-dessous.

### 3.2. Article

# Dipolar coupling between electronic transitions of the DNA bases and its relevance to exciton states in double helices

B. Bouvier, T. Gustavsson, D. Markovitsi\*, P. Millié\*

*Laboratoire Francis Perrin (CNRS FRE 2298), SPAM, CEA Saclay, F-91191 Gif-sur-Yvette, France*

Received 9 May 2001; in final form 15 July 2001

---

## Abstract

The present communication addresses the question of the magnitude of dipolar coupling between the lowest electronic transition moments of the DNA nucleosides and its relevance to Frenkel exciton states in double helices. The transition energies and moments of the nucleosides are determined from absorption spectra recorded for dilute water solutions. Dipolar interactions are computed for some typical nucleoside dimers according to the atomic transition charge distribution model. The properties of the exciton states of two particular double helices,  $(dA)_{20} \cdot (dT)_{20}$  and  $(dAdT)_{10} \cdot (dAdT)_{10}$ , are calculated considering three closely lying molecular electronic transitions ( $S_1$  and  $S_2$  for adenosine,  $S_1$  for thymidine). It is shown that (i) the oscillator strength is distributed over a small number of eigenstates, (ii) important mixing of the three monomer electronic transitions may occur, (iii) all eigenstates are spatially delocalised over the whole length of the double helix and (iv) the extent of exciton states over the two strands depends on the base sequence. © 2002 Elsevier Science B.V. All rights reserved.

---

## 1. Introduction

Various photoreactions occurring upon absorption of UV irradiation by the DNA bases (adenine, cytosine, guanine and thymine) are known to have lethal or mutagenic effects on the cells [1]. The yield of such photoreactions may be enhanced if they are preceded by a transport process (charge or electronic excitation energy) within the double helix. In view of the crucial role

the transport phenomena can play in the appearance of skin cancers, it is important to elucidate the associated mechanisms. Indeed, the study of charge transport in DNA has recently found a renaissance with the support of ultrafast spectroscopy and sophisticated theoretical calculations [2]. In contrast, such means have not yet been used to improve our understanding of energy migration in DNA for which no clear picture has emerged so far in spite of the efforts made for more than thirty years [3–8].

The main controversial point is whether the excited states involved in singlet excitation migration are localised on just one base or whether they extend over a certain number of them. The exciton theory has been used since the early sixties in the

---

\* Corresponding authors. Tel.: +33-1-6908-4644; fax: +33-1-6908-8707.

*E-mail addresses:* dmarkovitsi@cea.fr (D. Markovitsi), millie@pandora.saclay.cea.fr (P. Millié).

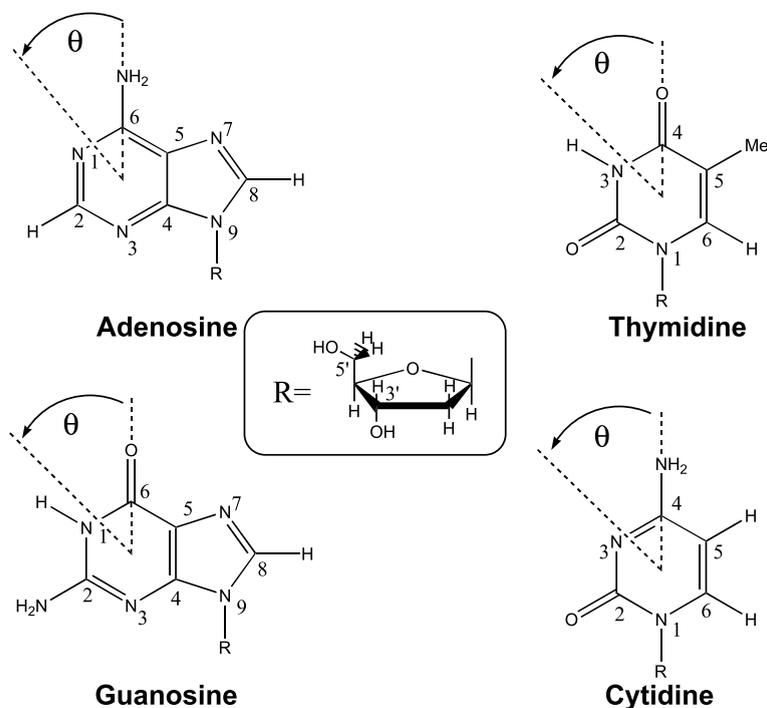


Fig. 1. The studied nucleosides: adenosine (dA), cytidine (dC), guanosine (dG) and thymidine (dT). In the quantum chemistry calculations the deoxyribose residue is replaced by a methyl group. The angle  $\theta$  denotes the direction of the transition moments.

analysis of the DNA electronic excited states [9–12]. Some authors draw the conclusion that the excited states remain localised since exciton splitting is not observed upon formation of the double helix [13]. This type of reasoning has also guided most of the “experimental”<sup>1</sup> investigations [3–8] of singlet energy migration where the photosensitised fluorescence from an energy trap is analysed in terms of excitation hopping via a Förster mechanism [14].

This controversy arises from the complexity of the system and is related to both theoretical and experimental difficulties, as well as to conceptual problems. Below, we enumerate some of them and we comment on the possibility of overcoming them. In what follows we use the term “chromophore” to denote the nucleosides, deoxyadenosine (dA), deoxycytidine (dC), deoxyguanosine (dG) and deoxythymidine (dT), composed of a base and

the deoxyribose residue (Fig. 1). Their optical properties differ from those of the bases but they are quite similar to those of nucleotides which are the monomeric units of nucleic acids [15,16].

1. The lowest absorption bands of the four nucleosides are very broad; they all peak in the same spectral region (260–270 nm) and they have similar intensities [15,16]. Moreover, fluorescence anisotropy and circular dichroism measurements have revealed that the adenosine and guanosine absorption bands correspond to more than one electronic transition [17–19]. Thus, it would in principle be possible to improve the modelling of electronic excitations in double helices by using experimental values of transition energies and transition moments of individual chromophores resulting from a more refined spectral analysis.

2. Since the excitation energy, the transition moment and the polarisation of the dipolar transitions are specific to each type of nucleoside, the collective properties of the double helices are expected to depend on the base sequence. Therefore,

<sup>1</sup> In the sense that they are model dependent.

a study of energy migration at the molecular level can only be performed for double helices with known base sequence and not for native DNA. The base sequence effect on the energy transfer efficiency to an energy trap has been experimentally evidenced for double-stranded decamers [20,21].

3. In most calculations of the DNA exciton states, the dipolar coupling between electronic transition moments was calculated by means of the point dipole approximation. This approximation is also implicit in the Förster formula [14] used in the “experimental” studies approaching energy migration as a hopping process. One would expect this approximation not to be valid for double helices where the interchromophore distance is of the same order of magnitude as the chromophore dimensions, as already pointed out in 1969 by Miyata and Yomosa [12]. The electronic coupling can be calculated in a much more accurate way using the method of atomic transition charges [12,22,23].

4. In all the studies dealing with DNA exciton states and energy migration in double helices, not only may the dipolar coupling be unsatisfactorily described but interactions due to orbital overlap are also neglected. Among these interactions, charge resonance and charge transfer terms may become important when the strength of dipolar transitions decreases [24–26]. In particular, these interactions are involved in the formation of excimers and exciplexes often invoked in discussions on nucleic acid components [27–31].

5. The strength of the electronic coupling responsible for the formation of collective excited states and excitation transport may be seriously affected by structural disorder. Various types of conformations are known to occur in double helices in solution. They depend on the base sequence and can be simulated using molecular modelling [32].

6. We have seen in point 2 that it is important to work with double helices with known base sequence so as to correctly model the excited states and the excitation transport. The experimental counterpart resides in the difficulty of obtaining well-defined double helices in solution. Although polynucleotides of known base sequence are commercially available, the pairing of comple-

mentary single strands may lead to mismatching. Structural characterisations have to be carried out before the photophysical properties measured can be unambiguously attributed to double-stranded helices.

7. Another serious experimental difficulty when working with nucleic acids is the low-fluorescence quantum yields [33] and the extremely short lifetimes of the singlet excited states [34,35]. Only recently has femtosecond transient absorption spectroscopy been used to study the photophysical properties of bases and nucleosides [36,37].

Within this context we have undertaken a study aimed at a better understanding of electronic excitations in polynucleotide double helices using methodologies we previously developed for the investigation of exciton states and excitation transport in columnar phases [23,26,38,39]. The present communication is our first approach to this problem. Our efforts are focused on the improvement of the points 1 to 3 mentioned above, e.g., use of accurate properties for chromophore transition moments, consideration of oligomers with given base sequence and calculation of dipolar coupling by means of the atomic transition charges.

Firstly, we determine the properties of the individual chromophores in the double helices. Experimentally, we record the absorption spectra of nucleosides in dilute water solutions to avoid aggregation. Using fluorescence anisotropy data from the literature [17,19], we determine the energy and the transition moments of the lowest electronic transitions. In parallel, we calculate these properties with the conformation spectra–intermediate neglect of differential overlap–configuration interaction by perturbative selected iterations (CS-INDO-CIPSI) method and we compare the results to both our experimental data and to the results of calculations as reported in the literature (Section 3).

Secondly, we are interested in the magnitude of dipolar coupling encountered in some typical dimers, both stacked and paired, as was recently done for the electronic coupling responsible for charge transport [40]. We compare the results obtained using the atomic transition charge distribution model with those calculated by means of

the point dipole approximation. Moreover, we examine whether it is possible to determine, for each nucleoside, a dipole length so that the extended dipole approximation [41], convenient for simulations, can be used to correctly describe the dipolar coupling in double helices (Section 4).

Thirdly, we calculate the properties of the exciton states for two simple DNA oligomers, each one consisting of twenty identical dA–dT pairs, by considering two electronic transitions for adenosine and one for thymidine. The two types of nucleosides are either located on different strands  $(\text{dA})_{20} \cdot (\text{dT})_{20}$ , or alternate in the same strand  $(\text{dAdT})_{10} \cdot (\text{dAdT})_{10}$ . We examine the energy, the oscillator strength and describe the spatial and electronic delocalisation of each eigenstate. Finally, we illustrate the topography of some typical eigenstates by representing the contribution of the electronic transitions of each particular nucleoside to these collective states (Section 5).

## 2. Methods

### 2.1. Experimental procedure

Nucleosides were purchased from Sigma Aldrich and used without further purification. They were dissolved in ultrapure water produced by a Millipore (Milli-Q) purification system. Absorption spectra were recorded in the UV region down to 200 nm with a Perkin Lambda 900 spectrophotometer using 1 and 5 cm quartz cells.

The spectral analysis was performed by using PeakFit software (Jandel Inc.). Each spectrum was decomposed into a sum of sub-bands. Each sub-band was represented by a log-normal function according to the equation:

$$\varepsilon(v) = \varepsilon_0 \exp \left[ - \frac{\ln(2) \left[ \ln \left( \frac{(v-v_0)(\gamma^2-1)}{\gamma \Delta v} + 1 \right) \right]^2}{[\ln \gamma]^2} \right], \quad (1)$$

where  $\varepsilon_0$  is the maximum value of the molar extinction coefficient,  $v_0$  the position of the absorption maximum,  $\Delta v$  the full width at half maximum

(FWHM) of the absorption band and  $\gamma$  the asymmetry parameter. It has been shown that a log-normal curve better describes an asymmetric structureless absorption band than a gaussian one [42].

The oscillator strengths  $f$  and the transition moments  $\mu$  (in Debyes) corresponding to a given log-normal curve were calculated according to the equations:

$$f = 4.317 \times 10^{-9} A, \quad (2)$$

$$\mu^2 = 9.181 \frac{A}{v_0} \times 10^{-3}, \quad (3)$$

where  $A = \int_{-\infty}^{\infty} \varepsilon(v) dv$  is the integral of the absorption band (in  $\text{l mol}^{-1} \text{cm}^{-2}$ ) and  $v_0$  is the peak transition energy (in  $\text{cm}^{-2}$ ).

### 2.2. Calculation procedure

#### 2.2.1. General formalism

The excited states of double-stranded DNA fragments are calculated in the framework of the exciton theory [43–45], in which the exact Hamiltonian of an  $n$ -molecular supersystem may be written:

$$H = H_0 + V. \quad (4)$$

$H_0$ , the unperturbed Hamiltonian of the supersystem, is evaluated as the sum of the individual Hamiltonians of the isolated molecules. As a consequence, zeroth-order eigenfunctions of the system consist of products of the eigenfunctions of these molecules:

$$H_0 = \sum_{\text{molecules}, k}^n H_k \quad \text{and} \quad |\Phi_m^i\rangle = \Psi_m^i \prod_{\text{molecules}, k \neq m}^n \Psi_k^0, \quad (5)$$

where  $\Psi_m^i$  denotes the  $i$ th excited state of chromophore  $m$ . The excited states of the supersystem are thus decomposed on states for which the excitation is localised on a given monomer, the other molecules being in their respective ground states.

$V$  is the perturbation operator, which was formalised by Longuet-Higgins [46] using a local charge density operator  $\rho^{(k)}(\vec{r}^{(k)})$  associated with each individual chromophore  $k$ :

$$V = \int \int \frac{\rho^{(k)}(\vec{r}^{(k)})\rho^{(l)}(\vec{r}^{(l)})}{|\vec{r}^{(k)} - \vec{r}^{(l)}|} d\vec{r}^{(k)} d\vec{r}^{(l)}. \quad (6)$$

Since these operators depend exclusively on the coordinates of the individual subsystems, separation of variables inside the integrand yields, as an expression of the general matrix element of  $V$ :

$$\langle \Psi_n^i \Psi_m^j | V | \Psi_n^k \Psi_m^l \rangle = \frac{\langle \Psi_n^i | \rho^{(n)} \vec{r} | \Psi_n^k \rangle \langle \Psi_m^j | \rho^{(m)} \vec{r} | \Psi_m^l \rangle}{|\vec{r}^{(n)} - \vec{r}^{(m)}|} d\vec{r}^{(n)} d\vec{r}^{(m)}. \quad (7)$$

This expression may be viewed as an interaction between appropriate transition charges which can be calculated at different levels of approximation (point dipoles, extended dipoles, atomic transition charges, ...).

The placement of a point dipole on a molecule devoid of inversion centre (such as the DNA bases) is problematic. Moreover, this approximation is inappropriate when the interchromophore distance is of the same magnitude as molecular dimensions. In such cases, the extended dipole approximation [41] is traditionally used. An electronic transition at this level of approximation is represented by two opposite charges  $+q$  and  $-q$  separated by a distance  $l$ , such that  $\vec{\mu} = ql$ , where  $\mu$  is the transition moment.

In the atomic transition charges model, the off-diagonal terms are subjected to a dipolar development [47]. The resulting molecular transition dipoles  $\vec{\mu} = \langle \Psi_m^k | \vec{r} | \Psi_m^0 \rangle$  are then decomposed onto the atomic orbitals of the molecule, in the framework of the INDO approximation [48]. This treatment yields transition charges located on each atom of the molecule, as well as local transition dipoles whose contribution is generally negligible for  $\pi \rightarrow \pi^*$  transitions (typically  $\approx 5\%$ ). The dipole moment associated with the charge distribution is approximately equal to the corresponding transition moment.

At the first order of perturbation, consideration of the expression of the total electronic Hamiltonian  $H$  in the basis formed by the eigenfunctions of  $H_0$  is sufficient. Diagonalisation of the matrix yields  $N$  eigenstates

$$\Phi_k = \sum_{\text{molecules } m} \sum_{\text{states } i} C_{k,m}^i | \Psi_1^0 \Psi_2^0 \dots \Psi_m^i \dots \Psi_n^0 \rangle. \quad (8)$$

The transition moments associated with the eigenstates are given by:

$$\vec{\mu}_k = \sum_{\text{molecules } m} \sum_{\text{states } i} C_{k,m}^i \vec{\mu}_m^i. \quad (9)$$

Eq. (9) implies that the model is conservative with regard to  $\mu^2$ : the sum of squared transition moments of the monomers equals the sum of squared eigenstate transition moments.

The localisation behaviour of the eigenstates is usually expressed by the inverse participation ratio  $L_k$  [49,50]. The number of coherently coupled chromophores in a given eigenstate  $k$  is given by the participation ratio  $1/L_k$ . When the eigenstates are built on more than one molecular states,  $L_k$  is written as [23]:

$$L_k = \sum_{\text{molecules } m} \left[ \sum_{\text{states } i} (C_{k,m}^i)^2 \right]^2. \quad (10)$$

An interesting comparison can be drawn between the exciton theory and direct calculation of the excited states of the supersystem as a whole. At the lowest level of complexity, a configuration interaction (CI) restricted to singly excited configurations (CIS) is applied, on the one hand, to each monomer sub-unit and followed by an exciton treatment, and on the other hand, to the entire supersystem. These two treatments are equivalent under the following conditions: (i) the perturbation is sufficiently small, (ii) the basis used consists of molecular orbitals that are localised on each molecule, and (iii) orbital overlap interactions (exchange, charge resonance, charge transfer) are small compared to Coulombic interactions. Conversely, Eq. (5) implies complete neglect of orbital overlap interactions.

Alternately, transition charges can be obtained using a higher level of theory than a CIS. Direct treatment of the excited states of the supersystem at this same level is often tedious or unfeasible, whereas the exciton theory yields correct results if the previous three conditions are fulfilled. An important difference between the two aforementioned approaches concerns the modelling of the well-known DNA hypochromism [51], that is, the decrease in intensity of the absorption spectrum upon pairing of two single strands resulting in the

formation of the double helix. The supersystem CIS method accounts for hypochromism [52] most probably because it includes charge transfer excitations, in contrast to our approach which we deem unable to reproduce this phenomenon. Tinoco's [9,11] explanation of hypochromism based on the exciton theory, already criticised by Rhodes [10], can be attributed to a large number of approximations, particularly an overestimation of the off-diagonal terms due to the use of the point dipole model (cf. Section 4).

### 2.2.2. Computational details

The deoxynucleosides were modelled by the corresponding 9-methylated purines or 1-methylated pyrimidines and described by their most abundant tautomeric form (Fig. 1). Their geometry was optimised at the AM1 level. So as to allow separation of the  $\sigma$  and  $\pi$  subsystems in the INDO treatment, and in agreement with calculations on the isolated bases [53], the heavy atom skeletons were kept planar (point group  $C_s$ ) during the optimisation process. Rotation of the substitutive methyl group around the N1–C or N9–C bond axis was also optimised. The geometry of the two Watson–Crick base pairs dG–dC and dA–dT were derived from those of the individual bases; relative placement parameters were taken from [54]. Positioning of the pseudo-dyad followed conventions described in Ref. [55]. For the calculation of coupling elements between all possible base pairs, a standard B-DNA geometry was generated from the Watson–Crick pairs ( $36^\circ$  twist,  $3.4 \text{ \AA}$  rise) [54], regardless of the bases considered. The oligonucleotides  $(dA)_{20} \cdot (dT)_{20}$  and  $(dAdT)_{10} \cdot (dAdT)_{10}$  were built using a standard B-DNA geometry. Sequence dependent changes in helical twist and rise were taken from X-ray diffraction studies on fibres [56] and single crystals [57]:  $36^\circ$  twist and

$3.29 \text{ \AA}$  rise for  $(dA)_{20} \cdot (dT)_{20}$ ,  $31^\circ/41^\circ$  twist and  $3.34 \text{ \AA}$  rise for  $(dAdT)_{10} \cdot (dAdT)_{10}$ . The sugar-phosphate backbones of DNA helices were not taken into account.

For the calculation of dipolar coupling, computed transition moments were rescaled so as to coincide with experimental values, whereas transition moment directions were those determined according to the atomic transition charges model. Diagonal elements of the exciton matrix were taken equal to the experimental excitation energy (Table 1). Successive transitions on the same purine were considered orthogonal to each other. Interactions were not restricted to nearest neighbours: all the off-diagonal matrix elements were evaluated.

The extended dipole for a given transition was positioned along the axis defined by the weighted centres of negative and positive atomic transition charges. It was placed at the weighted centre of the absolute values of the atomic transition charges. The values of charge  $q$  and distance  $l$  were chosen so as to verify the following conditions: (i)  $\mu = ql$ ,  $\mu$  being the experimental transition dipole moment, and (ii) the coupling matrix elements calculated on the basis of the interaction between two extended dipoles is equal to that calculated with the atomic transition charges model, for two superposed identical DNA bases whose planes of symmetry are parallel and separated by a distance of  $3.5 \text{ \AA}$ . These two conditions uniquely define parameters  $q$  and  $l$ .

The  $\sigma$  subsystem on each chromophore was localised on individual bonds, whereas the  $\pi$  subsystem was allowed to delocalise over the whole molecule. The subsequent CI was not limited to monoexcitations or to the  $\pi$  and  $\pi^*$  molecular orbitals, so as to ensure a better description of differential correlation effects and  $\sigma$  system

Table 1

Energy ( $\text{cm}^{-1}$ ) and transition moment ( $D$ ) of the three lowest singlet electronic transitions of nucleosides deduced from the experimental absorption spectra in  $10^{-5} \text{ M}$  water solutions

	dA <sup>a</sup>	dC	dG	dT
$S_0 \rightarrow S_1$	36 700 (1.60)	36 800 (3.45)	36 700 (3.31)	37 500 (3.68)
$S_0 \rightarrow S_2$	38 800 (3.70)	42 100 (2.18)	40 300 (3.31)	46 600 (2.75)
$S_0 \rightarrow S_3$	45 500 (1.20)	44 900 (2.22)	48 200 (2.44)	49 900 (1.63)

<sup>a</sup> From Ref. [19].

reorganisation. Individual  $\sigma_i \rightarrow \sigma_j^*$  transitions were allowed to contribute as long as molecular orbitals  $i$  and  $j$  share a common atomic centre. This approach typically accounts for 95% of the correlation energy. Resulting configurations were then sorted with regard to their weight, in the framework of the CIPSI [58] algorithm, and yielded a CI matrix containing five to ten thousand individual configurations. The wave functions generated by diagonalisation of the CI matrix were used to determine atomic transition charges, as well as transition dipole moments associated with each excited state.

### 3. Individual chromophores

In this Section, we focus on the electronic transitions of individual chromophores on which the magnitude of dipolar coupling and the properties of exciton states depend. The transition energies and the transition moments are obtained from the experimental absorption spectra, whereas transition directions and atomic transition charges

are calculated using the CS-INDO-CIPSI method, whose validity is checked by comparing the experimental and calculated values of transition energies and transition moments. As a wealth of theoretical information about the electronic transitions of the nucleic acid bases has been gathered over the years, we also compare our theoretical data with selected literature results.

Fig. 2 shows the absorption spectra of the four nucleosides obtained for  $10^{-5}$  M concentrations. We found that at higher concentrations the Beer-Lambert law is not obeyed. The concentration dependence of the photophysical properties of nucleosides will be the object of a forthcoming communication. The spectrum profiles in Fig. 2 resemble those reported in the literature and the molar extinction coefficients match within 10% the reported values, for which the concentrations are not mentioned [15,16].

We performed a decomposition of each spectrum into sub-bands in the region 30 000–50 000  $\text{cm}^{-1}$  according to the fitting procedure described in Section 2.1. All the log-normal curves used to fit each spectrum were constrained to have the same

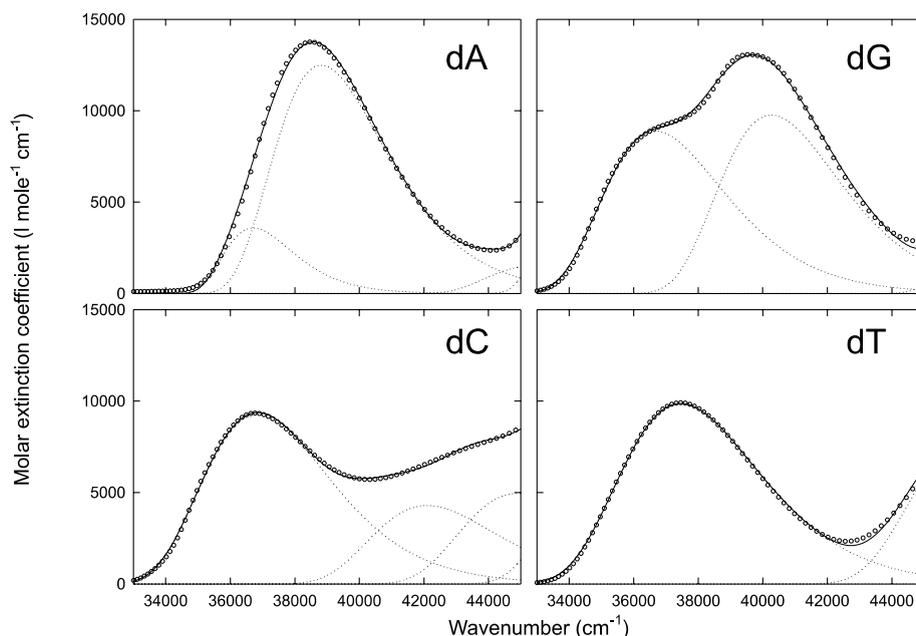


Fig. 2. Absorption spectra (circles) of the adenosine, cytidine, guanosine and thymidine in ultrapure water ( $10^{-5}$  M) fitted with a sum (solid lines) of log-normal curves (dotted lines).

width and we used the minimum number of curves giving an acceptable fit. Although our decomposition is phenomenological, we based it on fluorescence anisotropy and linear dichroism measurements reported in the literature. These measurements suggest that the lowest absorption band corresponds to a single electronic transition for cytidine and thymidine whereas that of adenosine and guanosine is an envelope of two transitions with different polarisations [17]. For adenosine, we fixed the first two transition energies to the values given in a recent detailed analysis [19], and allowed the first weak transition to be slightly narrower.

The resulting fitted curves are shown as solid lines in Fig. 2. As can be seen, the description is fully satisfactory. The sub-bands above 42 000  $\text{cm}^{-1}$  are only included to stabilise the description of the lowest transitions and are not further treated in the present paper. The numerical results of the fits for the three lowest transitions are shown in Table 1. In what follows, we will consider only one

transition for cytidine and thymidine and two for adenosine and guanosine. This choice is made because the other transitions are too high in energy and/or they have a weak transition moment. Therefore, they are not expected to mix with the lower ones within the same exciton band.

Experimental and calculated values of transition energies, along with associated directions and oscillator strengths, are compiled in Tables 2 and 3 for purines and pyrimidines, respectively. Transition moment directions, visualised by means of the angle  $\theta$  [59] (Fig. 1), are taken from the literature [19,60–62]. Selected literature results based on semi-empirical methods [53,63–65] and more recent studies based on ab initio calculations [53,66–69] are presented in the same Tables.

The theoretical and experimental values are in good agreement as far as excitation energies and oscillator strengths are concerned. The oscillator strength of cytidine is rather underestimated, probably because the global parameterisation of the CS–INDO method is not as well adapted for

Table 2

Energy ( $\text{cm}^{-1}$ ), oscillator strength and directions of the two lowest electronic transitions of purines. Comparison with literature data

Method/level	$S_0 \rightarrow S_1$			$S_0 \rightarrow S_2$		
	$E$	$f$	$\theta^a$	$E$	$f$	$\theta^a$
<i>Adenosine</i>						
CS–INDO	37 175	0.057	+1	39 840	0.216	+52
Experiment (Section 2.1)	36 700	0.05	+66 <sup>b</sup>	38 800	0.24	+19 <sup>b</sup>
INDO/S <sup>c</sup>	35 714	0.099	+30	38 168	0.305	+62
Rescaled CIS/6-31 G <sup>*d,e</sup>	37 100	0.41	+38	37 500	0.03	-72
Ab initio MRCI <sup>d,f</sup>	36 400	0.001	-33	38 600	0.318	+41
CASPT2 <sup>d,e</sup>	38 100	0.236	+43	38 200	0.001	-56
<i>Guanosine</i>						
CS–INDO	34 245	0.232	-57	40 160	0.203	+54
Experiment (Section 2.1)	36 700	0.19	-4/35 <sup>g</sup>	40 300	0.21	-75 <sup>g</sup>
INDO/S <sup>h</sup>	31 750	0.28	-44	36 300	0.38	+57
Rescaled CIS/6-31 G <sup>*d,e</sup>	35 600	0.29	-46	40 300	0.50	+63
Ab initio MRCI <sup>d,f</sup>	36 200	0.20	-64	42 700	0.09	+52
CASPT2 <sup>d,i</sup>	35 842	0.24	-52	36 900	0.09	+3

<sup>a</sup> See Fig. 1.

<sup>b</sup> From Ref. [19].

<sup>c</sup> From Ref. [63].

<sup>d</sup> Values corresponding to 9H-purine.

<sup>e</sup> From Ref. [53].

<sup>f</sup> From Ref. [66].

<sup>g</sup> From Ref. [61].

<sup>h</sup> From Ref. [65].

<sup>i</sup> From Ref. [68].

Table 3  
Energy ( $\text{cm}^{-1}$ ), oscillator strength and directions (Fig. 1) of the two lowest electronic transitions of pyrimidines—comparison with literature data

Method/level	$S_0 \rightarrow S_1$		
	$E$	$f$	$\theta^a$
<i>Thymidine</i>			
CS-INDO	36 360	0.220	-13
Experiment (Section 2.1)	37 500	0.24	+14/-19 <sup>b</sup>
CNDO/OPTIC-2 <sup>c</sup>	39 200	0.245	+14
Rescaled CIS/6-31 G <sup>*d,e</sup>	41 100	0.36	-3
CASPT2 <sup>d,f</sup>	39 400	0.17	+15
<i>Cytidine</i>			
CS-INDO	36 100	0.102	+38
Experiment (Section 2.1)	36 800	0.21	+9/+51 <sup>g</sup>
INDO/S <sup>c</sup>	31 750	0.28	-44
Rescaled CIS/6-31 G <sup>*d,e</sup>	35 600	0.29	-46
CASPT2 <sup>d,h</sup>	36 200	0.20	-64

<sup>a</sup> See Fig. 1.

<sup>b</sup> From Ref. [60].

<sup>c</sup> From Ref. [63].

<sup>d</sup> Values correspond to 1H-pyrimidines.

<sup>e</sup> From Ref. [53].

<sup>f</sup> From Ref. [67].

<sup>g</sup> From Ref. [62].

<sup>h</sup> From Ref. [69].

this particular chromophore as for the others. Nevertheless, the overall accord between both sets of results justifies the rescaling of transition moments explained in Section 2.2.2. At this point, it is interesting to compare our data with these obtained by Miyata and Yomosa who studied excitation states of oligonucleotides [12]. The atomic transition charges derived by these authors from a limited CIS calculation [70] of the bases lead to important errors in the oscillator strength values and to state inversions.

A vast number of semi-empirical methods, reviewed in Ref. [18], have been utilised, ranging from INDO/S to CNDO or MNDO. These calculations are based on CI between singly or singly and doubly excited states. As a rule, they have proved able to reproduce most experimental transition energies and oscillator strengths, but generally fail to provide consistent transition moment directions. Discrepancies in results between different methods may be attributed to varying parameterisations. The necessity to include at least doubly excited states in the CI space to ensure

correct treatment of conjugated molecules, such as the DNA bases, has been pointed out by Volosov [63].

Correct ab initio treatment of transition energies and moments requires to take into account the variation of the dynamic correlation between the ground and excited states. Energy calculations that meet this requirement can be achieved using the CASPT2 method. The choice of the active space at the CASSCF level determines the quality of the zeroth-order wave function used for CASPT2; as a consequence, CASPT2 results for near-degenerate excited states may not prove satisfactory, as is the case for the first two  $\pi \rightarrow \pi^*$  transitions of adenine. These transitions, clearly separated at the CASSCF level, are rendered almost degenerate by the PT2 correction [53]. Ab initio CIS, although less computationally expensive, takes no heed of differential correlation effects nor of the  $\sigma$  system's reorganisation. Hence, results obtained at this level need to be rescaled, considering the difference in electronic correlation between the ground and excited states as a constant. Furthermore, CIS calculations do not correctly describe excited states possessing a substantial doubly excited character or multiconfigurational structure, for which CASSCF, MCSCF or MRCI methods are better suited.

On the whole, a relatively good agreement between different quantum chemistry methods as well as between theory and with experiment can be noted for the energies and oscillator strengths of the first few  $\pi \rightarrow \pi^*$  states of the methylated DNA bases. This is not the case concerning transition moment directions, which still lack reliability; experimental determinations of these directions may also be rather imprecise, rendering comparison with experiment somewhat troublesome. In particular, the transition moment directions of 9-methylguanine are very similar at every level of theory but blatantly disagree with the values reported by experimentalists.

#### 4. Dipolar coupling

In this Section, we present the atomic transition charge distributions for the four nucleosides,

which we apply to the determination of the dipolar coupling. Using this sophisticated formalism as a reference, we test the validity of the point dipole and the extended dipole approximations for geometries generated by simple spatial transformations (translations and rotations). Finally, we determine intrastrand and interstrand coupling elements for nearest neighbours in typical B-DNA structures.

The calculated transition monopoles on the atomic centres of the methylated bases are schematised in Fig. 3. As can be seen, the first transition of both purines is rather localised on the

six-membered ring, in contrast with the more delocalised character of the second transition for which the role of the transition monopole on atom C8 is notably increased. The substituent methyl groups do not carry significant charge and probably do not influence transition moment directions or intensities as far as the first two transitions on each purine are concerned. This is in agreement with the theoretical finding [64] that differently substituted adenines and guanines are characterised by similar transition charge distributions. In addition, the monopole decomposition presented in Fig. 3 is very similar to that reported by these

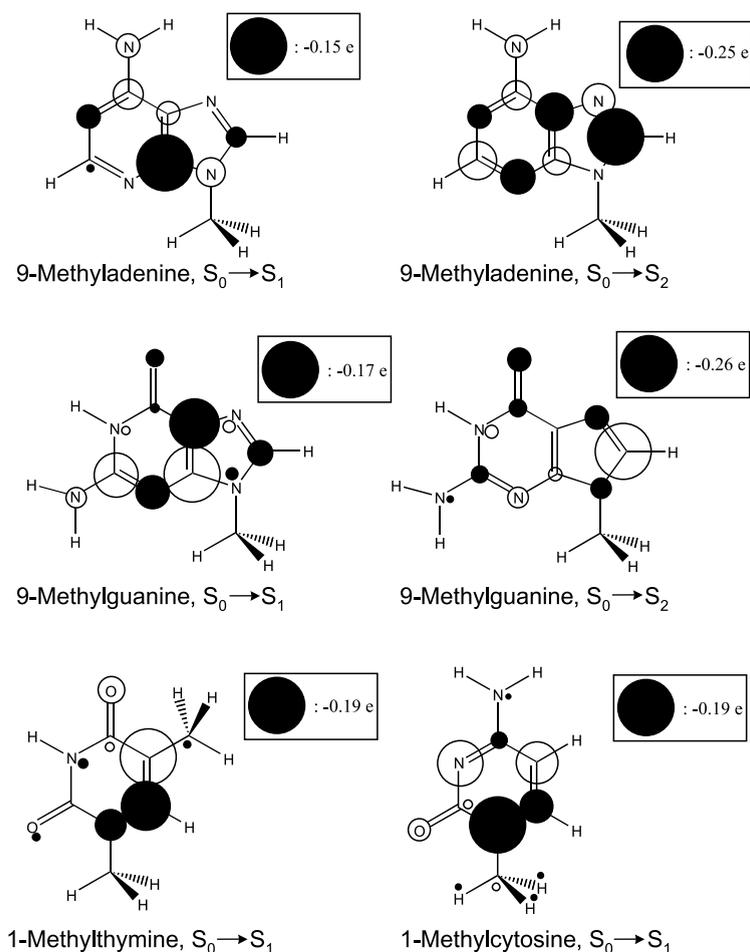


Fig. 3. Atomic transition charges found for the methylated DNA bases according to the CS-INDO-CIPSI method. Full circles correspond to negative charges and empty circles to positive charges; the circle diameter is proportional to the absolute value of the charge.

authors on the basis of the CNDO/OPTIC-2 semi-empirical method. The influence of carbonyl and amino groups on the transition moments may also be considered rather unimportant, as can be seen when comparing transition charge distributions on 1-methylthymine and 1-methylcytosine.

Table 4 presents the values of the extended dipole length  $l$  and the associated charge  $q$ , calculated as described in Section 2.2. We remark that values of  $l$  roughly correspond to typical dimensions of the part of the nucleoside concerned by the excitation. Furthermore, they are of the same order of magnitude compared to nearest-neighbour distances in double helices, justifying a priori the inadequacy of the point dipole model outlined further in this study.

The values of dipolar coupling  $V$  calculated at three levels of approximation are compared in Fig. 4 for two stacked and parallel thymidine chromophores, as a function of the stacking distance  $R$ . For clarity, we show the values of  $V$  scaled by  $R^3$ , giving a constant value for the point dipole model. The agreement between the atomic transition charges and the extended dipole models is remarkable at every distance. For values of  $R$  superior to 20 Å, both plots behave asymptotically and tend towards the point dipole curve. For stacking distances found in DNA (3.4 Å), the point dipole approximation results in a threefold error.

The aforementioned good performance of the extended dipole is understandable given the way in which  $q$  and  $l$  are defined (Section 2.2). The robustness of the approximation can be further tested by relative rotation of two chromophores around the helical axis. The result obtained for two cytidine chromophores is shown in Fig. 5. As the rotation angle increases, we observe that the values of dipolar coupling computed according to

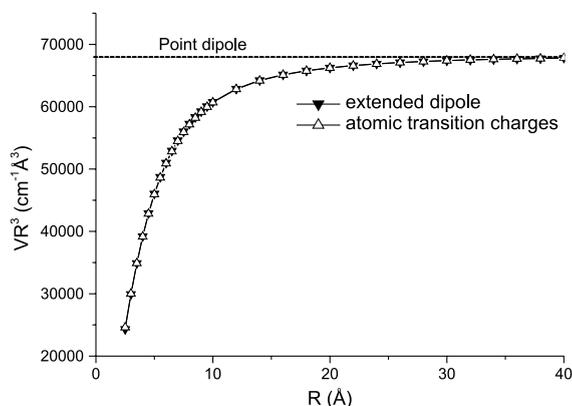


Fig. 4. Dipolar coupling  $V$  between the  $S_0 \rightarrow S_1$  transition moments of two stacked and parallel thymidine chromophores as a function of the stacking distance  $R$  calculated according to the point dipole (dashed line), the extended dipole (filled triangles) and the atomic transition charges (empty triangles) formalisms.

the extended dipole model diverge from those taken as reference. The dipolar interaction at 36°, typical of B-DNA, exhibits an error of  $\approx 100\%$ . Similar inaccuracies have been noted for other nucleoside dimers.

We have shown earlier that the extended dipole model yields quite similar results to atomic transition monopoles for triphenylene derivatives stacked in columnar phases and allowed to rotate around the column axis [23]. In this case, the chromophores have a centre of symmetry which coincides with the rotation axis. By contrast, chromophores in double helices are devoid of an inversion centre and are rotated around an off-centred axis. Hence, the extended dipole cannot account for varied relative positions of the monomers, due to its lack of theoretical foundations.

Given the insufficiencies of the point dipole and extended dipole models, we apply the atomic

Table 4

Extended dipole length  $l$  (Å) and absolute value of charges  $q$  (e) associated with the electronic transitions of nucleosides

	Adenosine		Thymidine	Guanosine		Cytidine
	$S_0 \rightarrow S_1$	$S_0 \rightarrow S_2$	$S_0 \rightarrow S_1$	$S_0 \rightarrow S_1$	$S_0 \rightarrow S_2$	$S_0 \rightarrow S_1$
$q$	0.15	0.16	0.19	0.15	0.16	0.28
$l$	2.32	4.63	4.10	4.72	4.24	2.53

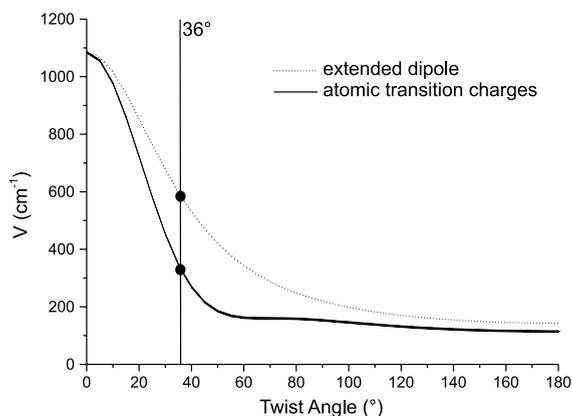


Fig. 5. Dipolar coupling  $V$  between the  $S_0 \rightarrow S_1$  transition moments of two cytidine chromophores located in parallel planes distant of 3.4 Å as a function of the twist angle calculated according to the extended dipole (dotted line) and the atomic transition charge (continuous line) formalisms. The full circles denote the coupling corresponding to a twist angle of 36°, typical of B-DNA.

transition charges model to the computation of coupling elements between neighbouring bases in a standard B-DNA geometry.

Coupling elements between consecutive intra-strand bases and interstrand Watson–Crick base pairs are compiled in Tables 5 and 6, respectively. They range, in absolute value, from 50 to 500  $\text{cm}^{-1}$ , with a mean value close to 250  $\text{cm}^{-1}$ ; interstrand Watson–Crick and intrastrand nearest-neighbour coupling elements exhibit matching values.

The average magnitude of these elements gives interesting hints on the localised or delocalised behaviour of excitations in a double-stranded DNA fragment. For example, the first transitions of guanosine and cytidine are separated by 170  $\text{cm}^{-1}$ ; whereas coupling elements between these transitions amount to 130 and 300  $\text{cm}^{-1}$  for intra-strand and interstrand geometries, respectively. As a consequence, excited states of  $(\text{dGdC})_n$ ,  $(\text{dGdC})_n$  tracts are expected to be rather delocalised on, at least, these two monomers. Adenosine and thymidine  $S_0 \rightarrow S_1$  transitions are distant of 790  $\text{cm}^{-1}$ , whereas inter and intra-strand coupling elements are roughly 200  $\text{cm}^{-1}$ . Broadening of the bands around the monomer transitions may grossly be considered equal to four times the coupling (800  $\text{cm}^{-1}$ ), so that the upper limit of the

Table 5

Dipolar coupling  $V$  ( $\text{cm}^{-1}$ ) calculated using the atomic transition charges for intrastrand nearest neighbours; rise 3.4 Å; twist 36°

Transition 1	Transition 2	$V$
dA $S_0 \rightarrow S_1$	dA $S_0 \rightarrow S_1$	170
	dA $S_0 \rightarrow S_2$	93
	dT $S_0 \rightarrow S_1$	148
	dG $S_0 \rightarrow S_1$	-323
	dG $S_0 \rightarrow S_2$	-106
	dC $S_0 \rightarrow S_1$	86
dA $S_0 \rightarrow S_2$	dA $S_0 \rightarrow S_1$	256
	dA $S_0 \rightarrow S_2$	405
	dT $S_0 \rightarrow S_1$	-68
	dG $S_0 \rightarrow S_1$	-202
	dG $S_0 \rightarrow S_2$	-401
	dC $S_0 \rightarrow S_1$	418
dT $S_0 \rightarrow S_1$	dA $S_0 \rightarrow S_1$	231
	dA $S_0 \rightarrow S_2$	543
	dT $S_0 \rightarrow S_1$	217
	dG $S_0 \rightarrow S_1$	68
	dG $S_0 \rightarrow S_2$	-500
	dC $S_0 \rightarrow S_1$	215
dG $S_0 \rightarrow S_1$	dA $S_0 \rightarrow S_1$	-46
	dA $S_0 \rightarrow S_2$	180
	dT $S_0 \rightarrow S_1$	-451
	dG $S_0 \rightarrow S_1$	380
	dG $S_0 \rightarrow S_2$	-177
	dC $S_0 \rightarrow S_1$	133
dG $S_0 \rightarrow S_2$	dA $S_0 \rightarrow S_1$	-245
	dA $S_0 \rightarrow S_2$	-380
	dT $S_0 \rightarrow S_1$	48
	dG $S_0 \rightarrow S_1$	227
	dG $S_0 \rightarrow S_2$	375
	dC $S_0 \rightarrow S_1$	-409
dC $S_0 \rightarrow S_1$	dA $S_0 \rightarrow S_1$	14
	dA $S_0 \rightarrow S_2$	264
	dT $S_0 \rightarrow S_1$	-90
	dG $S_0 \rightarrow S_1$	127
	dG $S_0 \rightarrow S_2$	-214
	dC $S_0 \rightarrow S_1$	329

Table 6

Dipolar coupling  $V$  ( $\text{cm}^{-1}$ ) calculated using the atomic transition charges for Watson–Crick base pairs

Transition 1	Transition 2	$V$
dT $S_0 \rightarrow S_1$	dA $S_0 \rightarrow S_1$	248
	dA $S_0 \rightarrow S_2$	266
dC $S_0 \rightarrow S_1$	dG $S_0 \rightarrow S_1$	306
	dG $S_0 \rightarrow S_2$	-243

adenosine  $S_0 \rightarrow S_1$  and the lower limit of the thymidine  $S_0 \rightarrow S_1$  bands are almost degenerate. Hence, the localised or delocalised behaviour of an

excitation in  $(dAdT)_n \cdot (dAdT)_n$  sequences is more difficult to predict, and requires taking into account not only nearest-neighbours matrix elements but all other possible interactions.

## 5. Exciton states of double-stranded helices

In this section, we focus on two double-stranded helices  $(dA)_{20} \cdot (dT)_{20}$  and  $(dAdT)_{10} \cdot (dAdT)_{10}$ . The properties of their exciton states are determined using the atomic transition charges model. The exciton matrix is built as described in Section 2 considering all possible interactions and taking into account the  $S_1$  and  $S_2$  states of adenosine and  $S_1$  state of thymidine. Diagonalisation of the matrix yields 60 eigenstates whose oscillator strengths and delocalisation behaviour are discussed.

Fig. 6 shows the oscillator strengths associated with the eigenstates of each one of the double helices studied. As can be seen for both oligonucleotides, the global oscillator strength is smeared over a small number of states which are globally distributed into two narrow bands for  $(dA)_{20} \cdot (dT)_{20}$  (around 37 750 and 39 750  $\text{cm}^{-1}$ ) and three narrow bands for  $(dAdT)_{10} \cdot (dAdT)_{10}$  (around 37 600, 38 750 and 39 500  $\text{cm}^{-1}$ ). It is noteworthy that the most intense transition of each system shares the same eigenstate index  $\langle 57 \rangle$ . A general blue shift of the transition energies compared to those of the individual molecules is observed, with the high-energy region of the spectra bearing the major part of the intensity. Low-energy states are associated with negligible oscillator strength.

Since more than one chromophore transition has been considered to build the exciton matrix, delocalisation of the eigenstates may be subdivided into two categories: transition-governed delocalisation (e.g., delocalisation over a specific transition of one type of chromophore), and spatial delocalisation.

Fig. 7 presents the contribution of each of the three types of chromophore transitions to the eigenstates of the double helices, for which three energy regions can be distinguished. Eigenstates with low energies are mostly localised on the first transition of adenosine. The first transition of thymidine is the major contributor to eigenstates

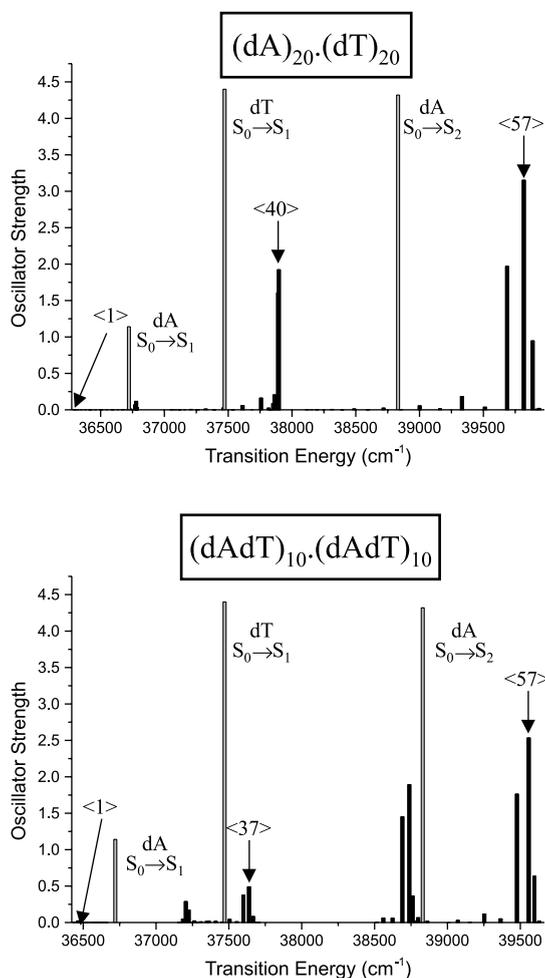


Fig. 6. Oscillator strengths associated with the eigenstates of the double-stranded eicosamers  $(dA)_{20} \cdot (dT)_{20}$  and  $(dAdT)_{10} \cdot (dAdT)_{10}$  as a function of the corresponding energies. The grey columns designate the position of the transitions of individual chromophores. The arrows denote the eigenstates whose topography is shown in Fig. 9.

in the second region. Finally, high-energy eigenstates are mainly built on the second transition of adenosine. Borders between these three regions are clearly defined, so that contributions from a given transition can vary tremendously from one eigenstate to the next in the corresponding energy interval. This is particularly clear for  $(dA)_{20} \cdot (dT)_{20}$ ;  $(dAdT)_{10} \cdot (dAdT)_{10}$  shows a similar, albeit somewhat more complex, trend. This base sequence effect may be related to differences

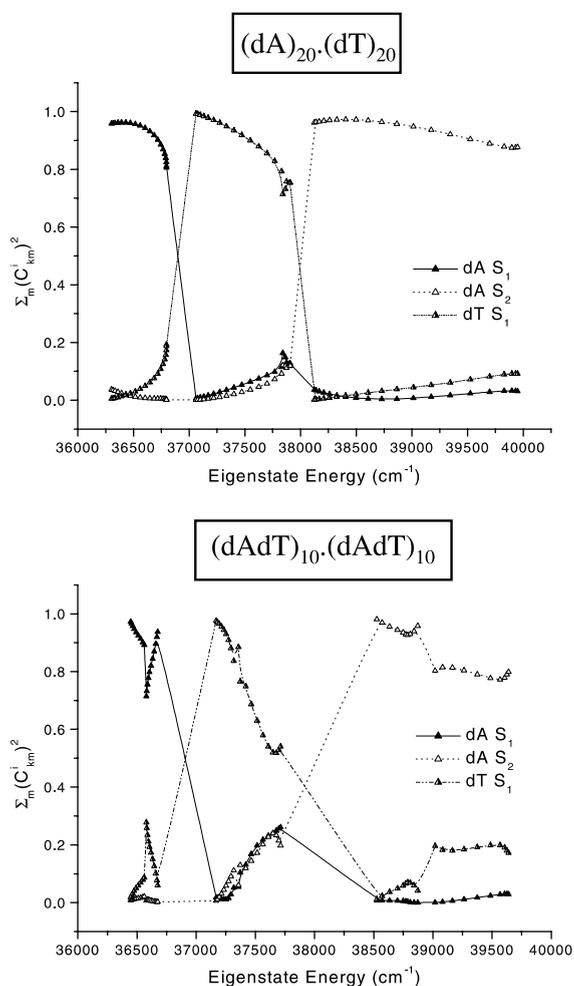


Fig. 7. Contribution of the  $S_1$  state of thymidine and the  $S_1$  and  $S_2$  states of adenosine to the eigenstates of the of double-stranded eicosamers  $(dA)_{20} \cdot (dT)_{20}$  and  $(dAdT)_{10} \cdot (dAdT)_{10}$  as a function the eigenstate energy.

between corresponding matrix elements: in  $(dA)_{20} \cdot (dT)_{20}$  transition-governed delocalisation occurs through intrastrand nearest-neighbour coupling, while in  $(dAdT)_{10} \cdot (dAdT)_{10}$  it rests on the interstrand coupling between molecules in consecutive Watson–Crick base pairs. The latter type of coupling ( $\approx 100 \text{ cm}^{-1}$ ), although smaller than the former (around  $250 \text{ cm}^{-1}$ ), is still large enough to induce delocalisation of the excitation. Hence, our model predicts excitation delocalisation in both oligonucleotides to be governed by

the type of monomer transition. This implies localisation of the excitation on one strand in  $(dA)_{20} \cdot (dT)_{20}$  and on the two strands in  $(dAdT)_{10} \cdot (dAdT)_{10}$ . A simplified explanation of this localisation behaviour may be attempted by considering the three exciton bands associated with each of the three chromophore transitions separately. For both systems examined, these bands do not overlap and the difference between their centres is larger than the maximum coupling values between the two corresponding transitions.

The plot of the participation ratio  $1/L_k$ , denoting the number of coherently coupled chromophores, versus the eigenstate index  $k$ , is presented in Fig. 8. We can see that the  $1/L_k$  values found for  $(dAdT)_{10} \cdot (dAdT)_{10}$  lie between 14 and 24, whereas those calculated for  $(dAdT)_{10} \cdot (dAdT)_{10}$  range between 12 and 28, showing a considerable spatial extent of all eigenstates. It has been shown [71], in the case of helical columnar aggregates consisting of  $n$  identical chromophores, that the maximum value of the normalised participation ratio  $1/nL_k$  is 0.7. In our case, the arrangement of one type of nucleoside within the double helix can be viewed as a one-dimensional aggregate. Hence, perfect spatial delocalisation over this system is expected to be characterised by a participation index of 14 ( $0.7 \times 20$ ). The fact that the majority of the eigenstates show participation ratios larger than 14 proves that they extend over both adenosine and thymidine chromophores. This is also quite apparent in Fig. 8: eigenstates located at the frontier between two regions (as appearing in Fig. 7) show a higher participation ratio, indicating contributions from more than one type of monomer transition. Mixing between adenosine and thymidine transitions in  $(dAdT)_{10} \cdot (dAdT)_{10}$  appears more pronounced than in  $(dA)_{20} \cdot (dT)_{20}$ , with frontier eigenstates presenting participation ratios up to 29.

Finally, we focus on the delocalisation properties of a few remarkable eigenstates of the two oligomers, each belonging to one of the three energy regions depicted in Fig. 7. The topography of these eigenstates, e.g., the relative contribution of individual chromophores, is schematised in Fig. 9. The contribution of a given chromophore is rep-

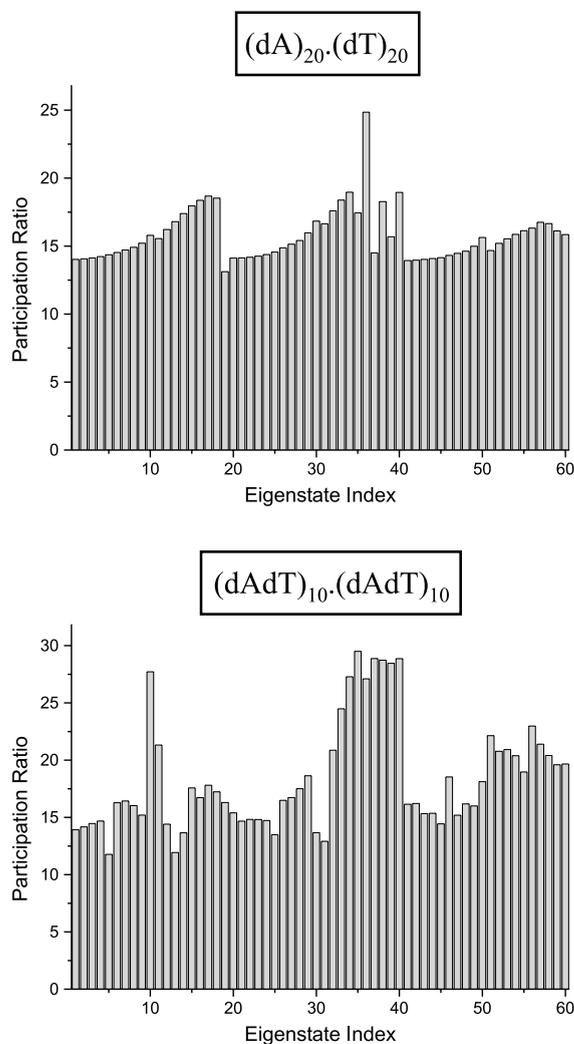


Fig. 8. Participation ratio of the eigenstates of the double-stranded eicosamers  $(dA)_{20} \cdot (dT)_{20}$  and  $(dAdT)_{10} \cdot (dAdT)_{10}$  as a function the eigenstate index.

resented upwards or downwards, depending on the strand to which it belongs.

The first eigenstate, important to fluorescence studies, is concentrated almost exclusively on the first transition of the adenosine chromophores for both double-stranded DNA fragments. This implies a delocalisation on a single strand for  $(dA)_{20} \cdot (dT)_{20}$  and on the two strands for  $(dAdT)_{10} \cdot (dAdT)_{10}$ . Similarly, eigenstate  $\langle 57 \rangle$ , bearing the highest oscillator strength in both systems, is mostly based on the adenosine  $S_0 \rightarrow S_2$  transition,

although contributions from thymidine are no longer negligible. Finally, the mixing of the excited states of adenosine and thymidine is more substantial in eigenstates  $\langle 40 \rangle$  of  $(dA)_{20} \cdot (dT)_{20}$  and  $\langle 37 \rangle$  of  $(dAdT)_{10} \cdot (dAdT)_{10}$ , which are also characterised by important oscillator strengths. All six eigenstates are distributed over the whole extent of the double helix.

## 6. Summary and comments

In the present communication, we have discussed problems related to the description of the electronic excited states and excitation transport in DNA double helices. In particular, we have pointed out how such studies can be improved using presently available experimental and computational techniques and the new methodologies which may result from their combination. The work reported here is our first step towards establishing a deeper understanding of phenomena which have significant biological interest.

The starting point of our study was the determination of the energies and transition moments of the lowest dipolar transitions of the individual chromophores forming the double helices. To this end, we recorded the absorption spectra of dilute ( $10^{-5}$  M) water solutions of the four standard deoxynucleosides and decomposed them on the basis of fluorescence anisotropy and circular dichroism measurements reported in the literature. In parallel, we calculated these properties with the CS-INDO-CIPSI method. The agreement between the experimental and computed data was satisfactory for all the electronic transitions considered, i.e.  $S_0 \rightarrow S_1$  for cytidine and thymidine,  $S_0 \rightarrow S_1$  and  $S_0 \rightarrow S_2$  for adenosine and guanosine.

Subsequently, we calculated the dipolar coupling for some typical stacked and paired dimers, using the atomic transition charge distribution model. We showed that neither the point dipole nor the extended dipole approximations provide a correct description of the dipolar coupling within double helices.

Finally, we calculated the properties of the exciton states of two particular helical oligonucleo-

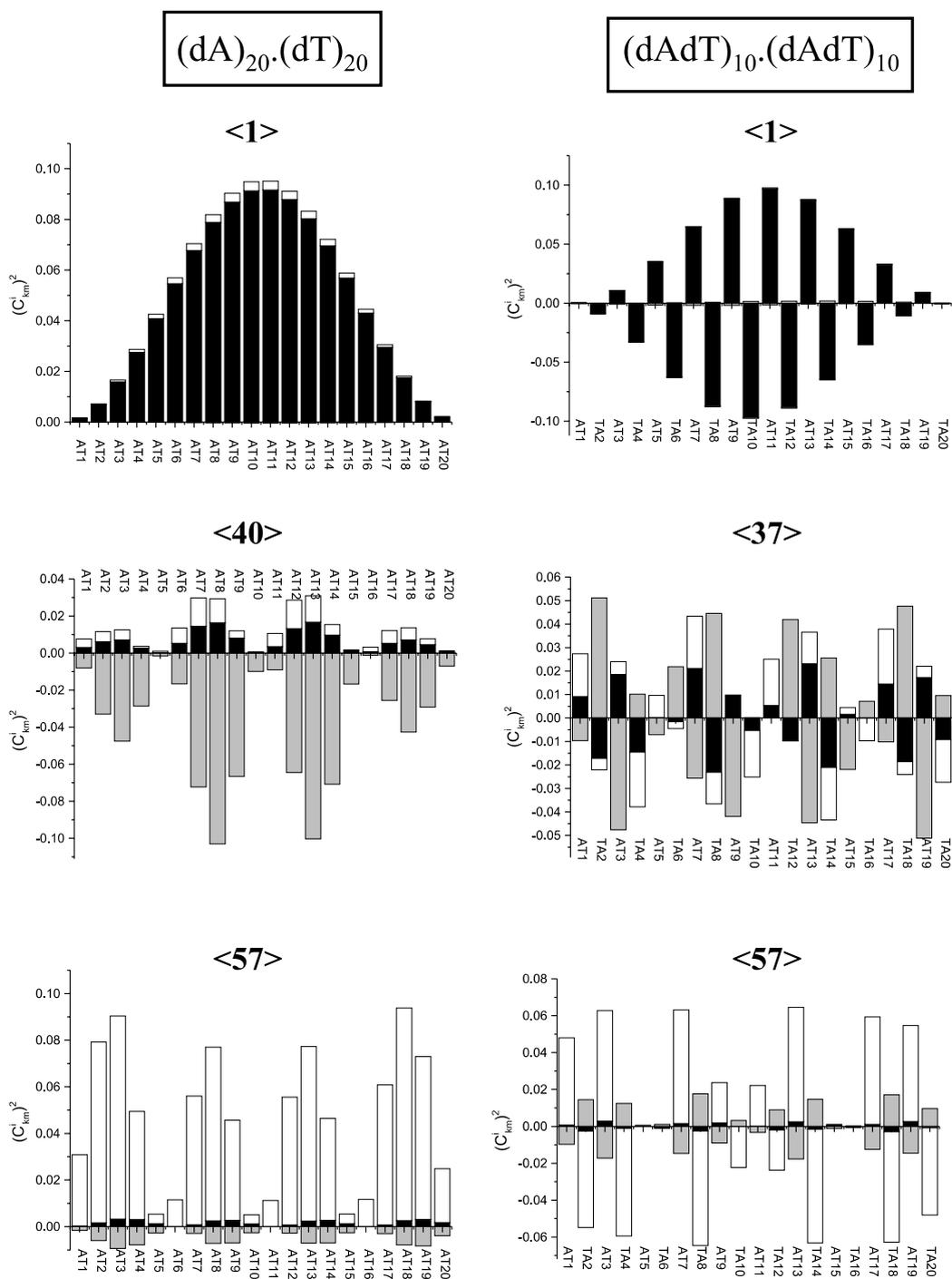


Fig. 9. Topography of three typical eigenstates of the double-stranded eicosamers  $(dA)_{20} \cdot (dT)_{20}$  and  $(dAdT)_{10} \cdot (dAdT)_{10}$ : the lowest in energy ( $\langle 1 \rangle$ ), those bearing the highest oscillator strength ( $\langle 57 \rangle$ ) and two intermediate ones ( $\langle 40 \rangle$  and  $\langle 37 \rangle$ ). The coefficients  $(C^i_{k,m})^2$  represent the contribution of chromophore  $m$  in its  $i$ th excited state ( $S_1$  of adenosine in black,  $S_2$  of adenosine in white,  $S_1$  of thymidine in grey) to eigenstate  $k$ . The upper and lower parts of each histogram refer to chromophores located on each of the two strands.

tides,  $(dA)_{20} \cdot (dT)_{20}$  and  $(dAdT)_{10} \cdot (dAdT)_{10}$ , considering three closely lying molecular electronic transitions, two for adenosine and one for thymidine. We found that the oscillator strength is distributed over a small number of eigenstates. All eigenstates are spatially delocalised over the whole length of the double helix. Most of them are built mainly on a single molecular state, but some, in particular in  $(dAdT)_{10} \cdot (dAdT)_{10}$ , result from the mixing of all three molecular states considered. A more thorough mixing of the excited states of the two nucleosides is expected in  $(dG)_n \cdot (dC)_n$  and  $(dGdC)_n \cdot (dGdC)_n$  tracts, since individual transitions of guanosine and cytidine lie much closer to each other than those of adenosine and thymidine.

The most important conclusion of our numerical calculations is that dipolar coupling alone may lead to a spatial delocalisation of the excitation within double helices having an idealised B-DNA geometry.

Prior to any comparison with experimental observations, we must take three important effects into account, which could notably reduce the spatial delocalisation of the electronic excitation found in the present study. Firstly, we need to model the site-dependent excitation energies originating from variations of the permanent dipole moment between the ground and the excited states [38]. Secondly, we must consider the plasticity of the double strands, by modelling solvent and temperature effects. We intend to reach both these goals by combining quantum chemistry calculations and molecular dynamics simulations both in the ground and in the excited states [32]. Thirdly, we have to include interactions due to orbital overlap, in particular the charge resonance and charge transfer terms, for which we are currently developing a new methodology. In parallel, we are also pursuing our experimental studies, making a particular effort to work only with well-defined molecular systems.

## Acknowledgements

The authors wish to express their gratitude to Dr. Richard Lavery for encouragement and fruitful discussions.

## References

- [1] (a) D.E. Brash, W.A. Haseltine, *Nature* 298 (1982) 189; (b) J. Cadet, M. Berger, T. Douki, B. Morin, S. Raoul, J.-L. Ravanat, S. Spinelli, *Biol. Chem.* 378 (1997) 1275.
- [2] (a) C. Chaozhi, T. Fiebig, O. Schiemann, J.K. Barton, A.H. Zewail, *Proc. Natl. Acad. Sci.* 97 (2000) 14052; (b) A.A. Voityuk, J. Jortner, M. Bixon, N. Rösch, *J. Chem. Phys.* 114 (2001) 5614, and references therein.
- [3] M. Guéron, J. Eisinger, R.G. Shulman, *J. Chem. Phys.* 47 (1967) 4077.
- [4] D.M. Rayner, A.G. Szabo, R.O. Loutfy, R.W. Yip, *J. Phys. Chem.* 84 (1980) 289.
- [5] S. Georghiou, S. Zhu, R. Weidner, C.-R. Huang, G. Ge, *J. Biomol. Struct. Dyn.* 8 (1990) 657.
- [6] G. Ge, S. Georghiou, *Photochem. Photobiol.* 54 (1991) 301.
- [7] S. Georghiou, G.R. Phillips, G. Ge, *Biopolymers* 32 (1992) 1417.
- [8] C.-R. Huang, S. Georghiou, *Photochem. Photobiol.* 56 (1992) 95.
- [9] I. Tinoco Jr., *J. Am. Chem. Soc.* 82 (1960) 4785.
- [10] W. Rhodes, *J. Am. Chem. Soc.* 83 (1961) 3609.
- [11] I. Tinoco Jr., R.W. Woody, D.F. Bradley, *J. Chem. Phys.* 38 (1963) 1317.
- [12] T. Miyata, S. Yomosa, *J. Phys. Soc. Jpn.* 27 (1969) 720.
- [13] J. Eisinger, R.G. Shulman, *Science* 161 (1968) 1311.
- [14] T. Förster, *Disc. Faraday Soc.* 27 (1959) 7.
- [15] G.H. Beaven, E.R. Holiday, E.A. Johnson, in: E. Chargaff, J.N. Davidson (Eds.), *The Nucleic Acids*, vol. 1, Academic Press, New York, 1955, p. 493.
- [16] D. Voet, W.B. Gratzer, R.A. Cox, P. Doty, *Biopolymers* 1 (1963) 193.
- [17] R.W. Wilson, P.R. Callis, *J. Phys. Chem.* 80 (1976) 2280.
- [18] P.R. Callis, *Ann. Rev. Phys. Chem.* 34 (1983) 329.
- [19] A. Holmén, A. Broo, B. Albinsson, B. Nordén, *J. Am. Chem. Soc.* 119 (1997) 12240.
- [20] T.M. Nordlund, D. Xu, K.O. Evans, *Biochemistry* 32 (1993) 12090.
- [21] D. Xu, T.M. Nordlund, *Biophys. J.* 78 (2000) 1042.
- [22] G.D. Scholes, K.P. Ghiggino, *J. Phys. Chem.* 98 (1994) 4580.
- [23] S. Marguet, D. Markovitsi, P. Millié, H. Sigal, S. Kumar, *J. Phys. Chem. B* 102 (1998) 4697.
- [24] R. Silbey, *Ann. Rev. Phys. Chem.* 27 (1976) 203.
- [25] G.D. Scholes, R.D. Harcourt, G.R. Fleming, *J. Phys. Chem. B* 101 (1997) 7302.
- [26] D. Markovitsi, S. Marguet, L.K. Gallos, H. Sigal, P. Millié, P. Argyrakis, H. Ringsdorf, S. Kumar, *Chem. Phys. Lett.* 306 (1999) 163.
- [27] M. Guéron, R.G. Shulman, J. Eisinger, *Proc. Natl. Acad. Sci.* 56 (1966) 814.
- [28] T. Montenay-Garestier, C. Hélène, A.M. Michelson, *Biochim. Biophys. Acta* 182 (1969) 342.
- [29] J.P. Morgan, M. Daniels, *Photochem. Photobiol.* 31 (1980) 207.

- [30] J.P. Ballini, P. Vigny, M. Daniels, *Biophys. Chem.* 18 (1983) 61.
- [31] S. Georghiou, S.M. Kubala, C.C. Large, *Photochem. Photobiol.* 67 (1998) 526.
- [32] B. Hartmann, R. Lavery, *Q. Rev. Biophys.* 29 (1996) 309.
- [33] P. Vigny, *C.R. Acad. Sc. Paris t272* (1971) 3206.
- [34] D.N. Nikogosyan, D. Angelov, B. Soep, L. Lindqvist, *Chem. Phys. Lett.* 252 (1996) 322.
- [35] T. Häupl, C. Windolph, T. Jochum, O. Brede, R. Hermann, *Chem. Phys. Lett.* 280 (1997) 520.
- [36] A. Reuther, H. Iglev, R. Laenen, A. Laubereau, *Chem. Phys. Lett.* 325 (2000) 360.
- [37] J.-M.L. Pecourt, J. Peon, B. Kohler, *J. Am. Chem. Soc.* 122 (2000) 9348.
- [38] C. Ecoffet, D. Markovitsi, P. Millié, J.P. Lemaistre, *Chem. Phys.* 177 (1993) 629.
- [39] D. Markovitsi, S. Marguet, J. Bondkowski, S. Kumar, *J. Phys. Chem. B* 105 (2001) 1299.
- [40] A.A. Voityuk, N. Rösch, M. Bixon, J. Jortner, *J. Phys. Chem. B* 104 (2000) 9740.
- [41] V. Czikkely, H.D. Forsterling, H. Kuhn, *Chem. Phys. Lett.* 6 (1970) 207.
- [42] D.B. Siano, D.E. Metzler, *J. Chem. Phys.* 51 (1969) 1856.
- [43] J. Frenkel, *Phys. Rev.* 37 (1931) 1276.
- [44] A.S. Davydov, *Theory of Molecular Excitons*, Plenum Press, New York, 1971.
- [45] E.I. Rashba, M.D. Sturge, in: V.M. Agranovitch, A.A. Maradudin (Eds.), *Excitons*, North-Holland, Amsterdam, 1982.
- [46] H.C. Longuet-Higgins, *Proc. R. Soc. London A* 235 (1956) 537.
- [47] P. Claverie, in: B. Pullman (Ed.), *Intermolecular Interactions—From Diatomics to Biopolymers*, Wiley, New York, 1978, p. 69.
- [48] J.A. Pople, D.L. Beveridge, P.A. Dobosh, *J. Chem. Phys.* 47 (1967) 2026.
- [49] P. Dean, *Rev. Mod. Phys.* 44 (1972) 127.
- [50] M. Schreiber, Y. Toyozawa, *J. Phys. Soc. Jpn.* 51 (1982) 1537.
- [51] M. Weissbluth, *Q. Rev. Biophys.* 4 (1971) 1.
- [52] E. Brown, E.S. Pysh, *J. Chem. Phys.* 56 (1972) 31.
- [53] A. Broo, A. Holmén, *J. Phys. Chem. A* 101 (1997) 3589.
- [54] O. Kennard, W.N. Hunter, in: W. Saenger (Ed.), *Nucleic Acids*, Landolt-Börnstein New Series, vol. VII/1a, Springer, Berlin, 1989, p. 255.
- [55] R.E. Dickerson, M. Bansal, C.R. Calladine, S. Diekmann, W.N. Hunter, O. Kennard, R. Lavery, H.C.M. Nelson, W.K. Olson, W. Saenger, Z. Shakked, H. Sklenar, D.M. Soumpasis, C.-S. Tung, E. von Kitzing, A.H.-J. Wang, V.B. Zhurkin, *J. Mol. Biol.* 205 (1989) 787.
- [56] S. Arnott, E. Selsing, *J. Mol. Biol.* 88 (1974) 509.
- [57] M.A. Viswamitra, Z. Shakked, P.G. Jones, G.M. Sheldrick, S.A. Salisbury, O. Kennard, *Biopolymers* 21 (1982) 513.
- [58] A. Germain, P. Millié, *Chem. Phys.* 219 (1997) 265.
- [59] H. DeVoe, I. Tinoco, *J. Mol. Biol.* 4 (1962) 518.
- [60] R.F. Stewart, N. Davidson, *J. Chem. Phys.* 39 (1963) 255.
- [61] L.B. Clark, *J. Am. Chem. Soc.* 99 (1977) 3934.
- [62] P.R. Callis, W.T. Simpson, *J. Am. Chem. Soc.* 92 (1970) 3593.
- [63] A. Volosov, *Int. J. Quant. Chem.* 36 (1989) 473.
- [64] A. Volosov, R.W. Woody, *J. Phys. Chem.* 96 (1992) 4845.
- [65] D. Theiste, P.R. Callis, R.W. Woody, *J. Am. Chem. Soc.* 113 (1991) 3260.
- [66] J.D. Petke, G.M. Maggiora, R.E. Christoffersen, *J. Am. Chem. Soc.* 112 (1990) 5452.
- [67] J. Lorentzon, M.P. Fülcher, B.O. Roos, *J. Am. Chem. Soc.* 117 (1995) 9265.
- [68] B.O. Roos, M.P. Fülcher, P.Å. Malmqvist, M. Merchán, L. Serrano-Andrés, in: S.R. Langhoff (Ed.), *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, Kluwer, Dordrecht, 1995, p. 420.
- [69] M.P. Fülcher, B.O. Roos, *J. Am. Chem. Soc.* 117 (1995) 2089.
- [70] T. Miyata, H. Suzuki, S. Yomosa, *J. Phys. Soc. Jpn.* 25 (1968) 1428.
- [71] D. Markovitsi, L.K. Gallos, J.P. Lemaistre, P. Argyrakos, *Chem. Phys.* 269 (2001) 147.



## 4. États Excités Collectifs de Fragments d'ADN : Approche Dynamique

### 4.1. Introduction

L'aspect dynamique de la caractérisation des états propres de doubles brins d'ADN de type poly-dA et poly-(dAdT), qui faisait défaut à la précédente étude, a été abordé sur des géométries hélicoïdales réalistes tirées de simulations de dynamique moléculaire, en tenant compte cette fois du désordre diagonal et des effets de contre-ions et de solvant lors de la constitution de la matrice excitonique. Les résultats obtenus aux deux niveaux de simulation pourront ainsi être comparés.

### 4.2. Paramétrage du modèle électrostatique

#### *4.2.1. Choix des méthodes*

L'emploi du champ de force AMBER<sup>255,256</sup> comme base du calcul de l'énergie d'interaction au sein des doubles brins à l'étude impose de n'avoir recours, pour retranscrire les propriétés électrostatiques des chromophores, qu'à un jeu de monopôles atomiques. Si ceux-ci existent pour l'état fondamental des bases de l'ADN, ils restent à déterminer pour les états excités de ces molécules. Il convient pour cela de distribuer l'information électrostatique obtenue par un calcul *ab initio* sur les différents centres d'interaction des chromophores, de la manière la plus rigoureuse possible. En effet, le rôle joué par les moments d'ordre supérieur aux courtes distances d'approche entre entités en interaction (telles que celles rencontrées entre sites interbrins voisins dans l'ADN) peut se révéler capital ; il conviendra dès lors d'inclure un nombre suffisant de moments dans le modèle électrostatique développé.

Plusieurs méthodes sont disponibles pour extraire du calcul *ab initio* un modèle électrostatique. Le choix le plus judicieux serait d'effectuer un développement multipolaire multicentrique à partir d'une analyse directe de la fonction d'onde. Un tel développement, mené jusqu'au quadrupôle sur les atomes et milieux de liaisons, a déjà été présenté et mis en

œuvre dans ce travail. Il ne peut en revanche être utilisé ici puisque, pour des raisons de coût en temps de calcul liées à la taille des systèmes traités, le champ de force AMBER se limite à une distribution de charges ponctuelles atomiques.

Deux approches sont dès lors envisageables pour extraire du calcul *ab initio* cette distribution de monopôles : une analyse simplifiée de la fonction d'onde, ou l'exploitation d'une propriété prédite par la fonction d'onde, en l'occurrence le potentiel électrostatique.

La première méthode consiste à tronquer le développement multipolaire précédemment évoqué au terme monopolaire. Un calcul *ab initio* de bonne qualité nécessite l'utilisation de fonctions de base de nombre quantique orbital supérieur à zéro (fonctions p, d...), qui présentent intrinsèquement des moments supérieurs au monopôle. Les orbitales moléculaires exploitées lors du développement multipolaire sont construites sur ces fonctions de base et impliquent en conséquence des dipôles et quadripôles atomiques généralement non négligeables. Il sera montré par la suite que les erreurs ainsi commises peuvent avoisiner les 100 % aux distances intermoléculaires typiques.

La seconde méthode, de loin la plus précise, permet d'inclure de manière implicite l'influence des moments d'ordre supérieur dans la distribution monopolaire. Elle consiste à ajuster les paramètres de la distribution de charges atomiques de façon à retranscrire le potentiel électrostatique calculé à partir de la fonction d'onde.

#### 4.2.2. Méthode Eigenvalue Analysis

L'algorithme (dénommé *Eigenvalue Analysis*) employé pour cet ajustement est une extension de la méthode RESP (*Restrained Electrostatic Potential*)<sup>257</sup> qui repose sur la minimisation contrainte de la fonction coût suivante :

$$C = \sum_{i=1}^N (V^0(P_i) - V(P_i))^2$$

dont le calcul nécessite la connaissance du potentiel électrostatique de référence  $V^0$  et celle du potentiel électrostatique  $V$  généré par la distribution de charges en cours d'ajustement, sur une grille de  $N$  points  $\{P_i, i \in [1..N]\}$ . Le potentiel électrostatique créé en  $P_i$  par la distribution monopolaire est calculé selon :

$$V(P_i) = \sum_{j=1}^n \frac{q_j}{r_{ij}}$$

#### 4. États Excités Collectifs de Fragments d'ADN : Approche Dynamique

où les charges atomiques  $\{q_j, j \in [1..n]\}$  sont les paramètres du problème, et  $r_{ij}$  la distance de la charge  $j$  au point  $P_i$ . La minimisation de  $C$  revient à résoudre un système d'équations linéaires qui peut s'écrire sous forme matricielle :

$$\mathbf{A}\vec{q} = \vec{b}$$

où  $\vec{q}$  est le vecteur  $\vec{q} = (q_i, i \in [1..n])$ . Il a été montré que la matrice  $\mathbf{A}$  est rendue quasi-singulière<sup>258</sup> par le fait que le nombre de paramètres disponibles est supérieur au nombre de paramètres effectivement liés au problème d'optimisation. Par exemple, la charge portée par un atome central au sein d'une molécule de grande taille aura peu d'effet sur le potentiel en un point  $P_i$  dont elle est systématiquement éloignée : cet atome est dit 'caché'<sup>259</sup> et sa charge est mal déterminée (elle peut prendre une large gamme de valeurs sans affecter pour autant la qualité de l'ajustement). De même, le potentiel électrostatique en un point  $P_i$  situé relativement loin d'une molécule de charge globale nulle est quasiment égal au potentiel que produirait le premier moment non nul de la molécule en ce point ; un dipôle peut être modélisé par deux charges, un quadrupôle par quatre, etc..., de sorte que tout paramètre supplémentaire est redondant et par conséquent indéfini.

La présente méthode pallie cette indétermination en s'appuyant sur un jeu de charges de référence choisies *a priori*. Soit  $\mathbf{P}$  la matrice du projecteur sur l'espace sous-tendu par les vecteurs propres de  $\mathbf{A}$  associés aux valeurs propres inférieures à un seuil  $\varepsilon$ , et  $\mathbf{Q}$  telle que  $\mathbf{Q} = \mathbf{1} - \mathbf{P}$ . En conséquence,  $\mathbf{P}\vec{q} + \mathbf{Q}\vec{q} = \vec{q}$ . Il peut être prouvé<sup>257</sup> que la fonction  $C$  ne dépend, aux termes en  $O(\varepsilon)$  près, que de  $\mathbf{P}\vec{q}$  et non de  $\mathbf{Q}\vec{q}$ . Il vient alors :

$$\vec{q} = \mathbf{P}(\mathbf{P}\mathbf{A}\mathbf{P} + \mathbf{Q})^{-1}\vec{b} + \mathbf{Q}\vec{q}_0$$

où  $\vec{q}_0$  est le vecteur constitué des charges de référence. Un certain nombre de contraintes sont imposées au cours de la minimisation, par le biais de multiplicateurs de Lagrange. En premier lieu, la somme des charges doit être égale à la charge globale de la molécule. En second lieu, il est possible d'imposer des égalités entre charges, pour rendre compte d'une symétrie moléculaire ou introduire une transférabilité entre fonctions chimiques par exemple. Le choix du paramètre  $\varepsilon$  contrôle le nombre de valeurs propres de  $\mathbf{A}$  qui sont éliminées au cours du processus et conditionne la ressemblance du jeu de charges obtenu avec la référence initialement choisie.

#### 4.2.3. Génération d'une grille de points d'évaluation du potentiel électrostatique

La grille de points  $\{P_i, i \in [1..N]\}$  utilisée pour l'ajustement est générée à l'aide de l'algorithme de Spackman<sup>260</sup>, lui-même adapté de celui de Kollman et Singh<sup>261</sup>. Les points qui la composent sont situés sur une surface constituée par fusion d'un ensemble de sphères, chacune étant centrée sur un atome et ayant pour rayon  $\gamma R$ ,  $R$  étant le rayon de van der Waals de l'atome en question. L'algorithme de Spackman permet un bon échantillonnage d'une telle surface pour tout choix de  $\gamma$ ; de plus, les options de tessellation disponibles permettent de limiter le nombre de points  $N$  à un nombre raisonnable (quelques milliers) même lorsque  $\gamma$  est grand.

#### 4.2.4. Procédure de calcul

La procédure adoptée pour la détermination des charges atomiques est la suivante :

- Calcul CASSCF des états électroniques (fondamental et excités) des molécules (MOLPRO<sup>262</sup>);
- Analyse de Stone<sup>133</sup> du potentiel électrostatique généré par la densité électronique CASSCF. Le développement est tronqué au quadrupôle et restreint aux atomes et milieux de liaisons (MOLPRO) ;
- La grille de points utilisée pour l'ajustement est générée (GAMESS<sup>263</sup>) ;
- Le potentiel électrostatique créé en ces points par la distribution de Stone est évalué ;
- La distribution de charges atomiques recherchée est ajustée de façon à reproduire ce potentiel.

L'étape supplémentaire que constitue le passage par une analyse de Stone en multipôles distribués est liée à des contraintes techniques empêchant le calcul direct du potentiel électrostatique créé par la densité électronique CASSCF. Il sera montré, dans le paragraphe suivant, que l'erreur qu'elle induit est négligeable par rapport à celle commise lors de l'ajustement des charges atomiques par la méthode *Eigenvalue Analysis*.

#### 4.2.5. Validation

La validité de la méthode et les incertitudes accumulées au cours du processus peuvent être évaluées simplement, sur le cas de la molécule d'eau. Deux grilles de points ont

#### 4. États Excités Collectifs de Fragments d'ADN : Approche Dynamique

été utilisées pour ce test : l'une à courte distance ( $\gamma = 1,4 - 1252$  points), l'autre à longue distance ( $\gamma = 10 - 960$  points).

Un calcul SCF de la molécule d'eau en base cc-pVDZ<sup>264</sup> est menée à l'aide du logiciel MOLPRO ; il est suivi d'une analyse de Stone tronquée au quadrupôle et restreinte aux atomes et milieux de liaisons. Le potentiel électrostatique issu de la fonction d'onde (P1) et celui généré par la distribution multipolaire (P2) sont comparés aux points de la grille (tableau 4.1) :

$\gamma$	1,4	10
Erreur relative (%)	0,14	0,07

Tableau 4.1

L'erreur commise est donc minime. Ces valeurs sont d'ailleurs typiques d'un tel développement multipolaire multicentrique<sup>112</sup>.

Les charges atomiques retranscrivant au mieux le potentiel électrostatique P2 sont alors ajustées au moyen de la méthode *Eigenvalue Analysis*, avec  $\varepsilon = 10^{-2.265}$ . Pour des raisons de symétrie, la contrainte d'égalité des charges des deux atomes d'hydrogène est imposée. Le nombre de charges indéterminées pour chaque valeur de  $\gamma$  est (tableau 4.2) :

$\gamma$	1,4	10
Nombre d'indéterminations	0	2

Tableau 4.2

À longue distance, la molécule d'eau se comporte comme un dipôle ponctuel, défini de manière univoque par un unique paramètre ; les deux paramètres restants sont donc indéfinis. Comme mentionné en 4.2.2, l'indétermination peut être levée en ayant recours à des charges de référence.

L'obtention de charges de référence raisonnables n'est cependant pas triviale. Un développement monopolaire direct de la densité associée à la fonction d'onde ne convient pas (elle provoque une erreur moyenne de 109 % sur l'évaluation du potentiel électrostatique à courte distance). L'emploi de charges de Mulliken, quant à lui, dégrade considérablement la

qualité de l'ajustement. Il a donc été décidé d'effectuer l'ajustement sur une surface ne donnant pas lieu à des indéterminations, ici celle associée à  $\gamma = 1,4$ ; le potentiel électrostatique (P3) créé par la distribution de charges est comparé à P2, à courte et longue distances (tableau 4.3) :

$\gamma$	<b>1,4</b>	<b>10</b>
<b>Erreur relative (%)</b>	7,5	1,7

Tableau 4.3

La qualité de l'ajustement du potentiel électrostatique est donc bonne, particulièrement en comparaison du développement monopolaire direct de la densité associée à la fonction d'onde.

#### 4.2.6. Description des états excités

Les calculs CASSCF de l'adénine et de la thymine (groupe ponctuel  $C_s$ ) sont réalisés en base cc-pVDZ<sup>264</sup>. Il s'agit de calculs CAS par états et non moyens. L'espace actif est le système  $\pi$  (12 électrons dans 11 orbitales moléculaires pour l'adénine, 12 électrons dans 10 orbitales pour la thymine). La correction PT2 à l'énergie CASSCF est également déterminée pour chaque état. Les énergies d'excitation sont présentées dans le tableau 4.4, où sont également rappelés les résultats théoriques (CS-INDO-CIPSI) et expérimentaux obtenus précédemment (partie 3.2).

$E_{exc}$ ( $cm^{-1}$ )	CASSCF	CASPT2	CS-INDO	Exp.
<b>Adénine <math>S_0 \rightarrow S_1</math></b>	41500	38800	37200	36700
<b>Adénine <math>S_0 \rightarrow S_2</math></b>	48700	38900	39900	38800
<b>Thymine <math>S_0 \rightarrow S_1</math></b>	42300	36400	36400	37500

Tableau 4.4

Comme mentionné en partie 3.2, les énergies d'excitation calculées aux différents niveaux de théorie diffèrent de manière prononcée. La méthode CASSCF surestime les

#### 4. États Excités Collectifs de Fragments d'ADN : Approche Dynamique

énergies d'excitation expérimentales beaucoup plus que ne le fait la méthode CS-INDO-CIPSI. En revanche, et contrairement à cette dernière, elle prédit correctement l'ordre des trois transitions. La méthode CASPT2 corrige partiellement la surestimation des énergies d'excitation mais confère une quasi-dégénérescence aux deux premiers états excités de l'adénine.

La connaissance des forces d'oscillateur associées aux transitions est une aide précieuse à leur identification. Elles sont présentées dans le tableau 4.5 :

<b>f</b>	<b>CASSCF</b>	<b>CASPT2</b>	<b>CS-INDO</b>	<b>Exp.</b>
<b>Adénine <math>S_0 \rightarrow S_1</math></b>	0,001	0,001	0,057	0,05
<b>Adénine <math>S_0 \rightarrow S_2</math></b>	0,207	0,166	0,216	0,24
<b>Thymine <math>S_0 \rightarrow S_1</math></b>	0,145	0,133	0,220	0,24

Tableau 4.5

Une assez bonne concordance apparaît entre les différentes méthodes théoriques et les résultats expérimentaux. Le caractère peu permis de la transition  $S_0 \rightarrow S_1$  est notamment correctement retranscrit dans tous les cas. L'emploi combiné, dans cette étude, de calculs CASSCF et CS-INDO-CIPSI ne se justifie qu'à condition que les états excités traités par les deux méthodes soient les mêmes ; les résultats des tableaux 4.4 et 4.5 sont autant de bons arguments à cet égard.

##### 4.2.7. Obtention des distributions de charges

La procédure détaillée ci-dessus est appliquée aux états fondamental et excités de l'adénine et de la thymine, à partir de la densité électronique obtenue au niveau CASSCF. Les conditions d'application de la méthode à ces molécules sera précisée sur l'exemple de l'état fondamental de l'adénine.

Il convient ici de bien différencier la grille de points utilisée lors de l'ajustement de celle employée pour le contrôle de la qualité de ce dernier. La première influe directement sur la distribution de charge obtenue ; la seconde n'intervient pas dans le paramétrage mais donne un aperçu des erreurs commises sur le potentiel électrostatique aux distances concernées, lorsque la distribution de charges est employée. Elles seront dénommées  $G_1$  et  $G_2$ ,

respectivement. Le but est de déterminer  $G_1$  de façon à rendre compte, avec une incertitude correcte, du potentiel électrostatique sur une large gamme de distances (c'est-à-dire sur des ensembles  $G_2$  les plus grands et les plus divers possibles).

À cet effet, onze surfaces de valeurs de  $\gamma$  croissantes sont construites ; elles sont notées  $H_\gamma$ . Le nombre de points sur chaque surface est compilé sur le tableau 4.6 :

$\gamma$	<b>1,4</b>	<b>1,6</b>	<b>1,8</b>	<b>2,0</b>	<b>2,2</b>	<b>2,4</b>
<b>Nombre de points</b>	3897	3458	3046	2768	2551	2337
$\gamma$	<b>2,6</b>	<b>2,8</b>	<b>3,0</b>	<b>5,0</b>	<b>10,0</b>	
<b>Nombre de points</b>	2162	1979	1885	1435	1174	

Tableau 4.6

Le jeu de charges de référence adopté pour l'ajustement est constitué des charges de Mulliken obtenues au niveau CASSCF pour l'état fondamental de l'adénine. Le potentiel électrostatique généré par la distribution de charges est comparé à celui engendré par la distribution multipolaire de Stone, dont il a été vu précédemment qu'elle n'induisait que de faibles différences par rapport au potentiel électrostatique *ab initio*. Le moment dipolaire équivalent à la distribution de charges est quant à lui comparé à sa valeur *ab initio*.

La manière la plus intuitive d'obtenir un jeu de charges atomiques valables à toutes distances est d'inclure dans  $G_1$  un grand nombre de points représentatifs de toutes les gammes de distances. Le cas échéant,  $G_1 = \bigcup_{\gamma=1,4}^{10,0} H_\gamma$  ; aucune indétermination ne survient lors du traitement *Eigenvalue Analysis*. L'erreur commise est ensuite évaluée sur chacune des surfaces  $H_\gamma$  individuelles (tableau 4.7) :

$\gamma$	<b>1,4</b>	<b>1,6</b>	<b>1,8</b>	<b>2,0</b>	<b>2,2</b>	<b>2,4</b>
<b>Erreur relative (%)</b>	219	196	174	160	147	136
$\gamma$	<b>2,6</b>	<b>2,8</b>	<b>3,0</b>	<b>5,0</b>	<b>10,0</b>	
<b>Erreur relative (%)</b>	129	124	117	72	44	

Tableau 4.7

#### 4. États Excités Collectifs de Fragments d'ADN : Approche Dynamique

Le constat est accablant : il ne semble pas possible de déduire une distribution de charges qui serait valable dans la gamme de distances considérée. Ceci s'explique par le fait que la nature des interactions varie beaucoup avec la distance : moments d'ordre supérieur et anisotropie marquée à courte distance, dipôle quasi-ponctuel à longue distance... Un unique jeu de charges moyennes ne saurait suffire à rendre compte de cette complexité.

L'ajustement et le test sont à présent réalisés sur la même surface ( $G_1 \equiv G_2 \equiv H_\gamma$  pour toutes les valeurs de  $\gamma$ ). Le nombre d'indéterminations, de même que les erreurs commises sur le potentiel électrostatique et le moment dipôle permanent, pour toutes les valeurs de  $\gamma$ , sont présentées dans le tableau 4.8.

$\gamma$	1,4	1,6	1,8	2,0	2,2	2,4
<b>Potentiel électrostatique –</b>	13,1	12,7	9,7	9,1	19,9	15,8
<b>Erreur relative (%)</b>						
<b>Moment dipolaire –</b>	0,09	1,25	0,32	0,27	3,42	0,94
<b>Erreur relative (%)</b>						
<b>Nombre d'indéterminations (/15)</b>	0	1	2	3	5	5
$\gamma$	2,6	2,8	3,0	5,0	10,0	
<b>Potentiel électrostatique –</b>	13,6	13,2	12,0	12,6	17,9	
<b>Erreur relative (%)</b>						
<b>Moment dipolaire –</b>	2,21	0,87	1,21	0,21	0,35	
<b>Erreur relative (%)</b>						
<b>Nombre d'indéterminations (/15)</b>	5	6	7	10	12	

Tableau 4.8

Si le moment dipolaire est correctement retranscrit dans tous les cas, la qualité de l'ajustement du potentiel électrostatique varie en fonction de la surface. Le nombre d'indéterminations (atomes cachés) croît rapidement avec la distance ; les charges de Mulliken employées comme référence montrent alors clairement leurs limites : l'ajustement à longue distance n'est pas meilleur qu'à courte, malgré le fait que les interactions y soient

d'une nature plus simple. Il faut noter, en effet, que la substitution de charges par leurs valeurs de référence influe non seulement sur les atomes cachés mais également sur ceux qui ne le sont pas, *via* la contrainte imposée sur la somme des charges.

L'absence d'un jeu de charges de référence raisonnable conduit donc à limiter, autant que possible, le nombre d'indéterminations résultant du paramétrage.

L'espace de test est à présent étendu à l'union de toutes les surfaces  $H_\gamma$  :  $G_1 = H_\gamma$  et

$G_2 = \bigcup_{\gamma=1,4}^{10,0} H_\gamma$ . Les résultats deviennent (Tableau 4.9):

$\gamma$	<b>1,4</b>	<b>1,6</b>	<b>1,8</b>	<b>2,0</b>	<b>2,2</b>	<b>2,4</b>
<b>Potentiel électrostatique –</b>	10,4	13,1	13,0	14,3	28,2	26,0
<b>Erreur relative (%)</b>						
$\gamma$	<b>2,6</b>	<b>2,8</b>	<b>3,0</b>	<b>5,0</b>	<b>10,0</b>	
<b>Potentiel électrostatique –</b>	25,7	26,2	26,3	33,4	46,2	
<b>Erreur relative (%)</b>						

Tableau 4.9

L'ajustement sur la surface  $H_{1,4}$ , qui ne donne lieu à aucune indétermination, assure les meilleurs résultats à toutes distances et a été retenu pour cette étude. La qualité de la modélisation se dégrade au fur et à mesure que les points d'ajustement s'éloignent de la molécule et que le nombre d'atomes cachés croît.

Les erreurs relatives sur le potentiel électrostatique par gammes de distances, pour l'ajustement retenu ( $G_1=H_{1,4}$  et  $G_2=H_\gamma$ ), sont les suivantes (Tableau 4.10) :

$\gamma$	<b>1,4</b>	<b>1,6</b>	<b>1,8</b>	<b>2,0</b>	<b>2,2</b>	<b>2,4</b>
<b>Erreur relative (%)</b>	13,1	10,0	7,8	6,4	5,4	4,7
$\gamma$	<b>2,6</b>	<b>2,8</b>	<b>3,0</b>	<b>5,0</b>	<b>10,0</b>	
<b>Erreur relative (%)</b>	4,2	3,9	3,5	1,8	1,3	

Tableau 4.10

#### 4. États Excités Collectifs de Fragments d'ADN : Approche Dynamique

L'ajustement du modèle électrostatique sur la grille  $H_{1,4}$  fournit donc une distribution de charges de bonne qualité, l'erreur sur le potentiel électrostatique aux distances intermoléculaires typiques ( $\gamma = 2,0$ ) avoisinant les 6 %. L'emploi de cette grille est d'ailleurs préconisé dans GAMESS comme '*empirical best value*'<sup>263</sup>.

La même démarche est employée pour la mise en place du modèle électrostatique des états électroniques excités de l'adénine et de la thymine. Ici encore, l'ajustement des charges à partir du potentiel électrostatique sur la grille  $H_{1,4}$  (composée dans ce cas de la thymine de 4151 points) se révèle le meilleur choix. Les erreurs relatives sur le potentiel électrostatique par gammes de distances sont présentées dans le tableau 4.11. Le tableau 4.12 compare quant à lui norme et direction (en convention de DeVoe et Tinoco<sup>266</sup>) des moments dipolaires des états, selon qu'elles sont déterminées *ab initio* ou issues de la distribution de charges atomiques.

$\gamma$	1,4	1,6	1,8	2,0	2,2	2,4
Adénine $S_0$	13,1	10,0	7,8	6,4	5,4	4,7
Adénine $S_1$	16,7	12,7	9,8	8,0	6,7	5,9
Adénine $S_2$	10,9	8,3	6,4	5,3	4,4	3,8
Thymine $S_0$	7,9	5,7	4,4	3,5	2,9	2,6
Thymine $S_1$	6,1	4,2	3,2	2,4	2,0	1,7
$\gamma$	2,6	2,8	3,0	5,0	10,0	
Adénine $S_0$	4,2	3,9	3,5	1,8	1,3	
Adénine $S_1$	5,3	4,9	4,5	2,7	2,3	
Adénine $S_2$	3,4	3,1	2,8	1,4	1,0	
Thymine $S_0$	2,2	2,0	1,9	1,2	1,1	
Thymine $S_1$	1,5	1,3	1,2	0,9	0,8	

Tableau 4.11

	$\ \bar{\mu}\ $ (D)			$\Theta$ (°)		
	CASSCF	Charges Atomiques	Erreur (%)	CASSCF	Charges Atomiques	Erreur (%)
<b>Adénine S<sub>0</sub></b>	2,42	2,42	< 0,1	-89,6	-88,9	0,8
<b>Adénine S<sub>1</sub></b>	2,10	2,09	0,3	-85,7	-84,3	1,6
<b>Adénine S<sub>2</sub></b>	2,52	2,53	0,7	-26	-25,6	1,5
<b>Thymine S<sub>0</sub></b>	4,06	4,09	0,8	-124,1	-126,0	1,5
<b>Thymine S<sub>1</sub></b>	6,96	6,94	0,4	-169,2	-166,4	1,7

Tableau 4.12

L'ajustement du potentiel électrostatique s'avère plus aisé dans la thymine que dans l'adénine. Ceci tient essentiellement au fait que les états de la thymine sont, en général, caractérisés par un moment dipolaire plus fort et un moment quadrupolaire moins fort que les états de l'adénine. La précision attendue de la distribution de charge demeure cependant très satisfaisante pour tous les états traités. Le moment dipolaire équivalent à la distribution de charges atomiques constitue pour sa part une excellente approximation du moment dipolaire permanent CASSCF, tant en norme qu'en direction.

#### 4.2.8. Variation des charges atomiques entre états

La description de la densité électronique des états fondamental et excités de l'adénine et de la thymine au moyen de monopôles ajustés sur le potentiel électrostatique *ab initio* est à la fois simple et satisfaisante. La variation des charges atomiques entre états électroniques, nécessaire pour le calcul des éléments diagonaux de la matrice excitonique, peut ainsi être déduite. Elle est représentée graphiquement en Figure 4.1, pour les trois transitions considérées.

#### 4. États Excités Collectifs de Fragments d'ADN : Approche Dynamique

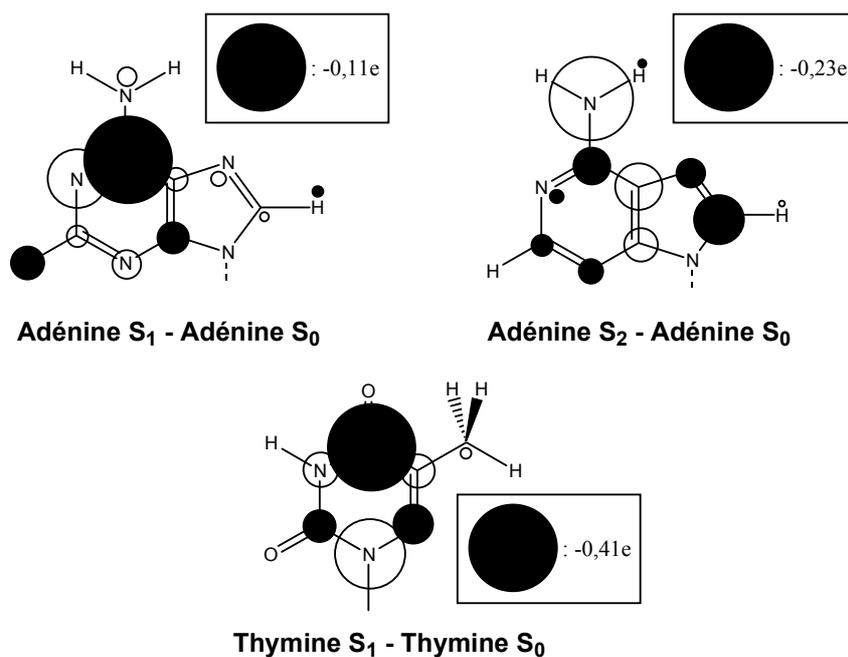


Figure 4.1

Les caractères, localisés ou à transfert de charge, des différents états excités de l'adénine et de la thymine apparaissent clairement sur cette figure. La première transition de l'adénine affecte la densité électronique de manière assez localisée, d'où une faible variation du moment dipolaire permanent, tant en norme qu'en orientation (tableau 4.12). La seconde transition de cette base concerne au contraire l'ensemble de la molécule, mais de manière modérée et bien répartie. Seul l'angle du moment dipolaire permanent en est réellement affecté. Enfin, un fort transfert de charge a lieu lors de l'excitation de la thymine depuis son état fondamental vers son premier état excité, modifiant le moment dipolaire permanent de manière considérable en norme (*ca.* 3 D) et en orientation (environ 45°).

#### 4.2.9. Exploitation

Le champ de force AMBER<sup>255,256</sup>, choisi dans ce travail pour décrire les énergies d'interaction au sein des doubles brins d'ADN, a pour paramètres un ensemble de charges atomiques adaptées aux systèmes à l'étude dans leur état électronique fondamental. Il ne permet pas, faute de charges atomiques adéquates, de traiter les cas où certaines bases se trouveraient dans des états électroniques excités. Ces jeux ont été constitués, dans le cadre de

cette étude, en ajoutant à la charge atomique AMBER pour un atome et une base donnés la différence de charge précédemment obtenue (4.2.8), pour ce même atome sur cette même base. Ce modèle devrait permettre de prendre en considération la réorganisation du système électronique des bases sous l'effet de l'excitation, tout en conservant la généralité des paramètres AMBER, bien adaptés aux acides nucléiques.

### 4.3. Génération de la matrice excitonique

#### 4.3.1. Éléments diagonaux

Comme démontré précédemment, l'élément diagonal de la matrice excitonique correspondant à la fonction d'onde  $\Phi_m^a$ , pour laquelle l'excitation est localisée sur le chromophore  $m$  (qui se trouve alors dans son état électronique  $a$ ) s'écrit :

$$\left\langle \Phi_m^a \left| \hat{H} \right| \Phi_m^a \right\rangle = E_0 + (\varepsilon_m^a - \varepsilon_m^0) + \sum_{n \neq m} (E_{m,n}^{a,0} - E_{m,n}^{0,0})$$

où  $E_0$  est l'énergie du système constitué de tous les chromophores à l'état fondamental, la molécule  $m$  a une énergie  $\varepsilon_m^a$  dans son état électronique  $a$  et  $\varepsilon_m^0$  à l'état fondamental,  $E_{m,n}^{a,0}$  est l'énergie d'interaction des molécules  $m$  et  $n$ , respectivement dans l'état  $a$  et à l'état fondamental, et  $E_{m,n}^{0,0}$  l'énergie d'interaction de ces deux mêmes molécules lorsque toutes deux sont dans leur état fondamental. Cette expression s'écrit également :

$$\left\langle \Phi_m^a \left| \hat{H} \right| \Phi_m^a \right\rangle = E_0 + \underbrace{(\varepsilon_m^a - \varepsilon_m^0)}_I + \underbrace{\left( \sum_{n \neq m} E_{m,n}^{a,0} + \sum_{\substack{n \neq m \\ p < n \\ p \neq m}} E_{n,p}^{0,0} \right)}_{II} - \underbrace{\left( \sum_{n \neq m} E_{m,n}^{0,0} + \sum_{\substack{n \neq m \\ p < n \\ p \neq m}} E_{n,p}^{0,0} \right)}_{III}$$

La référence des énergies pour les éléments de matrice diagonaux est prise égale à  $E_0$  (formellement,  $E_0 = 0$ ). Les éléments diagonaux représentent alors les énergies d'excitation associées aux formes diabatiques à excitation localisée, tandis que les valeurs propres de la matrice excitonique correspondent aux énergies d'excitation des états électroniques adiabatiques du système. Ce choix arbitraire n'entraîne aucune perte de généralité et n'influe pas sur les vecteurs propres décrivant la décomposition des états adiabatiques du système sur les états électroniques des monomères.

#### 4. États Excités Collectifs de Fragments d'ADN : Approche Dynamique

Cette convention une fois adoptée, chaque élément diagonal consiste en trois termes. Le premier (I) est l'énergie d'excitation du monomère  $m$ , depuis son état fondamental vers son état excité  $a$  ; il est pris égal à sa valeur expérimentale (4.2.6). Le deuxième (II) correspond à l'énergie d'interaction de l'agrégat dans lequel la molécule  $m$  est dans son état  $a$  et les autres molécules à l'état fondamental. Il est calculé avec le champ de force AMBER, dont les charges atomiques pour la molécule  $m$  sont remplacées par celles déterminées ci-dessus (4.2.9) pour l'état électronique  $a$  de cette molécule. Le troisième (III) est lui aussi issu d'un calcul AMBER, mais employant cette fois les charges AMBER standard. Les termes (II) et (III) prennent en considération l'influence du solvant et des contre-ions de manière implicite, *via* la méthode de Poisson-Boltzmann non linéaire<sup>170</sup>.

Il convient également de remarquer que la rigidité des bases de l'ADN, supposée jusqu'à présent, n'est pas explicitement imposée dans les calculs de dynamique moléculaire AMBER. Les structures qui en résultent peuvent donc être affectées par de très légères déformations des bases. Conformément à la philosophie du champ de forces AMBER, et compte tenu de la faible amplitude de ces déformations, les charges déterminées pour les états excités de l'adénine et de la thymine ont été supposées 'attachées' aux atomes, et employées quelle que soit la géométrie des bases.

Les termes (II) et (III) introduisent dans le modèle les effets d'environnement (ou désordre diagonal) qui faisaient défaut à la précédente étude. Ils sont dus d'une part aux fluctuations de la structure de la double hélice, d'autre part à l'interaction des bases avec le solvant et les contre-ions implicitement inclus dans le calcul. Le désordre diagonal tend à localiser l'excitation, en levant la dégénérescence des énergies des formes diabatiques associées à un même type d'excitation des monomères (voir article en 3.2).

##### 4.3.2. *Éléments non diagonaux*

Les éléments non diagonaux sont évalués, comme dans l'étude précédente (paragraphe 3), par le biais de l'interaction entre distributions de charges de transition sur les monomères concernés par le transfert d'excitation. Néanmoins, l'introduction *via* la dynamique moléculaire d'une flexibilité dans la double hélice induit un désordre structural qui agit sur les couplages. Ce phénomène, dénommé désordre non diagonal, peut tendre à

## Troisième Partie – Le Transfert d'Excitation

localiser ou à délocaliser l'excitation et rentre en jeu dans la compétition avec les éléments diagonaux pour déterminer le comportement final des états électroniques du système.

### 4.4. Article

L'article ci-après a été accepté pour publication au *Journal of Physical Chemistry*

*B.*

# **Influence of conformational dynamics on the exciton states of DNA oligomers**

**Benjamin Bouvier,<sup>a</sup> Jean-Pierre Dognon,<sup>a</sup> Richard Lavery,<sup>b</sup> Dimitra Markovitsi,<sup>a\*</sup>  
Philippe Millié,<sup>a</sup> Delphine Onidas<sup>a</sup> and Krystyna Zakrzewska<sup>b</sup>**

*Laboratoire Francis Perrin CEA/DSM/DRECAM/SPAM - CNRS URA 2453, CEA Saclay,  
91191 Gif-sur-Yvette, France*

*Laboratoire de Biochimie Théorique, CNRS UPR 9080, Institut de Biologie Physico-  
Chimique, 13, rue Pierre et Marie Curie, 75005 Paris, France*

---

\* Corresponding author: Dimitra Markovitsi ([dimitra.markovitsi@cea.fr](mailto:dimitra.markovitsi@cea.fr))

<sup>a</sup> Laboratoire Francis Perrin

<sup>b</sup> Laboratoire de Biochimie Théorique

## **Abstract**

The present communication examines how the dynamics of the double helix affects the Frenkel excitons corresponding to the low-energy absorption band of DNA. Two types of oligomers,  $(dA)_n \cdot (dT)_n$  and  $(dAdT)_{n/2} \cdot (dAdT)_{n/2}$ , are studied theoretically, in the frame of the exciton theory in combination with quantum chemical calculations. The properties of the exciton states (energy, oscillator strength, degree of delocalization, “anisotropy”...) found for canonical B-DNA geometries are compared to those obtained for conformations extracted from molecular dynamics simulations. It is shown that although structural fluctuations reduce both the mixing between different monomer transitions and the spatial extent of the eigenstates, excitations still remain delocalized over several bases. The presence of alternating base sequences makes the eigenstates of the double-stranded oligomers more sensitive to disorder. All these effects arise from a variation of the coupling terms, the diagonal energy being only slightly altered by the structural fluctuations. The experimental absorption spectra presented here corroborate the theoretical results according to which the absorption of  $(dA)_n \cdot (dT)_n$  is centered at higher energies than that of  $(dAdT)_{n/2} \cdot (dAdT)_{n/2}$ . Finally, it is shown that, in contrast to what is commonly admitted, the formation of collective excited states in double-stranded oligomers is not expected to induce large spectral shifts with respect to a homogeneous mixture of monomers.

## 1. Introduction

It is well known that reactions occurring after the absorption of UV radiation by the DNA bases may induce lethal or carcinogenic mutations. Although the major products of these reactions have been identified,<sup>1</sup> the physicochemical processes preceding DNA damage remain unclear. The very first step in such a series of events is the formation of Franck-Condon excited states. Their nature should determine, in large part, the fate of the excitation energy and, in particular, the possibility for it to be transferred amongst the bases, thus increasing the efficiency of the photoreactions. In this respect, DNA can be viewed as an organized molecular system composed of closely lying chromophores. Therefore, the question arises whether its electronic excited states are localized on single bases or more or less strongly delocalized.

The first communication on DNA excitons appeared in 1960 by Tinoco who described the excited states of the double helix theoretically by introducing dipolar coupling between the electronic transitions of the bases.<sup>2</sup> On the other hand, experimentalists observed that the spectra of “double-stranded DNA closely resembles the sum of the absorption spectra of the constituent purine and pyrimidine bases” and concluded that “the electronic interaction between the bases is weak enough so that it is proper to speak of the absorption of an ultraviolet photon by a single base”.<sup>3</sup> This statement, which guided subsequent studies dealing with energy transport,<sup>4-7</sup> implicitly postulates that the formation of exciton states in double helices should induce drastic changes in the profile of the absorption spectra.

The footprint of exciton states in experimental absorption spectra is not always clear because antagonistic factors may blur the underlying effect. For example, delocalization of the excitation in columnar stacks of chromophores induces a blue shift<sup>8,9</sup> but other geometrical arrangements generate different changes;<sup>10</sup> charge transfer or charge resonance interactions induce a red shift in the spectra,<sup>11,12</sup> and so does, in general, monomer “solvation” in organized molecular system. Hence, other observables, like those obtained by time-resolved spectroscopic measurements, are precious for comparing the behavior of double helices to that of the constituent monomers. Nevertheless, as the excited lifetimes of nucleic acids are extremely short, such measurements were not performed until very recently.<sup>13-17</sup> The first time-resolved investigation of DNA oligomers carried out with femtosecond resolution revealed that molecular organization of nucleotides to form single strands, and further to form double helices, slows down the fluorescence decays.<sup>18</sup> This observation suggests that important changes may occur in the excited states of the DNA oligomers as compared to the

corresponding monomers and raises questions concerning the collective excited states in DNA.

In this context, we have recently undertaken a systematic theoretical study of the Frenkel excitons of double-stranded oligomers in close relation with experiment. In a first investigation, we focused on two aspects: the monomer electronic transitions which have to be taken into account in the construction of the Hamiltonian matrix and the dipolar coupling.<sup>19</sup> We determined the energies and transition moments of the lowest dipolar transitions of the individual chromophores forming the double helices by decomposing the absorption spectra of nucleosides in aqueous solution on the basis of fluorescence anisotropy and circular dichroism measurements. In parallel, we calculated these properties with the CS-INDO-CIPSI method and found a satisfactory agreement with the experimental values. Subsequently, we determined the dipolar coupling using the atomic transition charge distribution model. Finally, we calculated the properties of the exciton states of two particular oligonucleotides,  $(dA)_{20} \cdot (dT)_{20}$  and  $(dAdT)_{10} \cdot (dAdT)_{10}$ , considering three closely-lying molecular electronic transitions, two for adenosine and one for thymidine.

The most important conclusion of our first approach was that dipolar coupling may lead to a spatial delocalization of the excitation within double helices having an idealized B-DNA geometry. But, as was pointed out in that study, the plasticity of the double helix could induce a localization of the electronic excitation. As matter of fact, a given DNA oligomer can adopt a range of conformations as a function of time which may lead to variations of the monomer excitation energy within the organized system (diagonal disorder) as well as the variation of the electronic coupling (off-diagonal disorder). The propensity of structural fluctuations of the double helix to localize the excitation depends on the relative amplitude of induced diagonal and off-diagonal disorder compared to the strength of the coupling.

Here, we follow the methodology developed in reference<sup>19</sup>, combining exciton theory and quantum chemical methods, to which we add ground state molecular dynamics calculations. In this way, we are able to model structural fluctuations by taking into account solvent, counter-ion and temperature effects. We deal again with the two types of double-stranded oligomers,  $(dA)_n \cdot (dT)_n$  and  $(dAdT)_{n/2} \cdot (dAdT)_{n/2}$ , composed of the same kind of base pairs but having, respectively, homopolymeric and alternating base sequences. In Section 2 we describe the procedure followed in the calculations. In Section 3 we present and discuss our results. First, we quantify the amplitude of diagonal and off-diagonal disorder induced by the dynamics of the double helix in the ground state and we correlate it with the variation of

some structural parameters (§3.1). Then we compare the properties of the Franck-Condon eigenstates (energy, oscillator strength, degree of delocalization, contribution of monomer electronic transitions) obtained with and without structural fluctuations and we tackle the problem of the “anisotropy” which should result in intraband scattering (§3.2). In §3.3, we present the absorption characteristics of  $(dA)_n.(dT)_n$  and  $(dAdT)_{n/2}.(dAdT)_{n/2}$  oligomers by examining in parallel calculated properties and the profile of experimental steady-state absorption spectra. Our conclusions and comments are presented in Section 4.

## 2. Calculation Procedure

A molecular dynamics simulation of the ground-state double helices, including solvent effects and counter-ions, was used to sample conformational space. The excited states of instantaneous conformations chosen from this simulation were then characterized using an effective Hamiltonian formalism.

### 2.1 Ground-state Molecular Dynamics Simulation.

Model building and simulations were performed using the AMBER 6 program<sup>20</sup> and the Parm98 parameter set.<sup>21</sup> Both oligomers were constructed using a standard B-DNA conformation and were subsequently neutralized with 22  $Na^+$  ions (placed using electrostatic potentials) and solvated with more than 6000 TIP3P water molecules in a truncated octahedral box. Molecular dynamics simulations were performed at constant temperature (300 K) and pressure (1 bar) using the Berendsen algorithm.<sup>22</sup> An integration time step of 2 fs was used and all bond lengths involving hydrogens were constrained using SHAKE.<sup>23</sup> Long-range electrostatic interactions were treated using the particle mesh Ewald (PME) approach<sup>24</sup> with a 9 Å direct space cut-off. The non-bonded pair-list was updated heuristically and the center of mass motion was removed every 10 ps during the simulation. Initially, the water molecules and ions were relaxed by energy minimization and allowed to equilibrate at 300 K around the fixed DNA for 100 ps at constant volume; the entire system was then heated from 100 to 300 K during 10 ps and equilibrated during 40 ps with harmonic restraints of 5.0 kcal/mol/Å<sup>2</sup> on the solute atoms at constant volume. Subsequently, the simulation was continued at constant pressure; the restraints were gradually removed over a period of 250 ps and an unrestrained simulation followed for over 4 ns. The coordinates were saved every 1 ps. The last nanosecond was used for the further study. 100 snapshots spaced by 10 ps were generated. In order to minimize bond length and valence angle distortions the snapshots were minimized in

AMBER for 1000 cycles before being used for Poisson-Boltzmann calculations of the electrostatic energy.

## 2.2 Calculation of exciton states

**Effective Hamiltonian Formalism.** The excited states of double-stranded DNA fragments are calculated in the framework of the exciton theory,<sup>25-27</sup> in which the exact Hamiltonian of an  $n$ -molecular supersystem may be written:

$$H = H_0 + V \quad (1)$$

$H_0$ , the unperturbed Hamiltonian of the supersystem, is evaluated as the sum of the individual Hamiltonians of the isolated molecules. As a consequence, the zeroth-order eigenfunctions of the system consist of products of the eigenfunctions of these molecules:

$$H_0 = \sum_{\text{molecules}, k}^n H_k \quad \text{and} \quad |\Phi_m^i\rangle = \Psi_m^i \prod_{\text{molecules}, k \neq m}^n \Psi_k^0 \quad (2)$$

where  $\Psi_m^i$  denotes the  $i^{\text{th}}$  excited state of chromophore  $m$ . The excited states of the supersystem are thus decomposed on diabatic states for which the excitation is localized on a given monomer, the other molecules being in their respective ground states.

$V$  is the perturbation operator, which was formalized by Longuet-Higgins<sup>28</sup> using a local charge density operator  $\rho^{(k)}(\vec{r}^{(k)})$  associated with each individual chromophore  $k$ :

$$V = \iint \frac{\rho^{(k)}(\vec{r}^{(k)})\rho^{(l)}(\vec{r}^{(l)})}{|\vec{r}^{(k)} - \vec{r}^{(l)}|} d\vec{r}^{(k)} d\vec{r}^{(l)} \quad (3)$$

Diagonalization of the exciton matrix  $H$  in the basis formed by the eigenfunctions of  $H_0$ , at the first order of perturbation, yields  $N$  eigenstates

$$|\Gamma_k\rangle = \sum_{\text{molecules } m} \sum_{\text{states } i} C_{k,m}^i |\Psi_1^0 \Psi_2^0 \dots \Psi_m^i \dots \Psi_n^0\rangle. \quad \text{The squares of the linear combination coefficients}$$

$C_{k,m}^i$  represent the contribution of the diabatic basis wavefunction  $|\Phi_m^i\rangle$  to the adiabatic system eigenstate  $|\Gamma_k\rangle$ .

**Diagonal Exciton Matrix Elements.** The diagonal element of  $H$  associated with the diabatic basis wavefunction  $|\Phi_m^i\rangle$  may be expressed as:

$$\begin{aligned} \langle \Phi_m^i | H | \Phi_m^i \rangle &= \langle \Phi_m^i | H_m | \Phi_m^i \rangle + \sum_{n \neq m} \langle \Phi_m^i | H_n | \Phi_m^i \rangle + \langle \Phi_m^i | V | \Phi_m^i \rangle \\ &= \varepsilon_m^i + \sum_{n \neq m} \varepsilon_n^0 + \sum_{n \neq m} E_{m,n}^{i,0} + \sum_{\substack{n \neq m \\ p < n \\ p \neq m}} E_{n,p}^{0,0} \end{aligned} \quad (4)$$

where  $E_{m,n}^{i,j}$  denotes the interaction energy of monomer  $m$  in its  $i^{\text{th}}$  electronic state with monomer  $n$  in its  $j^{\text{th}}$  electronic state, while  $\varepsilon_m^i$  is the electronic energy of chromophore  $m$  in its  $i^{\text{th}}$  electronic state, as defined by:

$$H_m \Psi_m^i = \varepsilon_m^i \Psi_m^i \quad (5)$$

Let  $E_0$  be the electronic energy of the system composed of all chromophores in their electronic ground state:

$$E_0 = \sum_m \varepsilon_m^0 + \sum_m \sum_{\substack{n < m \\ n \neq m}} E_{m,n}^{0,0} \quad (6)$$

The diagonal matrix element  $\langle \Phi_m^i | H | \Phi_m^i \rangle$  thus becomes:

$$\langle \Phi_m^i | H | \Phi_m^i \rangle = E_0 + \underbrace{(\varepsilon_m^i - \varepsilon_m^0)}_{\text{I}} + \underbrace{\left( \sum_{n \neq m} E_{m,n}^{i,0} + \sum_{\substack{n \neq m \\ p < n \\ p \neq m}} E_{n,p}^{0,0} \right)}_{\text{II}} - \underbrace{\left( \sum_{n \neq m} E_{m,n}^{0,0} + \sum_{\substack{n \neq m \\ p < n \\ p \neq m}} E_{n,p}^{0,0} \right)}_{\text{III}} \quad (7)$$

Choosing  $E_0$  as the reference energy (formally,  $E_0=0$ ), the diagonal matrix element consists of three terms. Term (I) represents the excitation energy of monomer  $m$  from its ground to its  $i^{\text{th}}$  electronic state. Term (II) corresponds to the interaction energy of the system in which monomer  $m$  is in its  $i^{\text{th}}$  state and all others in their respective ground states. It is computed as the electrostatic energy of the system in water by solving the non-linear Poisson-Boltzmann equation with AMBER atomic charges. In this calculation, atomic charge distributions associated with the excited states of the monomers (whose determination is explained in the next paragraph) were used. Finally, term (III) is the energy of the ground-state system, and is calculated by the Poisson-Boltzmann method.

**Off-Diagonal Exciton Matrix Elements.** Since the charge density operators involved in the perturbation operator  $V$  depend exclusively on the coordinates of the individual subsystems, separation of variables inside the integrand yields, as an expression of the general matrix element of  $V$ :

$$\langle \Psi_n^i \Psi_m^j | V | \Psi_n^k \Psi_m^l \rangle = \frac{\langle \Psi_n^i | \rho^{(n)} \vec{r} | \Psi_n^k \rangle \langle \Psi_m^j | \rho^{(m)} \vec{r} | \Psi_m^l \rangle}{|\vec{r}^{(n)} - \vec{r}^{(m)}|} d\vec{r}^{(n)} d\vec{r}^{(m)} \quad (8)$$

This expression may be viewed as an interaction between appropriate transition charges, obtained from a CS-INDO-CIPSI<sup>29</sup> calculation as explained in reference<sup>19</sup>.

As our model incorporates data from both CS-INDO and CASSCF calculations, the correspondence of monomer states at the two levels must be determined. Electronic transition energies and oscillator strengths computed at the CASSCF level may be found in Table 1, with the CS-INDO and experimental values transcribed from our previous work. Although transition energies show important variations between computational methods and with experiments, the agreement between oscillator strength values is much better and leaves little doubt as to the correct attribution of states.

	CASSCF		CS-INDO		Experiment <sup>a</sup>	
	$E_{\text{exc}}(\text{cm}^{-1})$	$f$	$E_{\text{exc}}(\text{cm}^{-1})$	$f$	$E_{\text{exc}}(\text{cm}^{-1})$	$f$
<b>Adenine <math>S_0 \rightarrow S_1</math></b>	41500	0.001	37200	0.057	36700	0.05
<b>Adenine <math>S_0 \rightarrow S_2</math></b>	48700	0.207	39900	0.216	38800	0.24
<b>Thymine <math>S_0 \rightarrow S_1</math></b>	42300	0.145	36400	0.220	37500	0.24

<sup>a</sup> Values determined for the nucleosides<sup>19</sup>

**TABLE 1: Excitation energies ( $E_{\text{exc}}$ ) and oscillator strengths ( $f$ ) associated with the first two transitions of adenine and the first transition of thymine**

**Degree of delocalization.** The localization behavior of the eigenstates is usually expressed by the inverse participation ratio  $L_k$ .<sup>30,31</sup> The number of coherently coupled chromophores in a given eigenstate  $k$  is given by the participation ratio  $N_k = 1/L_k$ . When the eigenstates are built on more than one molecular states,  $L_k$  is written as:<sup>8</sup>

$$L_k = \sum_{\text{molecules } m} \left[ \sum_{\text{states } i} (C_{k,m}^i)^2 \right]^2$$

**General remarks on the calculation procedure:**

1. As our aim was to obtain the most accurate results possible, in each step of the calculation procedure we used different ways to describe the monomer properties. Thus, in the calculation of the diagonal disorder induced by the structural fluctuations we used the properties of the bases (adenine and thymine) because the influence of the backbone was considered explicitly in the molecular dynamics simulations. In contrast, for the coupling terms, we used the properties of the nucleosides which are the effective chromophores; the transition energies and the oscillator strengths

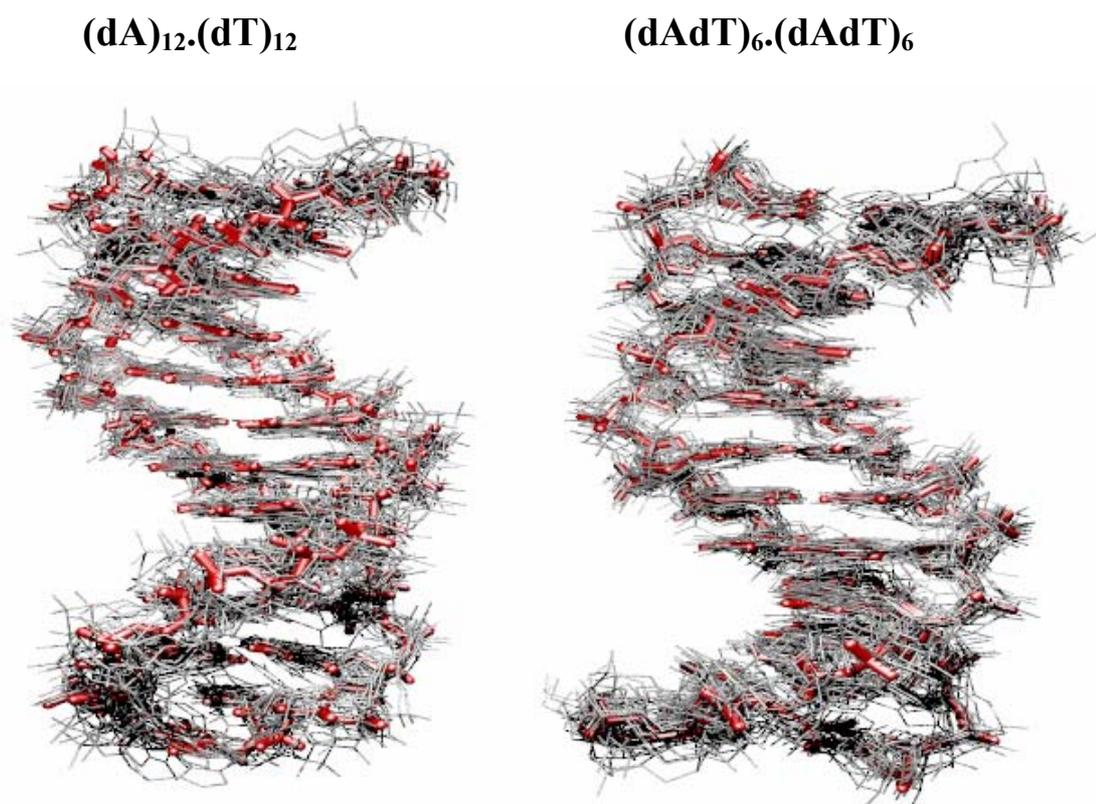
considered are those derived from the decomposition of the experimental absorption spectra recorded in aqueous solutions of adenosine and thymidine (Table 1), as described in our previous work.<sup>19</sup>

2. Molecular dynamics simulations were performed for oligomers with twelve base pairs, but the exciton matrix was constructed on only ten base pairs, neglecting the terminal pairs because of spontaneous base pair dissociation (“fraying”). Thus, all results associated with the excited states of disordered oligomers refer to (dA)<sub>10</sub>.(dT)<sub>10</sub> and (dAdT)<sub>5</sub>.(dAdT)<sub>5</sub>. Furthermore, experimental spectra (§3.3) refer to oligomers with twenty base pairs, because shorter ones, such as those used in the molecular dynamics simulations are not stable at room temperature.

### **3. Results and discussion**

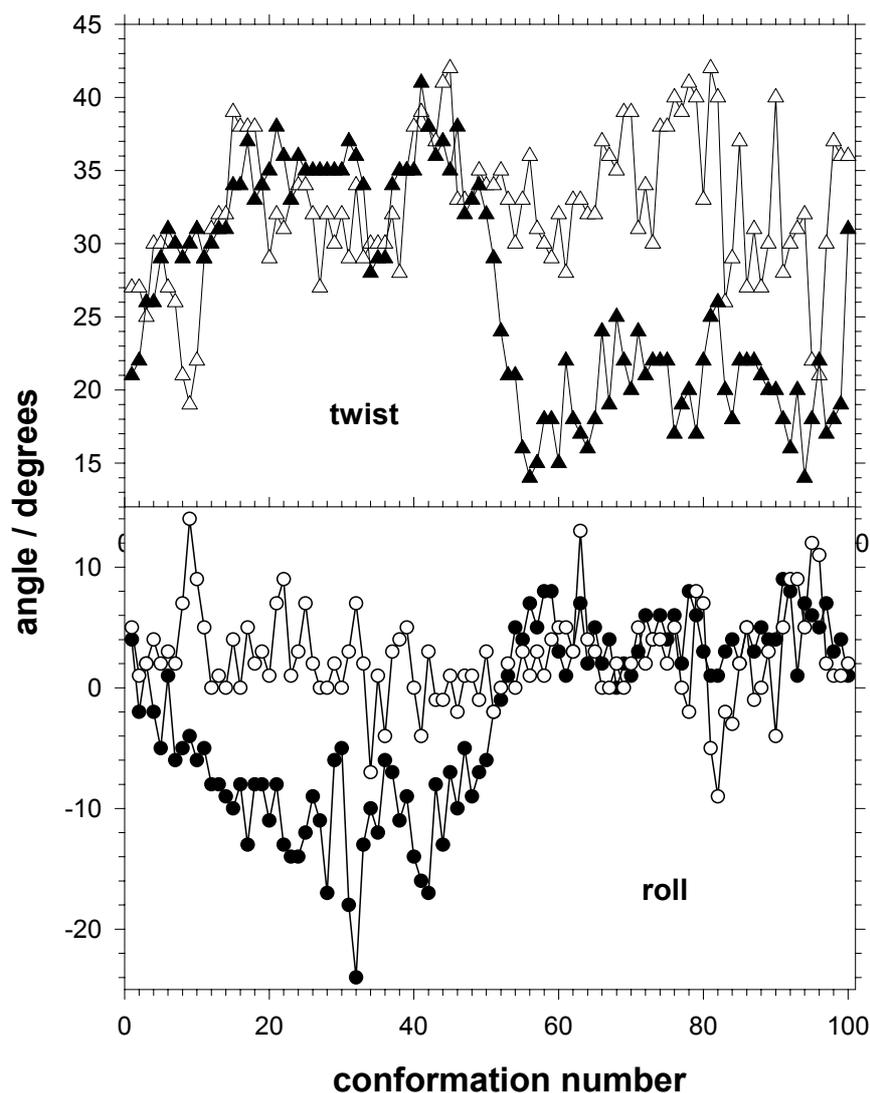
#### **3.1 Structural fluctuations and induced diagonal and off-diagonal disorder**

An analysis of the 100 conformations extracted from the last nanosecond of the molecular dynamics trajectories shows that, at room temperature, both (dA)<sub>12</sub>.(dT)<sub>12</sub> and (dAdT)<sub>6</sub>.(dAdT)<sub>6</sub> undergo significant thermal fluctuations, although they can both be classified as belonging to the B-DNA conformational family. Figure 1 illustrates the overall fluctuations by superimposing 20 representative conformations on the average structure obtained from the last nanosecond of simulation.



**Figure 1:** Average structure of the oligomers studied during the last nanosecond of simulation (red) superimposed on 20 representative conformations (grey) chosen from the same part of the trajectory: (a) (dA)<sub>12</sub>·(dT)<sub>12</sub>, (b) (dAdT)<sub>6</sub>·(dAdT)<sub>6</sub>.

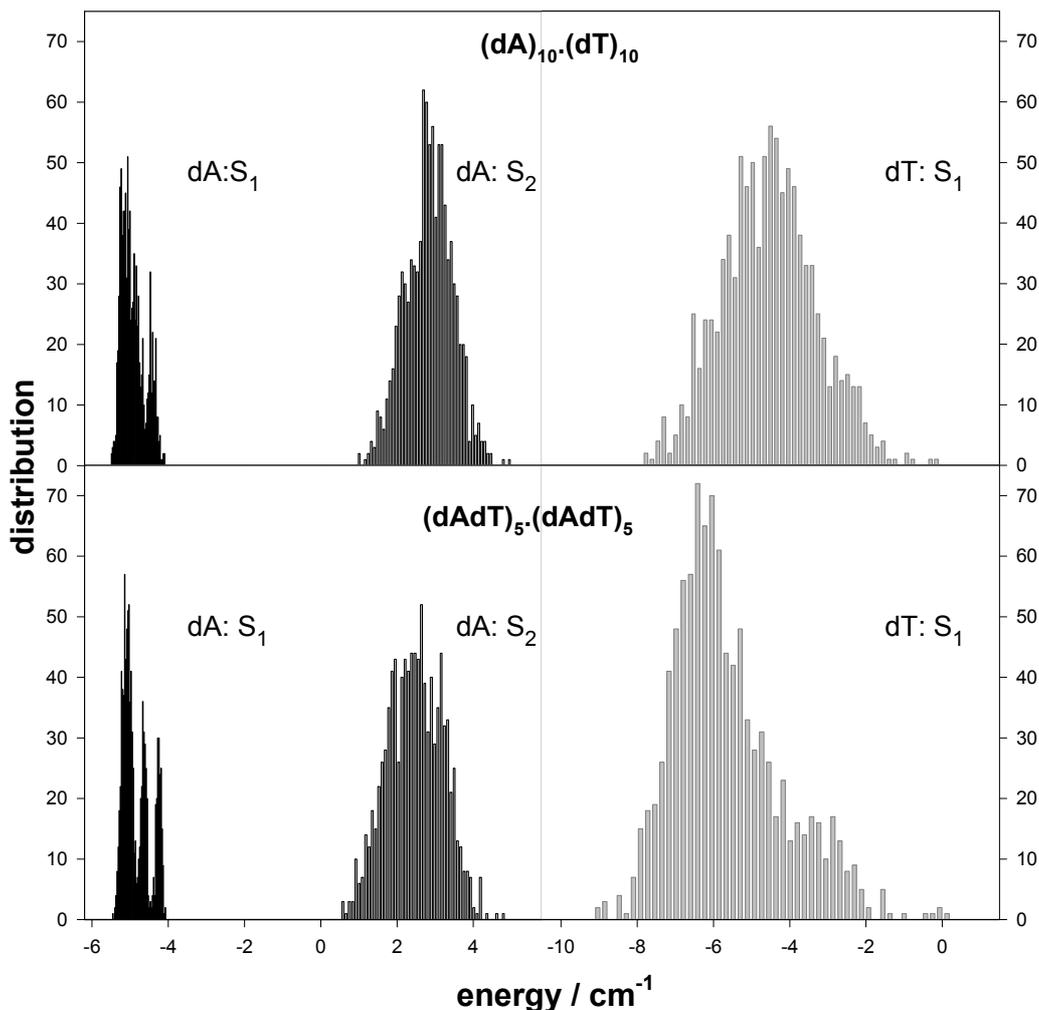
Figure 2 shows the detailed fluctuations of two helical parameters, twist and roll for the A5-A6 step of (dA)<sub>12</sub>·(dT)<sub>12</sub> and of (dAdT)<sub>6</sub>·(dAdT)<sub>6</sub>. Twist is the rotation around the helical axis between successive base pairs and roll is the angle formed between the base pair planes by rotation around their principal axes. Positive roll angles refer to opening up the base pairs on the major groove side of the duplex. The twist and the roll obtained for the homopolymeric sequence exhibit fluctuations on a 10 ps timescale of roughly  $\pm 10^\circ$  around mean values of  $32.5^\circ$  and  $2.4^\circ$ , respectively. Similar rapid fluctuations are observed for the alternating sequence. For the latter oligomer, larger and slower fluctuations also occur in twist and roll for the step analyzed. These large fluctuations in local base stacking are often linked to slow conformational changes occurring in the phosphodiester backbones of DNA.



**Figure 2:** Fluctuations of (a) twist and (b) roll for the A5-A6 step of  $(dA)_{12} \cdot (dT)_{12}$  (white) and of  $(dAdT)_6 \cdot (dAdT)_6$  (black). Twist is the rotation around the helical axis between successive base pairs and roll is the angle formed between the base pair planes by rotation around their principal axes; positive roll angles refer to opening up the base pairs on the major groove side of the duplex.

The fluctuations of the diagonal energy induced by the structural fluctuations are very weak; they do not exceed  $10 \text{ cm}^{-1}$ , that is, three orders of magnitude less than the absolute values of the excitation energy (Table 1). The distribution of the diagonal energy determined for the ten adenosine and ten thymidine chromophores in 100 conformations of  $(dA)_{10} \cdot (dT)_{10}$

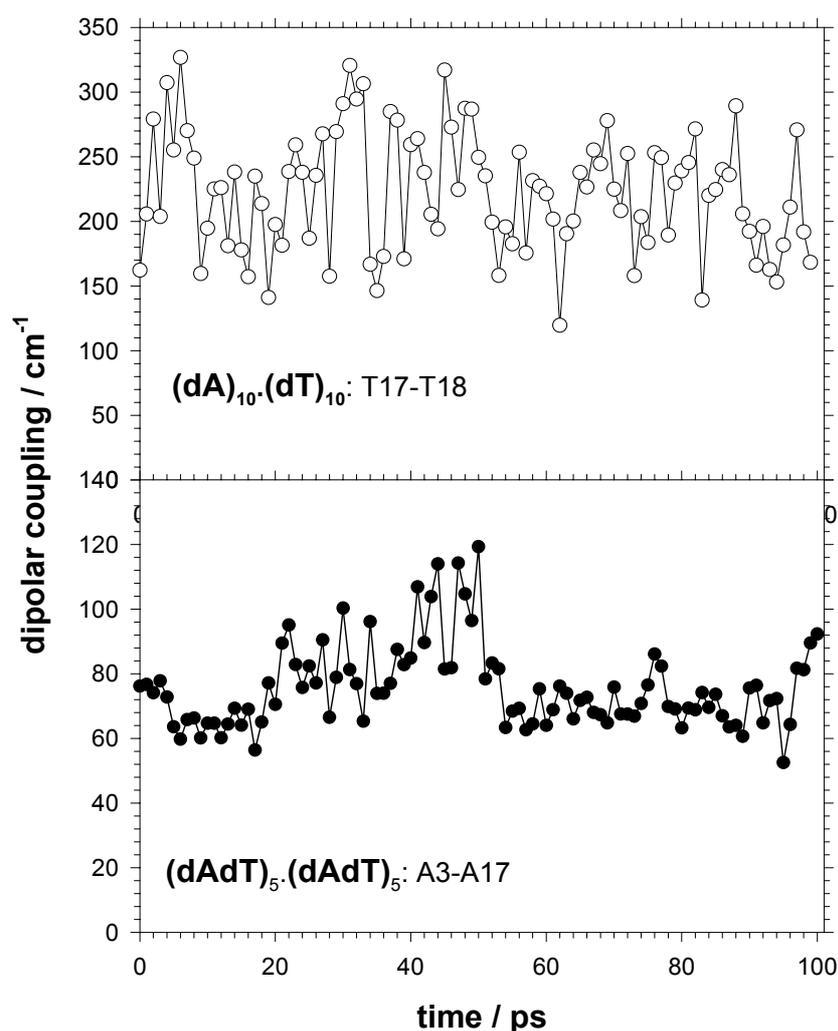
and  $(dAdT)_5 \cdot (dAdT)_5$  is presented in Figure 3. Note that the broadest distribution is observed for the  $S_0 \rightarrow S_1$  transition of thymidine which is explained by an important change in the atomic charge distributions associated with this transition. We also remark that the  $S_1$  states of both adenosine and thymidine are slightly destabilized within the double strands, whereas the  $S_2$  state of adenosine is weakly stabilized.



**Figure 3:** Distribution of the diagonal energy determined for the ten adenosine and ten thymidine chromophores in 100 conformations of  $(dA)_{10} \cdot (dT)_{10}$  and  $(dAdT)_5 \cdot (dAdT)_5$  extracted from the molecular dynamics simulations.

The fluctuations observed for the off-diagonal terms of the Hamiltonian matrix are much more important than those found for the diagonal terms. An example is given in Figure 4, which shows the dipolar coupling associated with one pair of chromophores for each type

of oligomer: the coupling between the  $S_0 \rightarrow S_1$  transitions of the 3<sup>rd</sup> and 4<sup>th</sup> thymidine chromophores in  $(dA)_{10} \cdot (dT)_{10}$  and that between  $S_0 \rightarrow S_2$  transitions of the 3<sup>rd</sup> adenosine located in one strand of  $(dAdT)_5 \cdot (dAdT)_5$  and the 17<sup>th</sup> adenosine located in the other strand. In the case of  $(dA)_{10} \cdot (dT)_{10}$ , the amplitude of the variations in the coupling amounts to 35%. A more complex pattern is found for the coupling fluctuations in  $(dAdT)_5 \cdot (dAdT)_5$ : relatively weak variations (less than 10%) are superimposed on larger a one, reaching 45%. This pattern reflects the behavior of twist and roll observed for the alternating oligomer (Figure 2).



**Figure 4:** Dipolar coupling determined for 100 conformations extracted from the molecular dynamics simulations. Upper plot: coupling between the  $S_0 \rightarrow S_1$  transitions of the 17<sup>th</sup> and 18<sup>th</sup> thymidine chromophores in  $(dA)_{10} \cdot (dT)_{10}$ . Lower plot: coupling between  $S_0 \rightarrow S_2$  transitions of the 3<sup>rd</sup> adenosine located in one strand of  $(dAdT)_5 \cdot (dAdT)_5$  and the 17<sup>th</sup> adenosine located in the other strand.

### 3.2 Eigenstate properties

The eigenstate properties presented here concern exclusively double-stranded oligomers composed of ten base pairs,  $(dA)_{10}.(dT)_{10}$  and  $(dAdT)_5.(dAdT)_5$ . For each type of oligomer, we compare the oscillator strength, the transition energy, the contribution of the monomer transitions and the participation ratio associated with all thirty eigenstates calculated for perfectly ordered conformations<sup>‡</sup> to those found for dynamically disordered conformations. In the latter case, we examine both properties averaged over an ensemble of 100 conformations and the behavior of individual conformations.

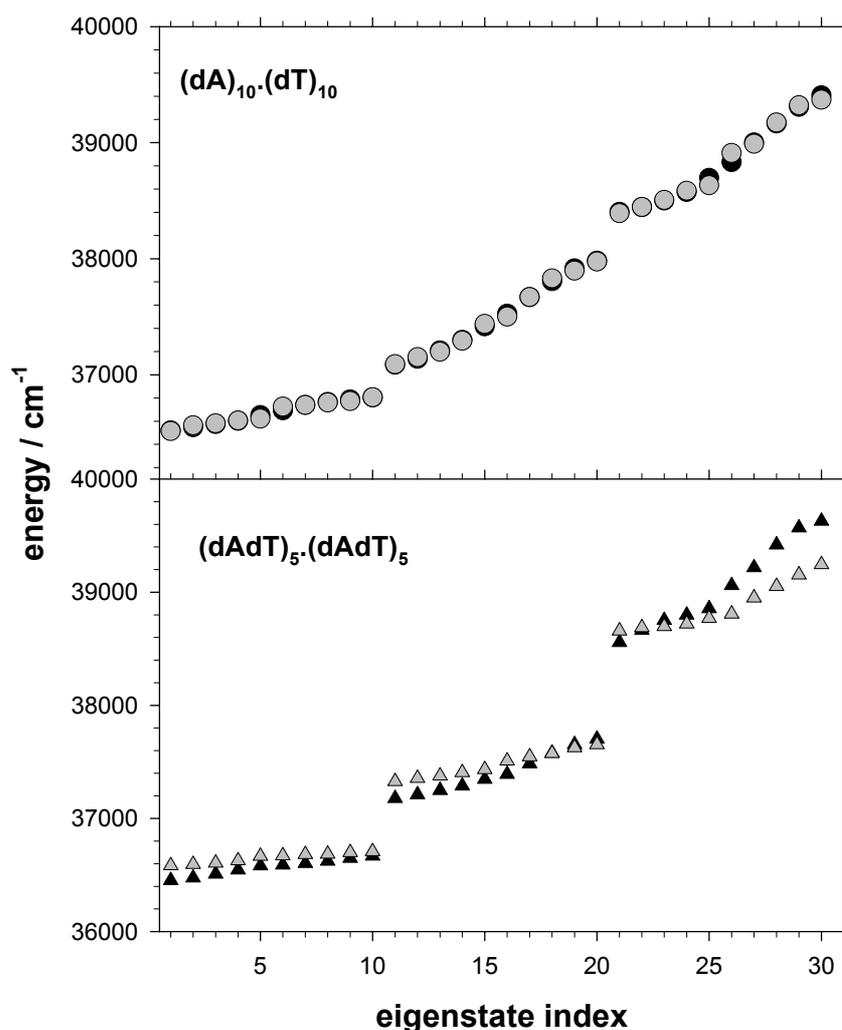
The distribution of the oscillator strength ( $f$ ) over the thirty eigenstates observed for  $(dA)_{10}.(dT)_{10}$  is not much affected by structural fluctuations: for both the ordered conformation and the disordered conformations, 90% of the oscillator strength remains concentrated on the same eigenstates, namely  $\langle 18 \rangle$ ,  $\langle 19 \rangle$ ,  $\langle 20 \rangle$ ,  $\langle 28 \rangle$ ,  $\langle 29 \rangle$  and  $\langle 30 \rangle$ . The picture is different for  $(dAdT)_5.(dAdT)_5$ , for which disorder tends to disperse the oscillator strength over a larger number of eigenstates; in the absence of disorder, 90% of  $f$  is borne by only seven eigenstates, whereas it is spread over twelve of them under the effect of fluctuations. Despite the dispersion of the oscillator strength over several eigenstates, we found that, for perfectly ordered conformations of both  $(dA)_{10}.(dT)_{10}$  and  $(dAdT)_5.(dAdT)_5$ , the maximum value of  $f$  is borne by eigenstate  $\langle 29 \rangle$ . In the case of  $(dA)_{10}.(dT)_{10}$ , the latter eigenstate continues to play a predominant role even in the presence of structural fluctuations. In contrast, the maximum oscillator strength found for disordered conformations of  $(dAdT)_5.(dAdT)_5$  is more frequently associated with eigenstate  $\langle 22 \rangle$ .

The variations in the energy of a given eigenstate of  $(dA)_{10}.(dT)_{10}$  as a function of the structural changes do not exceed  $\pm 100 \text{ cm}^{-1}$  and, in general, they are smaller for the lowest states. The average energy obtained for each of the thirty eigenstates is represented in Figure 5. For all types of double-stranded decamers, ordered or disordered, homopolymeric or alternating, we note the existence of three ensembles of ten eigenstates per oligomer, with abrupt changes in energy at the frontier between any two ensembles. This partition is related to the composition of the eigenstates, that is, the contribution to each eigenstate of the three monomer transitions considered in the exciton matrix, *i.e.*  $S_0 \rightarrow S_1$  and  $S_0 \rightarrow S_2$  of adenine and  $S_0 \rightarrow S_1$  of thymine. The monomer contribution, expressed as the sum of the squares of

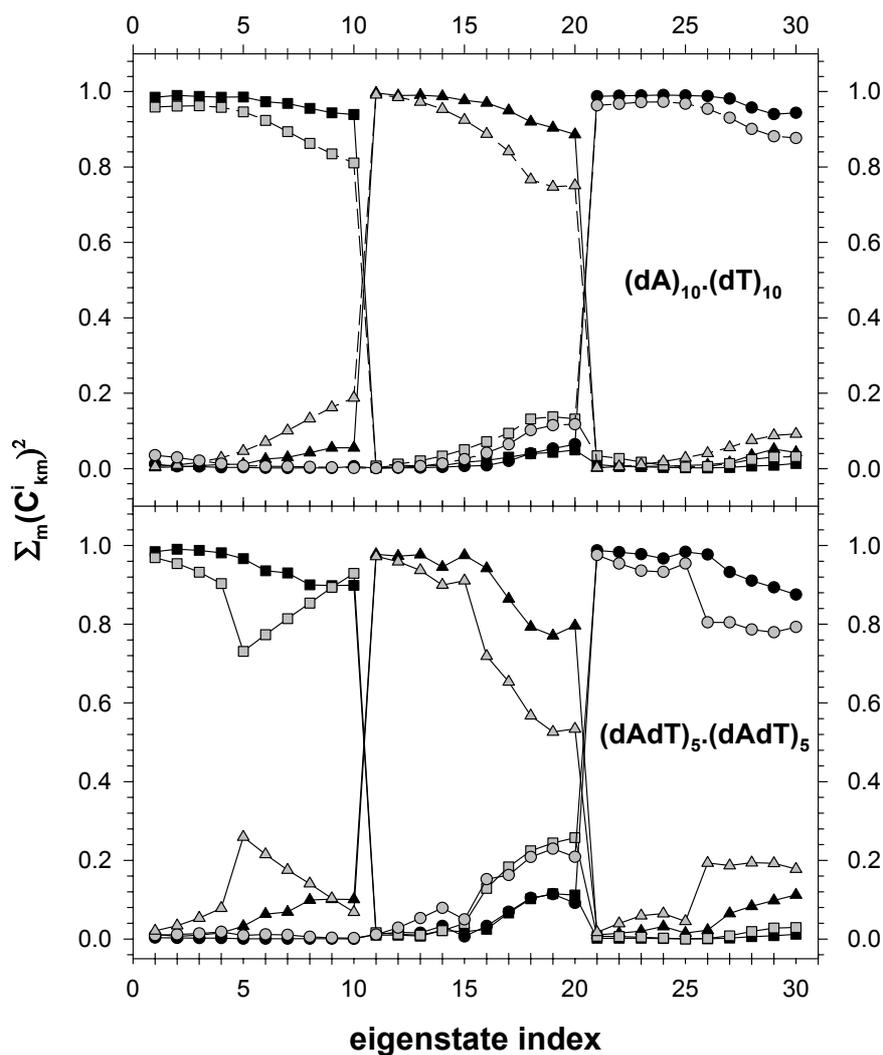
---

<sup>‡</sup> The structural parameters of ordered oligomers are those used in our previous work.<sup>19</sup>

the linear combination coefficients, is illustrated in Figure 6. We remark that, exactly as in Figure 5, the thirty eigenstates can be divided into three sets of ten, each one associated with a particular monomer transition. The mixing of the monomer transitions is relatively more efficient in the eigenstates of the alternating oligomers; this is true, in particular, for the five upper eigenstates of each set. For both types of oligomer, structural fluctuations tend to reduce the mixing between different monomer transitions; thus the eigenstates of disordered conformations are built to a larger extent (at least at 90% for  $(dA)_{10} \cdot (dT)_{10}$  and 75% for  $(dAdT)_5 \cdot (dAdT)_5$ , on a specific monomer electronic state. Thus the “global” exciton band of each oligomer can be seen as composed by three “partial” exciton bands. In the case of  $(dA)_{10} \cdot (dT)_{10}$ , the widths of both the “global” and the three “partial” exciton bands are not altered by structural fluctuations, whereas they decrease for  $(dAdT)_5 \cdot (dAdT)_5$  (Figure 5).

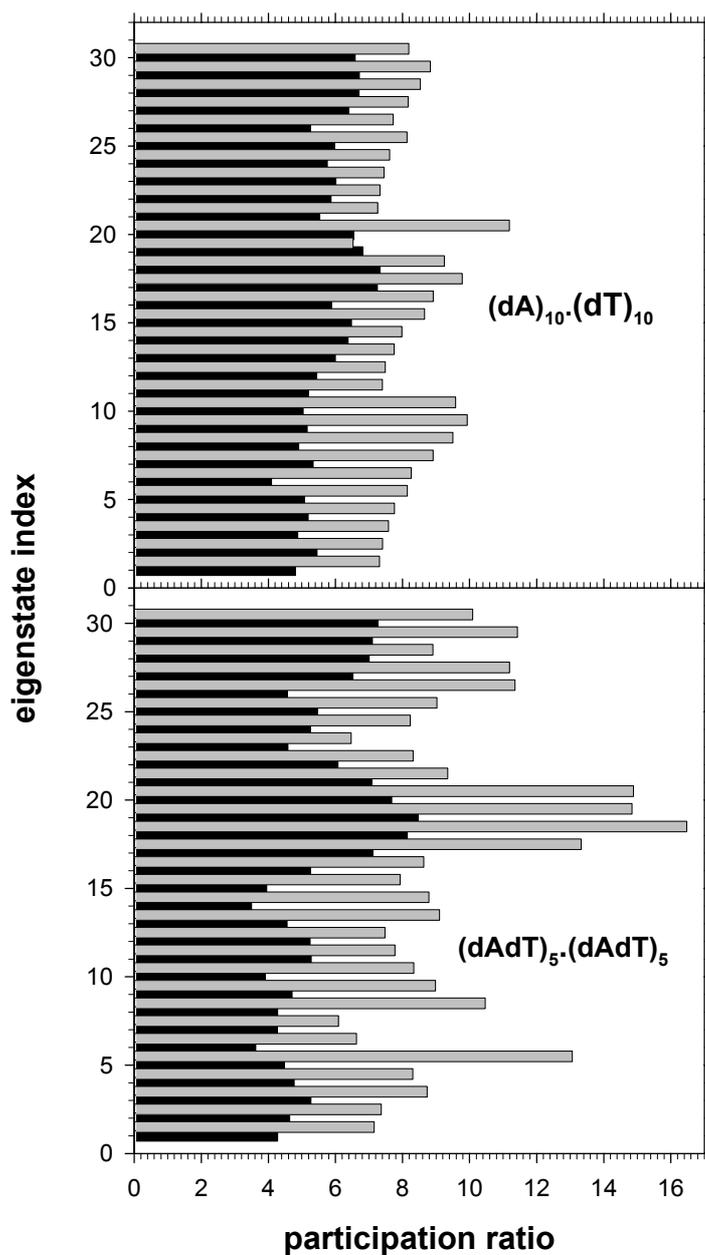


**Figure 5:** Eigenstate energy determined for  $(dA)_{10} \cdot (dT)_{10}$  (circles) and  $(dAdT)_5 \cdot (dAdT)_5$  (triangles). In black: average values for 100 conformations extracted from the molecular dynamics simulations. In grey: ordered conformations.

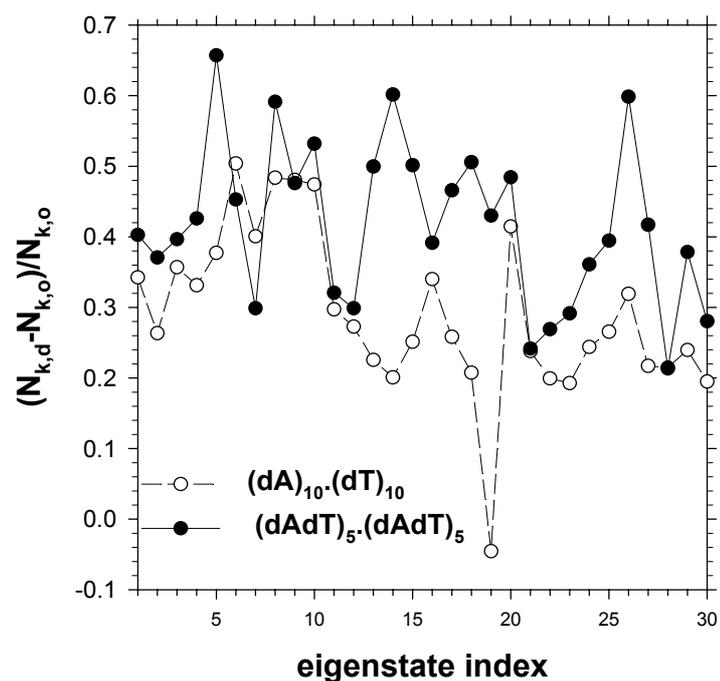


**Figure 6:** Contribution of the  $S_1$  state of thymidine (triangles) and the  $S_1$  (squares) and  $S_2$  (circles) states of adenosine to the eigenstates of the double-stranded decamers  $(dA)_{10} \cdot (dT)_{10}$  and  $(dAdT)_5 \cdot (dAdT)_5$ . In black: average values extracted from the molecular dynamics simulations. In grey: ordered conformations.

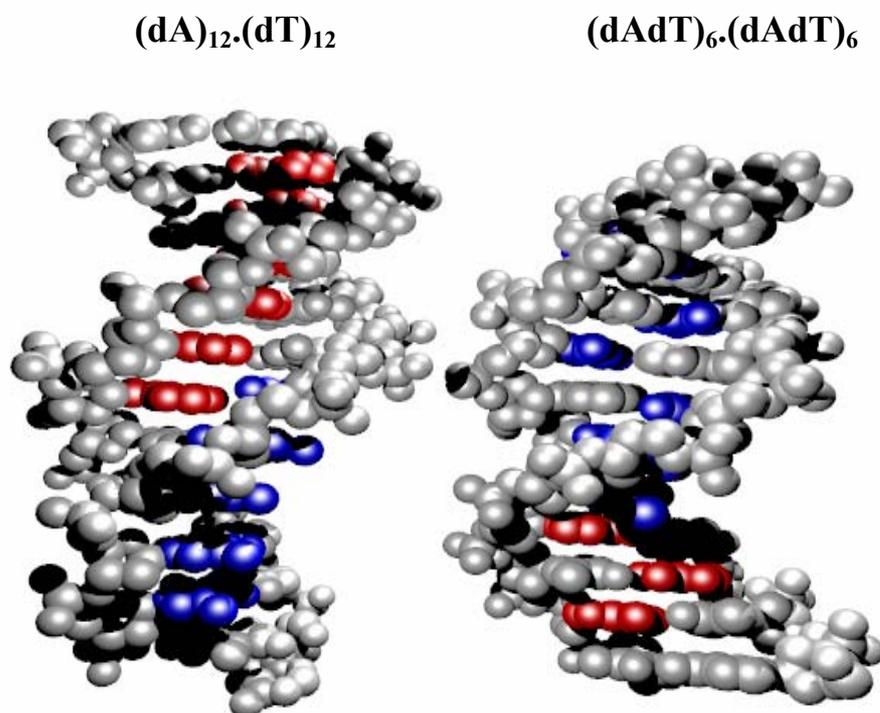
Regarding the spatial extent of the eigenstates, a first remark is that none of the eigenstates of the disordered oligomers examined are localized on a single base. The participation ratios are all higher than two and, on average, they range between four and eight depending on the eigenstate index. In Figure 7, we show that although the average  $N_k$  values found for disordered oligomers are relatively high, they are nearly all lower than the  $N_k$  values of the same eigenstates of the corresponding ordered oligomers. We express the reduction of the spatial extent of the excitations induced by structural fluctuations using the quantity  $(N_{ko}-N_{kd})/N_{ko}$ , where the subscripts d and o designate average disordered and ordered conformations, respectively. This quantity is plotted in Figure 8, where we can see that for nearly all the eigenstates, the decrease in the participation ratio is larger for the alternating oligomer  $(dAdT)_5.(dAdT)_5$ . The sensitivity originates from more important structural fluctuations leading to larger off-diagonal disorder, as illustrated in §3.



**Figure 7:** Influence of the double helix dynamics on the participation ratio corresponding to the eigenstates of  $(dA)_{10}-(dT)_{10}$  and  $(dAdT)_5-(dAdT)_5$ . In black: average values for 100 conformations. In grey: ordered conformations.



**Figure 8:** Reduction of the spatial extent of the eigenstates induced by structural fluctuations.  $N_k$  is the participation ratio, the subscripts o and d designate the values obtained for ordered and an average of 100 dynamically disordered conformations.

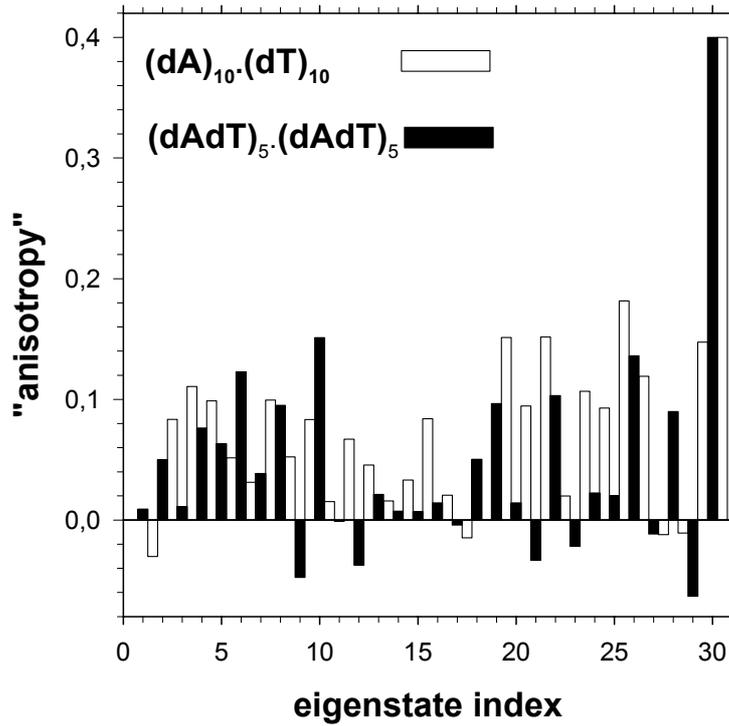


**Figure 9:** Excited bases corresponding to the eigenstates  $\langle 11 \rangle$  (in blue) and  $\langle 1 \rangle$  (in red) of one disordered configuration of the double-stranded homo-oligomer (left) and the alternating oligomer (right). The oligomers shown correspond to the dodecamers used in the dynamics simulations and, with respect to those used for the calculation of eigenstate properties, have an additional base pair at each end.

In the case of ordered oligomers, eigenstates extend along the double strand in a more or less symmetrical and/or periodic way (*cf.* Figure 9 in reference<sup>19</sup>). As expected, structural fluctuations break down such a regular topography, so that different eigenstates may involve different parts of the oligomers. An example is shown on Figure 9, where the bases participating in eigenstates  $\langle 11 \rangle$  and  $\langle 1 \rangle$  of the homo-oligomer and the alternating oligomer are represented in blue and in red, respectively.

Finally, we focus on the polarization of the  $\langle 0 \rangle \rightarrow \langle k \rangle$  transition associated to each eigenstate  $\langle k \rangle$ . By analogy with the fluorescence anisotropy used as evidence of excitation transfer in molecular systems, we define the “anisotropy” associated with each eigenstate  $\langle k \rangle$  with respect to a reference eigenstate  $\langle \lambda \rangle$ , as  $r_{k,\lambda} = 1/5(3\cos^2\theta_{k,\lambda} - 1)$ , where  $\theta_{k,\lambda}$  is the angle formed between the polarization of the transitions  $\langle 0 \rangle \rightarrow \langle k \rangle$  and  $\langle 0 \rangle \rightarrow \langle \lambda \rangle$ . Figure 10 shows the mean eigenstate “anisotropy” over 100 conformations of  $(dA)_{10} \cdot (dT)_{10}$  and  $(dAdT)_5 \cdot (dAdT)_5$ . We remark that for all eigenstates  $\langle k \rangle \neq \langle 30 \rangle$ , the  $r_{k,30}$  values are lower than 0.2. The same result is found if another eigenstate is taken as a reference.

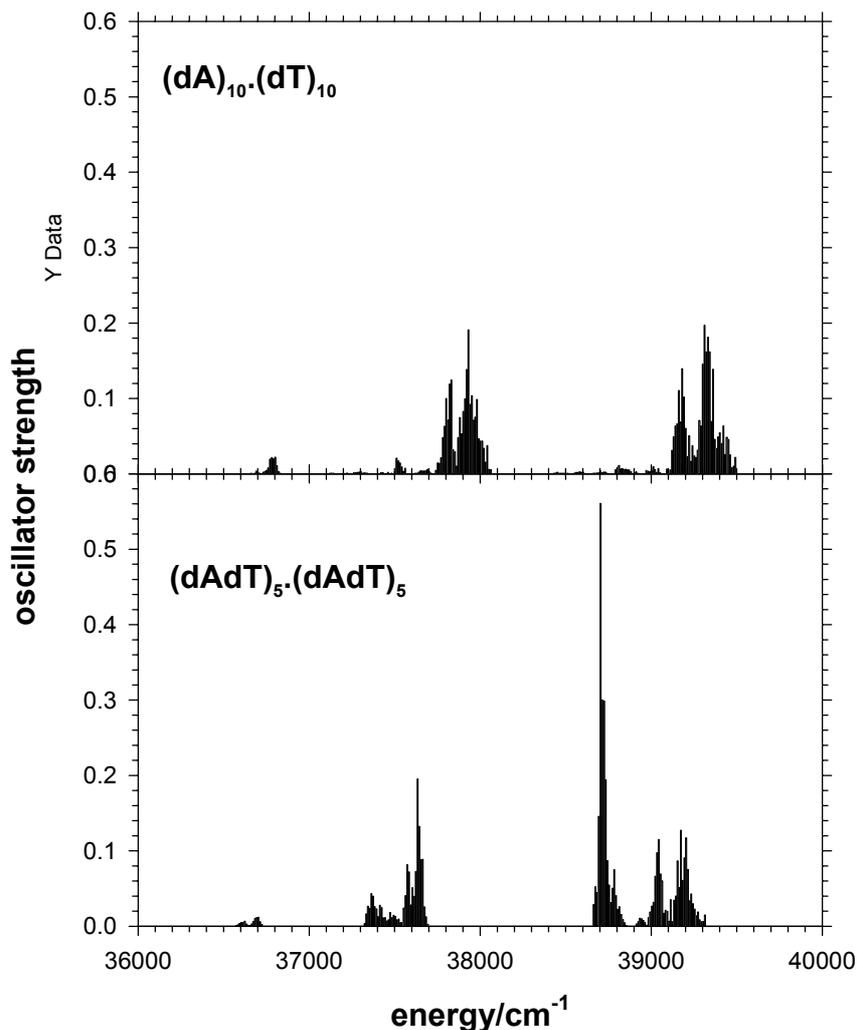
Remark: The eigenstate properties presented above have been calculated under the assumption that the monomer electronic transitions are delta functions. In the experimental spectra of the nucleosides in solution (*cf.* Figure 2 in reference<sup>19</sup>), the absorption bands corresponding to the three electronic transitions considered in the present work greatly overlap. As a result, we expect a much more efficient mixing of the three monomer electronic states within the eigenstates. This effect will tend to increase delocalization in spite of a much larger dispersion of the diagonal energy. Regarding the electronic coupling, the terms corresponding to the edge of the monomer absorption bands would decrease, whereas those related to the neighborhood of the absorption maximum would increase, thus compensating each other.



**Figure 10:** “Anisotropy” associated with each eigenstate  $\langle k \rangle$  with respect to the highest eigenstate  $\langle 30 \rangle$ . This is defined as  $r_{i,30} = 1/5(3\cos^2\theta_{i,30} - 1)$ , where  $\theta_{i,30}$  is the angle formed between the polarization of the transitions  $\langle 0 \rangle \rightarrow \langle k \rangle$  and  $\langle 0 \rangle \rightarrow \langle 30 \rangle$ .

### 3.3 Spectral Properties: comparison between theory and experiment

The absorption “spectra” calculated for  $(dA)_{10} \cdot (dT)_{10}$  and  $(dAdT)_5 \cdot (dAdT)_5$  are represented in Figure 11. They have been simulated by plotting the oscillator strength corresponding to the thirty eigenstates of one of the 100 disordered conformations. We can distinguish five main “bands” in the  $(dAdT)_5 \cdot (dAdT)_5$  “spectrum”, whereas only four bands appear in that of  $(dA)_{10} \cdot (dT)_{10}$ .



**Figure 11:** Average absorption “spectra” calculated for 100 conformations of  $(dA)_{10} \cdot (dT)_{10}$  and  $(dAdT)_5 \cdot (dAdT)_5$ . The total oscillator strength is normalized per base pair; the width of the individual subdivisions equals  $10 \text{ cm}^{-1}$ .

In order to quantify the relative position of the calculated “spectra” on the energy axis, we use as a criterion the spectrum barycenter, defined as  $\nu_b = \sum f_k E_k / \sum f_k$ . The  $\nu_b$  values found for the two types of oligomers,  $(dA)_n \cdot (dT)_n$  and  $(dAdT)_{n/2} \cdot (dAdT)_{n/2}$ , ordered or disordered, are shown in Table 2. They are all located in the energy range  $38500\text{-}39000 \text{ cm}^{-1}$ , that is, a few hundred wavenumbers higher in energy than the value found for the three transitions of the two independent chromophores ( $38013 \text{ cm}^{-1}$ ). We remark that for both oligomers, the structural fluctuations tend to shift the absorption barycenter to lower energies, but this change is more important with the homopolymeric sequence ( $356 \text{ cm}^{-1}$ ) than with the

alternating sequence ( $281 \text{ cm}^{-1}$ ). An opposite, but smaller, shift is observed when the length of the oligomers is increased from 10 to 20 base pairs.

System	Type	Barycenters	Maxima <sup>a</sup>
dA + dT	experimental <sup>b</sup>	38013	38168
(dA) <sub>20</sub> .(dT) <sub>20</sub>	experimental		38580
(dA) <sub>20</sub> .(dT) <sub>20</sub>	ordered	38983	
(dA) <sub>10</sub> .(dT) <sub>10</sub>	ordered	38891	
(dA) <sub>10</sub> .(dT) <sub>10</sub>	disordered	38601	
(dAdT) <sub>10</sub> .(dAdT) <sub>10</sub>	experimental		38110
(dAdT) <sub>10</sub> .(dAdT) <sub>10</sub>	ordered	38916	
(dAdT) <sub>5</sub> .(dAdT) <sub>5</sub>	ordered	38857	
(dAdT) <sub>5</sub> .(dAdT) <sub>5</sub>	disordered	38501	

<sup>a</sup> experimental error:  $\pm 30 \text{ cm}^{-1}$ , <sup>b</sup> derived from decomposition of the experimental spectra as explained in reference 19

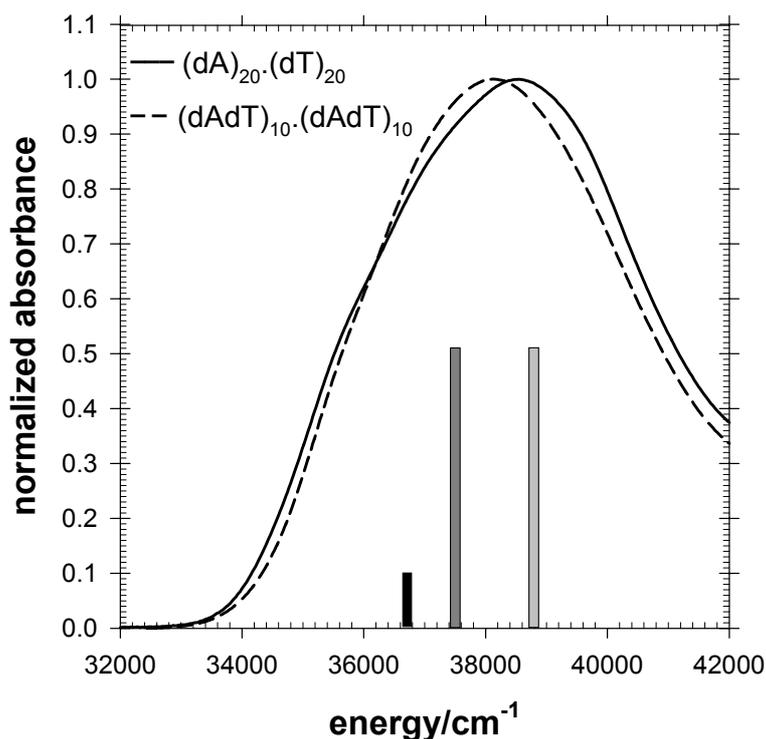
**TABLE 2: Absorption characteristics**

In Figure 12, we compare the profiles of the absorption spectra recorded at room temperature for aqueous solutions of (dA)<sub>20</sub>.(dT)<sub>20</sub> and (dAdT)<sub>10</sub>.(dAdT)<sub>10</sub>.<sup>§</sup> It can clearly be

---

<sup>§</sup> The nucleotides were purchased from Sigma Aldrich. The double stranded oligomers (dA)<sub>20</sub>.(dT)<sub>20</sub> and (dAdT)<sub>10</sub>.(dAdT)<sub>10</sub> were obtained from Eurogentec; they were purified by polyacrylamide gel electrophoresis (PAGE). All measurements were carried out in a phosphate buffer (pH=6.8; 0.1 M NaH<sub>2</sub>PO<sub>4</sub>, 0.1 M Na<sub>2</sub>HPO<sub>4</sub>, 0.25 M NaCl) at room temperature ( $20 \pm 1 \text{ }^\circ\text{C}$ ). Absorption spectra were recorded using a Perkin Lambda 900 spectrophotometer.

seen that the spectrum of the homo-oligomer is hypsochromically shifted with respect to that of the alternating oligomer. The difference of the absorption maxima is  $470\text{ cm}^{-1}$ . The absorption maximum of  $(\text{dA})_{20}(\text{dT})_{20}$  is located  $412\text{ cm}^{-1}$  toward higher energies as compared to the absorption maximum of the sum of the monomer spectra ( $\text{dA} + \text{dT}$ ). In contrast, to the precision of the measurements, the maximum of the  $(\text{dAdT})_{10}(\text{dAdT})_{10}$  spectrum is the same as that corresponding to the sum of the monomer spectra.



**Figure 12:** Normalized absorption spectra of  $(\text{dA})_{20}(\text{dT})_{20}$  (solid line) and  $(\text{dAdT})_{10}(\text{dAdT})_{10}$  (dotted line) in aqueous buffered solutions. The vertical bars represent the position and oscillator strength corresponding to the  $S_0 \rightarrow S_1$  (black) and  $S_0 \rightarrow S_2$  adenosine (grey) and  $S_0 \rightarrow S_1$  thymidine (dark grey) electronic transitions.

Prior to any comparison between calculated and experimental absorption features, we must consider at least two important issues. Firstly, there is a lower limit to the possible agreement between the calculated absorption barycenters and the experimental absorption maxima, intrinsic to the way in which these quantities are determined. This limit can be determined from the difference between the  $\nu_b$  and  $\nu_{\text{max}}$  values corresponding to the sum of

the monomers, both derived from the absorption spectra of aqueous solutions of nucleosides, and is equal to  $150 \text{ cm}^{-1}$ . Secondly, the energy of the chromophore transitions used for the determination of the diagonal terms of the Hamiltonian matrix is that corresponding to the nucleosides in aqueous solutions and not that corresponding to the gas phase, because the latter values are not known. However, on the basis of experiments performed for 9-methyladenine, we can estimate that the energy difference between the gas and aqueous phase amount to a few hundred wavenumbers.<sup>32,33</sup> Thirdly, an increase in the length of the oligomer may induce structural fluctuations of larger amplitude.<sup>34</sup>

With these limitations in mind, our model predicts that the shift of the oligomer spectra with respect to that of the sum of the monomers should not be larger than a few hundred wavenumbers (*i.e.*, no more than three nanometers in the considered spectral area), which is consistent with experimental observations. Previous theoretical studies predicted a shift of several thousand wavenumbers.<sup>35</sup> Moreover, according to our model the homopolymeric decamer absorbs at higher energies than the alternating decamers, in agreement with the spectra shown in Figure 12. Finally, not only the relative position of the spectra presented here, but also the absolute values of the calculated absorption characteristics are in accord with experiment. In fact, the maxima of experimental spectra of all the oligomers are located in the  $38100\text{-}38600 \text{ cm}^{-1}$  range, whereas the barycenters are found between  $38500$  and  $39000 \text{ cm}^{-1}$ . This agreement between experimental and theoretical values, which improves when structural fluctuations are considered, is the best observed so far.

#### 4. Conclusions and Comments

The main conclusions of the present work, where the influence of conformational dynamics on the exciton states of DNA oligomers was examined for the first time, can be summarized as follows.

Although structural fluctuations reduce both the mixing between different monomer transitions and the spatial extent of the eigenstates, excitations still remain delocalized over several bases. Such behavior, which contrasts with the complete charge localization reported quite recently,<sup>36</sup> originates from the nature of the electronic coupling responsible for the formation of Frenkel excitons. Actually, dipolar coupling operates over relatively long distances and is not restricted to nearest neighbors as is the coupling due to orbital overlap associated with charge delocalization. In the presence of structural fluctuations, some interatomic distances involving different DNA bases increase, but others decrease;

consequently, some terms of the dipolar coupling (which is a function of the atomic charge distribution) decrease, whereas others increase, preserving the delocalization of the excitation. The same reasoning could be applied for DNA fragments within nucleosomes, if the electronic structure of the bases is not seriously perturbed by interactions with the core histones.

The sensitivity of the spatial extent of the excitons towards structural fluctuations depends on the base sequence. It would be interesting to examine other sequences in order to establish general trends regarding this behavior.

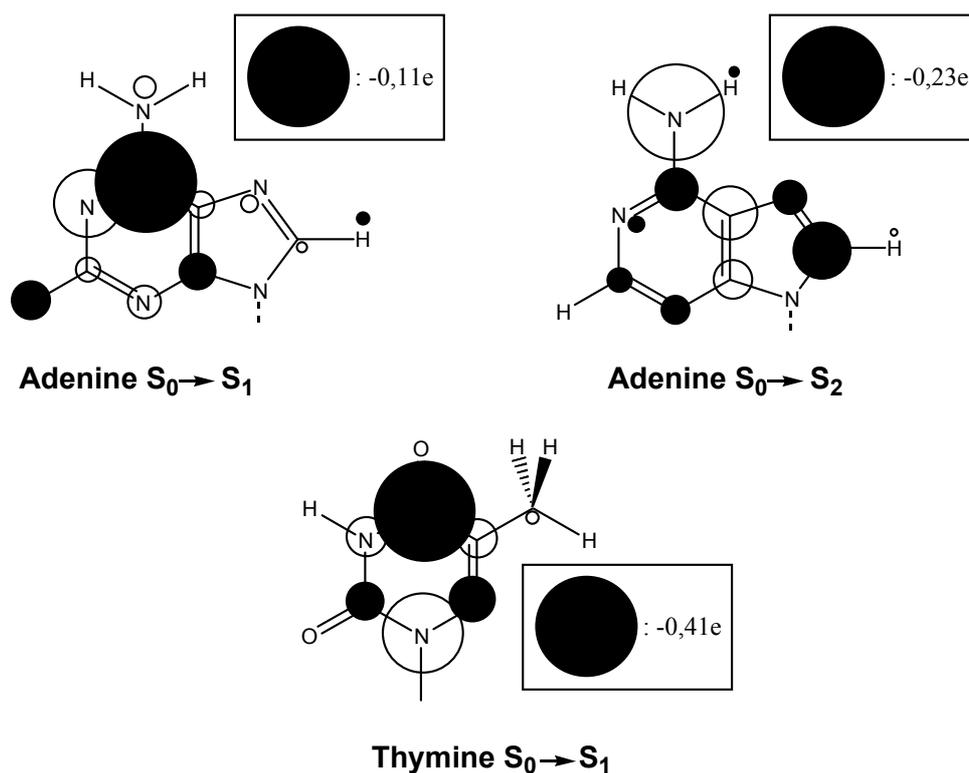
The formation of collective excited states in double-stranded oligomers is not expected to induce large spectral shifts with respect to a homogeneous mixture of monomers. Thus, it is not surprising that collective states are not easily detectable on UV spectra, in particular when these are recorded on a nanometer scale.

### **Acknowledgement**

Financial support from the CEA program "Toxicologie Nucléaire" and the ACI "Physicochimie de la matière complexe" is gratefully acknowledged. We thank Dr Michel Mons for helpful discussions.

### **Appendix**

*Calculation of the excited state charges.* As AMBER charges are only available for ground-state bases, atomic charges for the excited states of the monomers thus had to be constructed from *ab initio* calculations. We describe the change in the monomers' electronic wavefunction upon excitation as a set of atomic charge differences, computed by subtraction of the CASSCF/RESP charges on the atoms of the ground state molecule from those corresponding to the excited state. These changes are shown schematically in Figure 13 for the three transitions studied. The charges for the excited states of thymine and adenine were obtained by adding the corresponding charge difference to the standard AMBER ground-state charges. This should account for the reorganization of the electronic system of the monomers upon excitation, while retaining the generality of the AMBER force field, which is very well suited to the study of nucleic acids.<sup>37</sup>



**Figure 13:** Difference between the excited state and the ground state atomic charges corresponding to the  $S_0 \rightarrow S_1$  transition of thymine and the  $S_0 \rightarrow S_1$  and  $S_0 \rightarrow S_2$  transitions of adenine; negative and positive charges appear as black and white circles, respectively. The circle diameter is proportional to the absolute value of the change in charge.

In calculations of atomic charge differences, atomic charges for the ground and excited states of the DNA bases adenine and thymine were derived from the electrostatic potential calculated on a grid of points from the corresponding individual state CASSCF electronic wavefunction, using an extended version of the RESP method.<sup>38</sup>

The points on the grid used for charge fitting belong to an ensemble of fused spheres, centered on atoms and of radius  $\gamma R$ ,  $R$  being the atomic van der Waals radius of each individual atom. The grid was generated using an algorithm due to Spackman.<sup>39</sup> The value of 1.4 for  $\gamma$  was found to provide the best precision while preventing the so-called “hidden atom” phenomenon, *i.e.*, indeterminations occurring during the fitting procedure.<sup>40</sup> The resulting sets were composed of 4151 points for thymine and 3897 for adenine.

Permanent dipole moments equivalent to the charge distributions were compared with their CASSCF counterparts to assess the quality of the fit, and may be found in Table 3. Both the norm and the direction of these dipole moments show excellent agreement with the *ab initio* results.

	$\ \bar{\mu}\ $ (D)			$\Theta$ ( $^\circ$ ) <sup>a</sup>		
	CASSCF	Atomic Charges	Error (%)	CASSCF	Atomic Charges	Error (%)
<b>Adenine S<sub>0</sub></b>	2.42	2.42	< 0.1	-89.6	-88.9	0.8
<b>Adenine S<sub>1</sub></b>	2.10	2.09	0.3	-85.7	-84.3	1.6
<b>Adenine S<sub>2</sub></b>	2.52	2.53	0.7	-26.0	-25.6	1.5
<b>Thymine S<sub>0</sub></b>	4.06	4.09	0.8	-124.1	-126.0	1.5
<b>Thymine S<sub>1</sub></b>	6.96	6.94	0.4	-169.2	-166.4	1.7

<sup>a</sup> According to the DeVoe-Tinoco convention<sup>45</sup>

**TABLE 3: Norms ( $\|\bar{\mu}\|$ ) and angles ( $\Theta$ ) of the permanent dipole moments of the electronic states of adenine and thymine**

Non-linear solutions of the Poisson-Boltzmann equation were obtained with the DELPHI program (version 2.1).<sup>41</sup> An ionic strength of 0.145 was employed. Calculations involved 145 grid points and 80% of the box filled by the molecule, corresponding to 2.7 grid points/Å. The internal dielectric constant was taken as 2, and the external value as 80. The solvent accessible surface of the molecule was defined using a probe radius of 1.4 Å.

**Computational Details of Quantum Chemistry Calculations.** The geometry of adenine and thymine, described by their most abundant tautomeric forms, were optimized at the AM1 level. CASSCF calculations on the monomers ( $C_s$  point group, cc-pVDZ basis set<sup>42</sup>) used the  $\pi$  subsystem as the active space (12 electrons in 11 molecular orbitals for adenine, 12 electrons in 10 molecular orbitals for thymine). They were carried out using the MOLPRO 2002 package.<sup>43</sup> The grid of points used to fit the electrostatic potential was created using GAMESS(US).<sup>44</sup> All calculations were run on a Compaq ES45 AlphaServer workstation. The details on the CS-INDO calculations are given in reference<sup>19</sup>.

## References

- (1) Cadet, J.; Vigny, P. The Photochemistry of Nucleic Acids. In *Bioorganic Photochemistry*; Morrison, H., Ed.; John Wiley & Sons: New York, 1990; pp 1.
- (2) Tinoco Jr., I. *J. Am. Chem. Soc.* **1960**, *82*, 4785.
- (3) Eisinger, J.; Shulman, R. G. *Science* **1968**, *161*, 1311.
- (4) Guéron, M.; Eisinger, J.; Shulman, R. G. *J. Chem. Phys.* **1967**, *47*, 4077.
- (5) Rayner, D. M.; Szabo, A. G.; Loutfy, R. O.; Yip, R. W. *J. Phys. Chem.* **1980**, *84*, 289.
- (6) Georghiou, S.; Zhu, S.; Weidner, R.; Huang, C.-R.; Ge, G. *J. Biomol. Struct.* **1990**, *8*, 657.
- (7) Huang, C.-R.; Georghiou, S. *Photochem. Photobiol.* **1992**, *56*, 95.
- (8) Marguet, S.; Markovitsi, D.; Millié, P.; Sigal, H.; Kumar, S. *J. Phys. Chem. B* **1998**, *102*, 4697.
- (9) Markovitsi, D.; Gallos, L. K.; Lemaistre, J. P.; Argyrakakis, P. *Chem. Phys.* **2001**, 269.
- (10) Kasha, M.; Rawls, H. R.; El-Bayoumi, M. A. *Pure and Applied Chemistry* **1965**, *11*, 371.
- (11) Scholes, G. D.; Ghiggino, K. P. *J. Phys. Chem.* **1994**, *98*, 4580.
- (12) Scholes, G. D.; Harcourt, R. D.; Ghiggino, K. P. *J. Chem. Phys.* **1995**, *102*, 9574.
- (13) Pecourt, J.-M. L.; Peon, J.; Kohler, B. *J. Am. Chem. Soc.* **2000**, *122*, 9348.
- (14) Peon, J.; Zewail, A. H. *Chem. Phys. Lett.* **2001**, *348*, 255.
- (15) Onidas, D.; Markovitsi, D.; Marguet, S.; Sharonov, A.; Gustavsson, T. *J. Phys. Chem. B* **2002**, *106*, 11367.
- (16) Gustavsson, T.; Sharonov, A.; Markovitsi, D. *Chem. Phys. Lett.* **2002**, *351*, 195.
- (17) Gustavsson, T.; Sharonov, A.; Onidas, D.; Markovitsi, D. *Chem. Phys. Lett.* **2002**, *356*, 49.
- (18) Markovitsi, D.; Onidas, D.; Sharonov, A.; Gustavsson, T. *ChemPhysChem* **2003**, *3*, 303.
- (19) Bouvier, B.; Gustavsson, T.; Markovitsi, D.; Millié, P. *Chem. Phys.* **2002**, *275*, 75.
- (20) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Metz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weimer, P. K.; Kollman, P. A. *AMBER 6; University of California: San Francisco.* **1999**.
- (21) Cheatham, T. E.; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845.
- (22) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
- (23) Ryckaert, J. P.; Ciccotii, G.; Berendsen, H. J. C. *J. Comp. Phys.* **1977**, *23*, 327.
- (24) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
- (25) Frenkel, J. *Phys. Rev.* **1931**, *37*, 1276.
- (26) Rashbah, E. I.; Sturge, M. D. *Excitons*; North-Holland: Amsterdam, 1982.
- (27) Davydov, A. S. *Theory of Molecular Excitons*; Plenum Press: New York, 1971.
- (28) Longuet-Higgins, H. C. *Proc. R. Soc. A* **1956**, *235*, 537.
- (29) Momicchioli, F.; Baraldi, I.; Bruni, M. C. *Chem. Phys.* **1983**, *82*, 229.
- (30) Dean, P. *Rev. Mod. Phys.* **1972**, *44*, 127.
- (31) Schreiber, M.; Toyosawa, Y. *J. Phys. Soc. Jpn.* **1982**, *51*, 1537.
- (32) Plutzer, C.; Nir, E.; de Vries, M. S.; Kleinermanns, K. *Phys. Chem. Chem. Phys.* **2001**, *3*, 5466.

- (33) Nir, E.; Kleinermanns, K.; Grace, L.; Vries, M. S. d. *J. Phys. Chem. A* **2001**, *105*, 5106.
- (34) Ha Duong, T.; Zakrzewska, K. *J. Comp. Chem.* **1997**, *18*, 796.
- (35) Miyata, T.; Yomosa, S. *J. Phys. Soc. Jpn.* **1969**, *27*, 727.
- (36) Lewis, J. P.; Cheatham, T. E., III; Starikov, E. B.; Wang, H.; Sankey, O. F. *J. Phys. Chem. B* **2003**, *107*, 2581.
- (37) Giudice, E.; Lavery, R. *Acc. Chem. Res.* **2002**, *35*, 350.
- (38) Levy, B.; Enescu, M. *Theochem-J. Mol. Struct.* **1998**, *432*, 235.
- (39) Spackman, M. A. *J. Comput. Chem.* **1996**, *17*, 1.
- (40) Bayly, C. L.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.
- (41) Sharp, K. A.; Honig, B. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301.
- (42) Dunning Jr., T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
- (43) Werner, H.-J.; Knowles, P. J.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Celani, P.; Cooper, D. L.; Deegan, M. J. O.; J., D. A.; Eckert, F.; Hampel, C.; Hetzer, G.; Korona, T.; Lindh, R.; Lloyd, A. W.; McNicholas, S. J.; Manby, F. R.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Rauhut, G.; Schütz, M.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T. MOLPRO 2002.3, 2002.
- (44) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.
- (45) DeVoe, H.; Tinoco, I. *J. Mol. Biol.* **1962**, *4*, 518.

#### 4. États Excités Collectifs de Fragments d'ADN : Approche Dynamique



# 5. Dynamique Semi-Classique du Transfert de l'Excitation

## 5.1. Introduction

L'adjonction progressive des aspects dynamiques du transfert d'excitation au modèle présenté ici constitue sans doute la voie d'évolution la plus intéressante pour ce travail. Une première application à des brins d'ADN de géométrie statique a permis la mise au point et la validation du modèle. L'approche dynamique a été introduite, dans un second temps, en raisonnant non plus sur des configurations idéalisées mais sur des structures issues de calculs de dynamique moléculaire à l'état fondamental, soumises à une excitation verticale. Au cours de cette phase, la dynamique moléculaire classique a été employée comme outil d'échantillonnage de la surface de potentiel de l'agrégat neutre. En revanche, la succession temporelle des configurations issues de la dynamique n'est pas directement exploitée : ces structures sont envisagées sous la forme d'un ensemble statistique, donnant accès à des propriétés moyennées comparables aux mesures expérimentales. De plus, l'étude dynamique est restreinte à l'état fondamental du système : aucune information temporelle n'est disponible sur la relaxation post-excitation électronique du système. L'accès à cette information nécessite non seulement de pouvoir mener un calcul de dynamique moléculaire sur les surfaces d'énergie potentielle des états excités, mais également de savoir rendre compte des phénomènes de conversion entre énergies électronique et vibrationnelle qui ont lieu au cours de la désexcitation non radiative du système. Cela implique la prise en compte de degrés de liberté quantiques (l'état électronique du système à un instant donné), et leur couplage avec les degrés de liberté classiques (énergies cinétiques de translation et de rotation des molécules).

Parmi les méthodes de dynamique semiclassiques adaptées à ce type d'étude, la méthode à saut de surface (*Trajectory Surface Hopping* ou TSH), due à Tully<sup>181-184</sup>, a été retenue pour ce travail. En effet, le formalisme d'hamiltonien effectif et la description des états adiabatiques d'un système comme combinaison linéaire d'états diabatiques y trouvent naturellement leur place. La méthode TSH est décrite *in extenso* en première partie de ce mémoire.

Dans l'état actuel des méthodes et moyens de calcul, l'application directe de la méthode TSH à un fragment d'ADN de taille raisonnable est inenvisageable, et ce pour (au moins) deux raisons.

La première est liée au coût en temps de calcul. La constitution d'une matrice hamiltonienne de dimension  $N$  requiert  $N$  calculs d'énergie d'interaction intermoléculaire (l'énergie des formes à excitation localisée) et  $\frac{1}{2}N(N-1)$  calculs de couplage. Or, cette matrice doit être constituée, en toute rigueur, à chaque pas de temps quantique (ledit pas de temps quantique étant *a priori* bien inférieur à celui de la dynamique des noyaux). Certes, il est possible, comme dans cette étude, d'interpoler les éléments de matrice entre pas de temps classiques ; cette approximation, généralement bonne, peut cependant nécessiter une réduction de ce pas de temps. Le coût de la détermination des couplages non adiabatiques et de l'intégration de l'équation de Schrödinger dépendant du temps doivent également être prises en compte. Enfin, la nécessité d'une double statistique sur les configurations initiales et le générateur de nombres aléatoires impose de reproduire un grand nombre de fois chaque simulation. En conséquence, et même avec un potentiel intermoléculaire et des couplages simples, le traitement de systèmes étendus est difficilement envisageable.

La seconde raison est méthodologique et découle de la première. L'excitation d'un brin d'ADN concerne uniquement les bases azotées, et n'influe pas sur le squelette sucre-phosphate. Il est donc tentant de remplacer ce dernier par un 'ressort' équivalent, qui pourrait également intégrer les effets du solvant et des ions dont la prise en compte dans une dynamique moléculaire sur l'ADN est indispensable. Ceci permettrait de limiter le système effectivement traité à ses composants pertinents, réduisant du même coup sensiblement les temps de calcul. Cependant, et malgré quelques travaux encourageants<sup>267</sup>, une telle modélisation de l'ADN n'en est pour l'instant qu'à ses débuts.

Les agrégats de naphthalène ont donc été choisis comme systèmes cible pour une première application de la méthode TSH dans le cadre de cette étude. Une certaine expérience de la modélisation de ces systèmes a en effet été acquise au cours des études de transfert de charge présentées en deuxième partie ; de surcroît, et contrairement au benzène, les premières transitions électroniques du naphthalène ne sont pas interdites par symétrie. L'étude de tels systèmes devrait fournir un bon aperçu des possibilités de la méthode TSH couplée à un potentiel intermoléculaire relativement complexe (issue de la physique du solide, la méthode TSH est souvent employée avec des potentiels très simples), en vue d'extensions futures.

## 5.2. Mise en place du potentiel et des couplages

### 5.2.1. Potentiel modèle

Le potentiel modèle dont la mise au point a été présentée en deuxième partie, bien que précis et bien adapté aux agrégats aromatiques, est relativement coûteux. L'essentiel de ce coût provient du terme de polarisation, calculé de manière autocohérente. Si la prise en compte de cette contribution est indispensable dans le cas d'agrégats chargés (dans lesquels elle explique la compétition polarisation-délocalisation), son importance dans des agrégats neutres (à l'état fondamental ou excité) est sans doute beaucoup plus marginale. Pour s'en assurer, l'énergie d'interaction d'un dimère de naphthalène empilé, calculée à l'aide de ce potentiel modèle, a été découpée en ses diverses contributions, le tout en fonction de la distance séparant les monomères (figure 5.1).

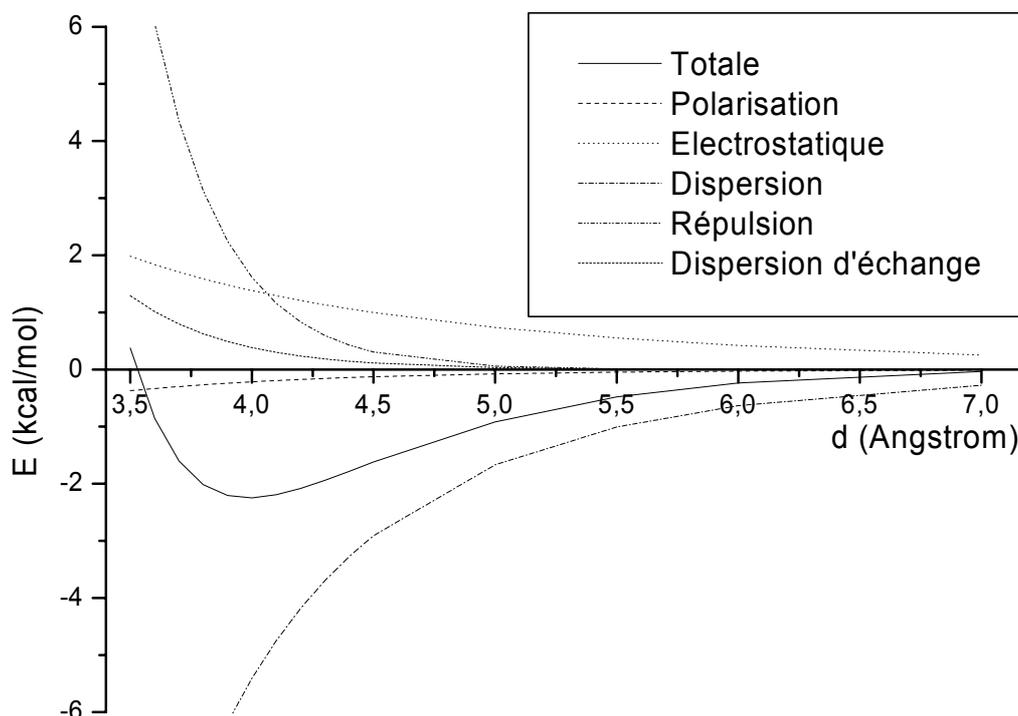


Figure 5.1.

Il est clair que, par rapport aux contributions électrostatiques ou de dispersion notamment, l'énergie de polarisation est négligeable. La figure 5.2 montre, sous forme d'un pourcentage, le rapport entre l'énergie de polarisation et la contribution la plus importante, à la distance considérée.

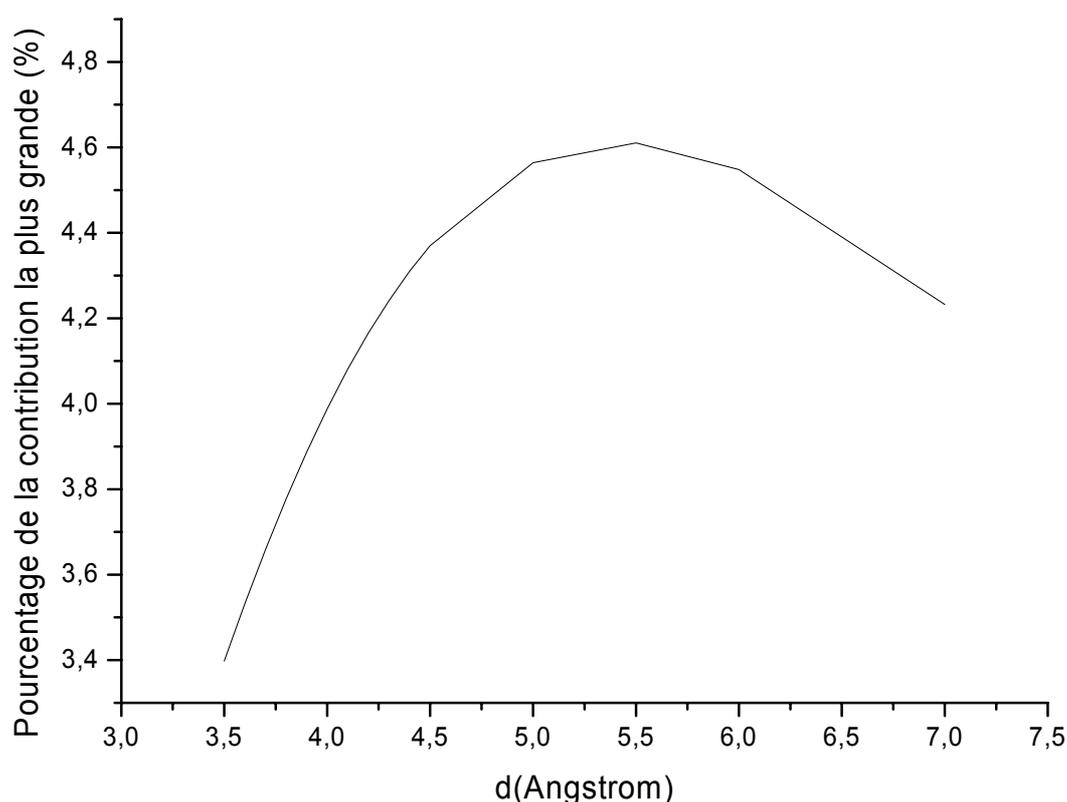


Figure 5.2.

La contribution de polarisation n'est jamais supérieure à 5 % de la contribution la plus importante, même dans la zone située entre 4,5 et 6,5 Å où son effet est le plus ressenti. Il est donc très raisonnable de négliger le terme de polarisation lors de l'étude d'agrégats de naphthalène neutres à l'état fondamental. Il en va sans doute de même pour des états excités de ces mêmes espèces, pour deux raisons : ces états ne sont pas dotés d'un caractère à transfert de charge intermoléculaire dans le présent modèle, et les moments dipolaires permanents des états électroniques du naphthalène sont nuls par symétrie.

Il a également été décidé de simplifier le modèle électrostatique, afin de n'utiliser qu'une distribution de monopôles atomiques en lieu et place de la distribution de monopôles,

## 5. Dynamique Semi-Classique du Transfert de l'Excitation

dipôles et quadrupôles sur les atomes et les milieux de liaisons habituellement utilisée au niveau potentiel modèle. Aux courtes distances d'approche, l'interaction est dominée par les effets de dispersion et de répulsion ; la contribution électrostatique à l'énergie d'interaction ne se manifeste donc réellement qu'aux moyennes et longues distances, pour lesquelles la perte de précision liée à l'emploi d'une distribution simplifiée est faible. Le gain en temps de calcul réalisé constitue une motivation additionnelle (la rotation des dipôles et quadrupôles locaux au cours de la dynamique peut se révéler assez coûteuse).

Comme précédemment dans le cas de l'ADN (paragraphe 4), la méthode *Eigenvalue Analysis* appliquée à la densité électronique CASSCF de l'état fondamental et des états excités du naphthalène a été employée pour obtenir les distributions de charge atomiques. L'espace actif des calculs MCSCF par états est le système  $\pi$  (10 électrons dans 10 orbitales) ; ils sont menés en base cc-pVDZ<sup>264</sup> dans le groupe ponctuel  $D_{2h}$ . Les énergies et forces d'oscillateur associées aux transitions  $S_0 \rightarrow S_1$  et  $S_0 \rightarrow S_2$  ainsi déterminées sont compilées dans le tableau 5.1 ci-dessous.

Transition	$\Delta E$ (CASSCF) (eV)	$\Delta E$ (CASPT2) (eV)	f (CASSCF)	f (CASPT2)
$S_0 \rightarrow S_1$	4,23	3,95	$1.10^{-4}$	$1.10^{-4}$
$S_0 \rightarrow S_2$	7,91	7,27	1,93	1,78

Tableau 5.1.

La première transition est quasiment interdite, la seconde au contraire est permise. L'ajustement des distributions de charges atomiques s'effectue par rapport aux valeurs du potentiel électrostatique *ab initio* sur une grille de points de  $\gamma = 1,4$  (4440 points pour le naphthalène). Le premier moment permanent non nul pour chaque état est ici le moment quadrupolaire ; la valeur *ab initio* du tenseur quadrupolaire (en Buckingham) est comparée au quadrupôle équivalent à la distribution de monopoles atomiques dans le tableau 5.2 (l'emploi de tenseurs sans trace rend la donnée de l'élément *zz* redondante).

État	<i>Ab Initio</i>		Monopôles	
	<i>xx</i>	<i>yy</i>	<i>xx</i>	<i>yy</i>
<b>S<sub>0</sub></b>	6,15	6,03	5,97	5,92
<b>S<sub>1</sub></b>	5,49	5,73	5,39	5,61
<b>S<sub>2</sub></b>	3,70	6,34	3,68	6,23

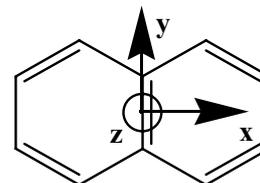


Tableau 5.2.

Le quadrupôle moléculaire est globalement bien reproduit par la distribution de charges atomiques.

Afin de compenser l'absence du terme de polarisation et la simplification du modèle électrostatique, les valeurs des paramètres atomiques régissant la répulsion et la dispersion (voir le paragraphe 1.3.4, deuxième partie) ont été légèrement retouchées, en référence à des calculs *ab initio* MP2/6-311G(2d,2p) corrigés de BSSE : la valeur de  $k$  pour le carbone aromatique vaut 1,03 contre 1,11 précédemment ; celle de  $R$  vaut 1,77 contre 1,78 auparavant. La comparaison des énergies d'interaction *ab initio* (RHF et MP2/6-311G(2d,2p) corrigés de BSSE) et potentiel modèle dans des dimères de naphthalène empilés est l'objet de la figure 5.3.

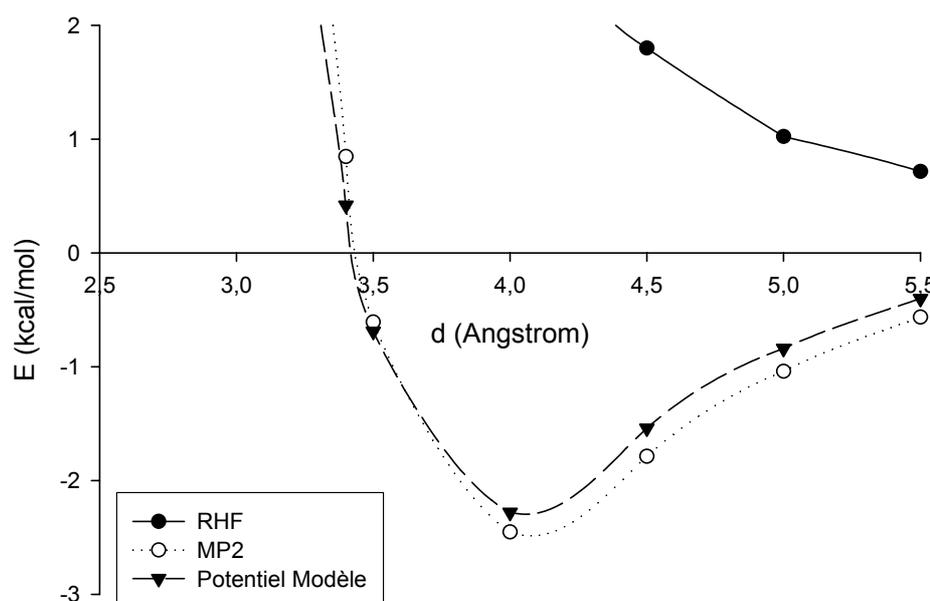


Figure 5.3

## 5. Dynamique Semi-Classique du Transfert de l'Excitation

Ici encore, le calcul Hartree-Fock ne peut rendre compte des effets de dispersion liés à la corrélation électronique, et prédit la dissociation de l'agrégat de naphthalène<sub>2</sub> empilé. Les niveaux MP2 et potentiel modèle, qui retranscrivent ces effets, prévoient une configuration stable pour cet agrégat. Malgré l'absence du terme de polarisation dans le potentiel modèle, il est possible d'obtenir un bon accord avec l'énergie d'interaction MP2 à toute distance.

### 5.2.2. Couplages

La détermination des couplages, c'est-à-dire des termes non diagonaux de la matrice excitonique, s'effectue par le biais de l'interaction de distributions de charges atomiques de transition sur les deux entités échangeant l'excitation (paragraphe 3.2). Celles-ci sont fournies, comme précédemment, par la méthode CS-INDO-CIPSI. La comparaison des moments de transition avec les valeurs MCSCF correspondantes (en Debyes) fait l'objet du tableau 5.3 ; les deux transitions sont polarisées suivant le grand axe du naphthalène (l'axe *x* de la figure du tableau 5.2).

Transition	CASSCF	CS-INDO-CIPSI
S <sub>0</sub> →S <sub>1</sub>	0,09	0,61
S <sub>0</sub> →S <sub>2</sub>	-8,03	-9,73

Tableau 5.3.

Bien que différant sensiblement entre elles, la norme et le signe des moments dipolaires de transition aux deux niveaux de calcul laissent peu de doute quant à l'identité des états concernés. Les distributions de charge de transition sont schématisées en figure 5.4.

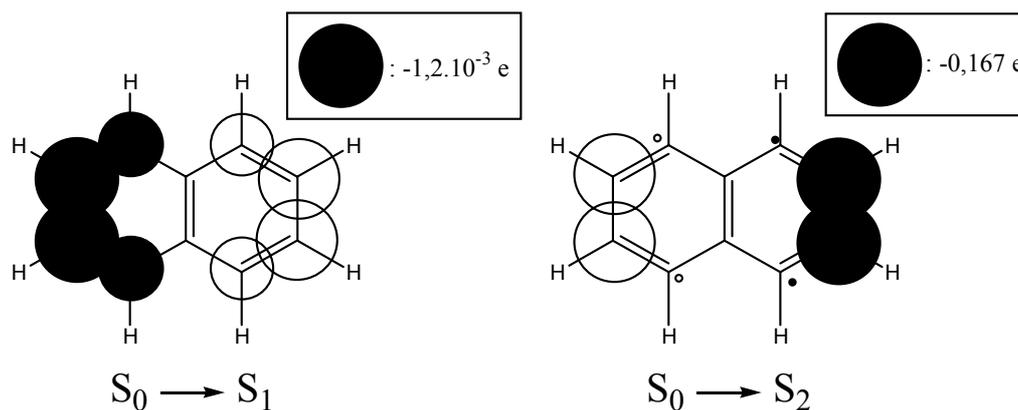


Figure 5.4.

Les sens de polarisation opposés et les intensités très différentes des deux transitions apparaissent clairement sur cette représentation schématique. En outre, ni les atomes de carbone pontants ni les atomes d'hydrogène ne jouent un quelconque rôle dans l'une ou l'autre.

### 5.3. Désexcitation non radiative du dimère de naphthalène excité

En tant que première application de la méthode TSH aux agrégats de naphthalène excités, il a semblé intéressant d'étudier la désexcitation non radiative de l'agrégat naphthalène<sub>2</sub>, selon que les états excités dudit système sont bâtis sur les états  $S_1$  ou  $S_2$  des monomères ; ces deux cas ont fait l'objet de simulations distinctes. Ici et par la suite, la possibilité d'évacuer l'énergie électronique vers des degrés de liberté intramoléculaires est négligée.

L'hamiltonien effectif du dimère, de rang deux, donne accès aux deux premiers états excités adiabatiques du système. L'agrégat est préparé, dans son second état excité électronique et à une température de 10 K, *via* le thermostat de Berendsen. Toute transition vers le premier état excité est interdite au cours de cette phase d'équilibration. Une fois les températures de translation et de rotation stabilisées, l'agrégat est autorisé à relaxer vers son premier état excité. Sont suivies, au cours du TSH, les évolutions des densités des deux états et des probabilités de saut, les températures vibrationnelle et rotationnelle de l'agrégat, l'état électronique actuel, et les énergies associées à la configuration actuelle du système sur chacune des deux surfaces d'énergie potentielle.

## 5. Dynamique Semi-Classique du Transfert de l'Excitation

Ces données sont schématisées en figure 5.5, pour une simulation de TSH typique dans laquelle les états excités du supersystème sont bâtis sur l'état électronique  $S_2$  des monomères.

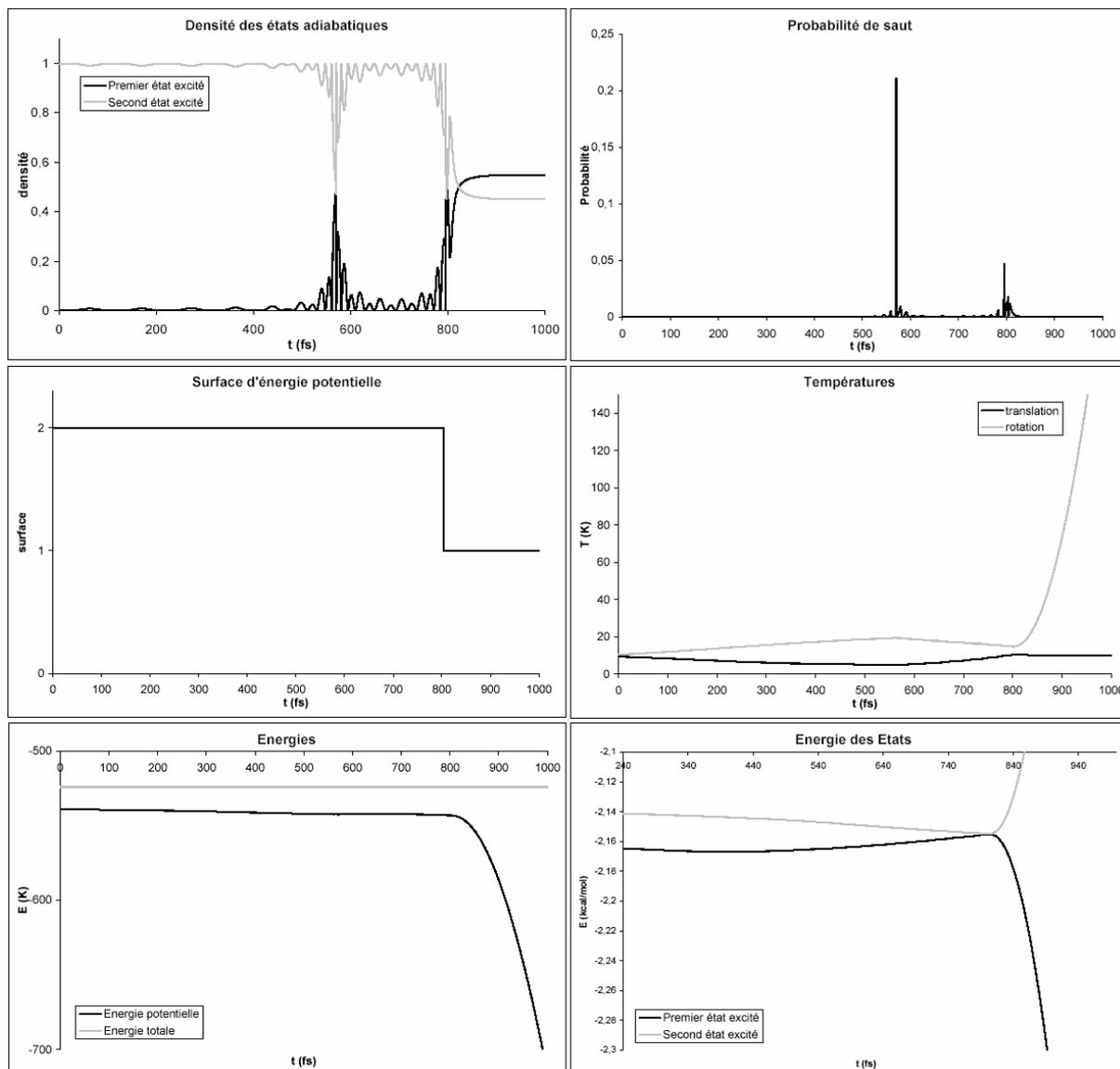


Figure 5.5.

La transition électronique vers le premier état excité se produit environ 803 fs après le début de la simulation. Aux alentours de cet instant, les degrés de liberté classiques du système sont tels que les surfaces d'énergie potentielle des deux états sont très proches énergétiquement, ce qui favorise la probabilité de saut. Environ 570 fs après l'origine temporelle, la probabilité de saut affiche également un maximum ; sa présence ne peut s'expliquer que par un couplage non adiabatique fort, étant donné que les surfaces d'énergie

potentielle des deux états sont assez distantes en ce point. Simultanément, la densité du deuxième état décroît au profit de celle du premier. Le saut n'a cependant pas lieu à 570 fs mais lors du second maximum de probabilité (803 fs). Le changement de surface d'énergie potentielle s'accompagne d'une évolution rapide des populations électroniques des deux états : le premier se peuple au détriment du second. La conservation de l'énergie imposée par l'ensemble microcanonique dans lequel la simulation a lieu se traduit par la conversion d'énergie électronique en énergie vibrationnelle. Il s'ensuit une diminution brutale de l'énergie potentielle du système après le saut. Simultanément, la température rotationnelle de l'agrégat s'élève rapidement, la température translationnelle demeurant quant à elle quasiment stable. Les modes orthogonaux à la surface d'énergie potentielle de départ, c'est-à-dire ceux qui promeuvent le saut et recueillent l'énergie électronique libérée, sont donc ici exclusivement des modes de rotation, les modes de translation (plus durs) restant spectateurs. Une fois sur la surface de potentiel du premier état excité, le système évolue vers des régions de l'espace des configurations dans lesquelles les deux surfaces d'énergie potentielle sont de topologies très différentes. Ceci se traduit par un écartement des énergies des deux états au cours du temps. Au vu des différences énergétiques atteintes, le retour vers l'état deux semble peu probable dans cette région de l'espace des configurations.

Des trajectoires distinctes peuvent apparaître lorsqu'une statistique sur le générateur de nombres aléatoires est menée. Tant qu'aucun saut de surface n'a lieu, les trajectoires restent confondues. En revanche, les sauts d'une surface d'énergie potentielle à l'autre ont lieu à différents instants selon le générateur, différenciant les trajectoires pour le reste de la simulation. Celles-ci peuvent également différer par le nombre de sauts effectués : un ou plusieurs échanges entre surfaces peuvent avoir lieu dans les zones de croisement évités. Une statistique de 50 trajectoires a été réalisée sur l'agrégat naphthalène<sub>2</sub>, dont les états excités sont construits sur les états S<sub>2</sub> des monomères. Initialement préparé dans son deuxième état excité adiabatique à 10 K, l'agrégat relaxe au cours d'une simulation d'une durée totale de 2 ps. La figure 5.6 regroupe quelques tendances observées sur cette statistique. Le pourcentage de trajectoires sur la surface d'énergie potentielle du second état excité, initialement de 100 %, décroît avec le temps ; à 2 ps, il demeure encore des trajectoires sur cette surface, bien que la contribution majoritaire provienne de la surface d'énergie potentielle du premier état excité. Il peut donc exister des différences importantes entre trajectoires, quant au temps nécessaire à la relaxation. Ceci est corroboré par les variations importantes du nombre de sauts entre surfaces qui ont lieu au cours des 2 ps : celui-ci varie en effet de 1 à 22 ! Le nombre de trajectoires comportant un nombre de sauts impairs, c'est-à-dire celles qui

## 5. Dynamique Semi-Classique du Transfert de l'Excitation

terminent sur la surface d'énergie potentielle du premier état excité et correspondent en conséquence à des relaxations, est plus important que le nombre de trajectoires à nombre pair de sauts, qui finissent sur la surface d'énergie potentielle du second état excité, ce qui explique l'évolution générale des populations. En revanche, au sein du premier groupe, le temps de relaxation est d'autant plus long que le nombre de sauts est important, et rien n'indique que des trajectoires du deuxième groupe ne rejoignent pas le premier à des temps supérieurs à 2 ps. Par le biais des sauts, les diverses trajectoires sont amenées à parcourir des itinéraires très divers sur les deux surfaces de potentiel. Ceci apparaît clairement sur l'évolution des différences entre énergies des deux états au cours du temps, caractérisées par un nombre variable d'extrema qui sont autant de points remarquables sur l'une ou l'autre des deux surfaces d'énergie potentielle.

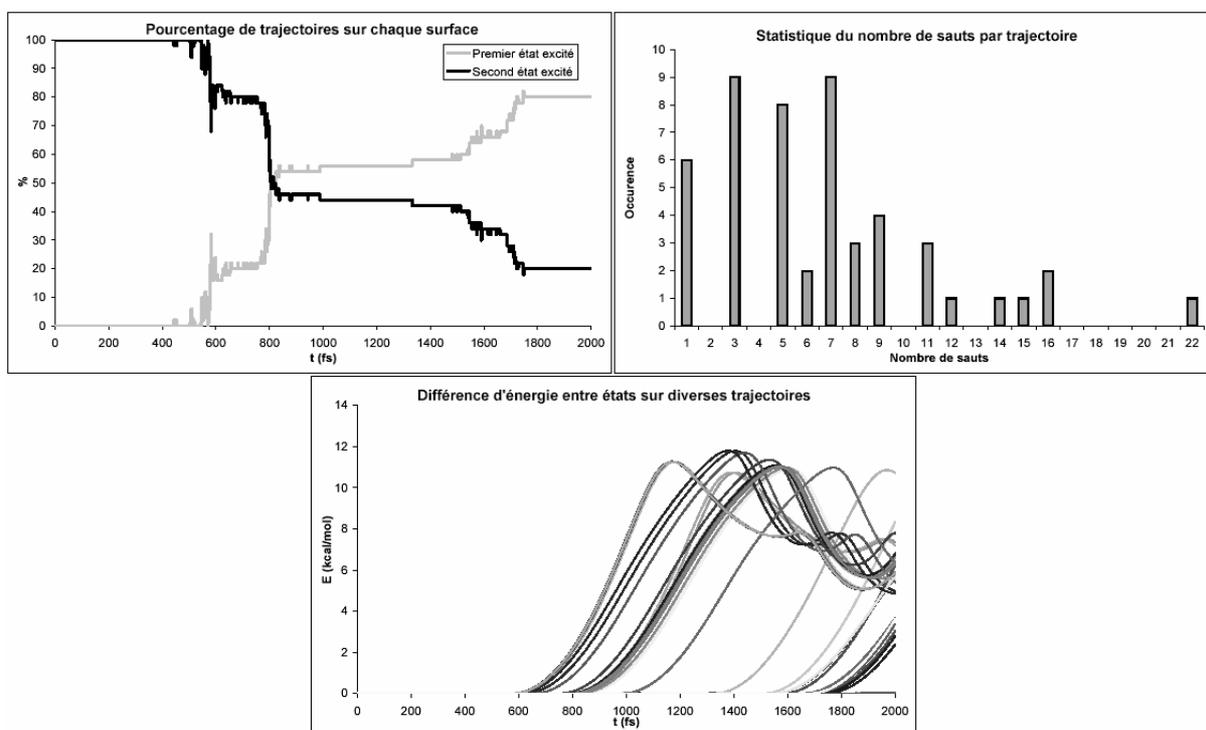


Figure 5.6.

Un ensemble de dynamiques de durées plus importantes (10 ps) permet d'atteindre la relaxation du système (figure 5.7). Au bout de ce temps, toutes les trajectoires évoluent sur la surface d'énergie potentielle du premier état excité électronique de l'agrégat. Pour autant, la densité de cet état n'est pas de 1 mais, en moyenne, d'environ 0,5. Les différences d'énergie entre les deux états excités en fin de dynamique sont conséquentes : supérieures à 3,6

kcal/mol dans tous les cas, elles sont en moyenne d'environ 7,5 kcal/mol et peuvent dépasser 14 kcal/mol. Un retour vers le second état excité est donc fort peu probable. L'extension de la simulation à des durées plus élevées permettrait sans doute d'observer la convergence de la densité de l'état 1 vers 1,0. Il est donc possible d'affirmer que le temps de désexcitation du dimère de naphthalène de son second vers son premier état excité est d'environ 8 ps (ce qu'une statistique portant sur plus de 50 trajectoires permettrait sans doute de confirmer).

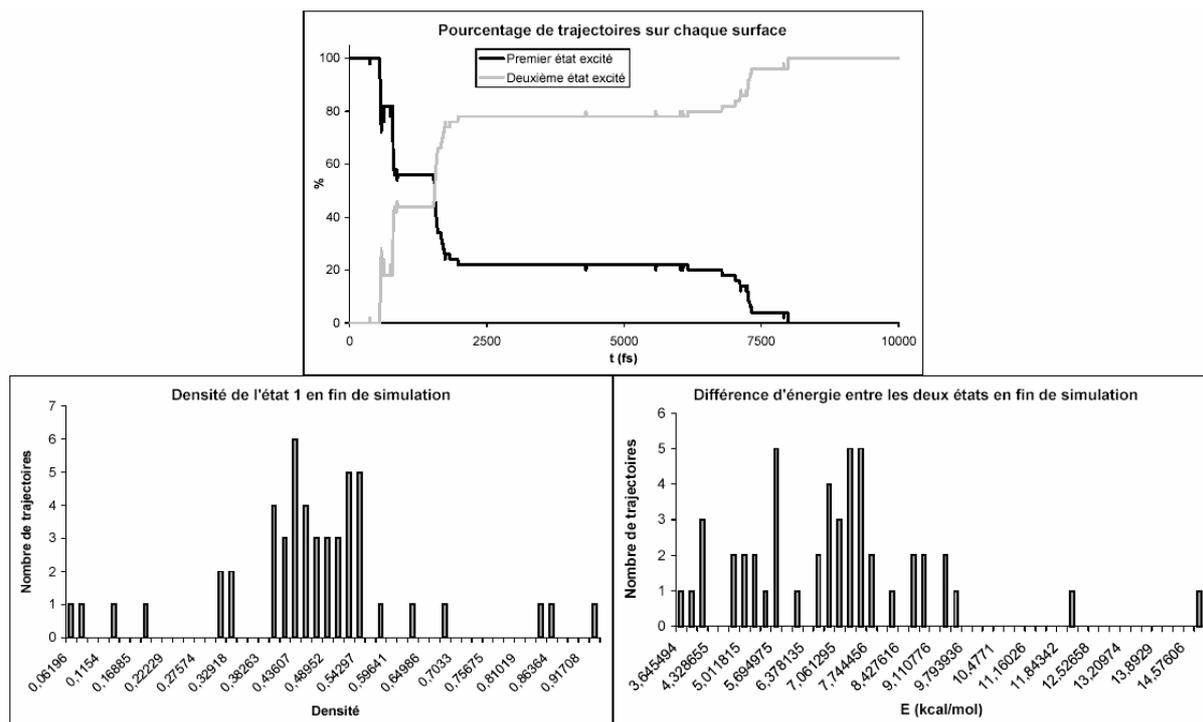


Figure 5.7.

Jusqu'à présent, les états excités du dimère de naphthalène ont été supposés bâtis sur l'état excité  $S_2$  des monomères. Il est également possible de simuler, dans les mêmes conditions, la désexcitation d'un agrégat dont les états propres seraient construits sur leur état  $S_1$ . Néanmoins, au cours des simulations réalisées, aucun saut du second vers le premier état excité du supersystème n'a pu être constaté. Ceci s'explique par la faiblesse des couplages dipolaires entre les deux formes diabatiques, les charges associées à la transition  $S_0 \rightarrow S_1$  étant en moyenne inférieures de deux ordres de grandeur à celles de la transition  $S_0 \rightarrow S_2$  (voir figure 5.4). Il existe donc, à une température moyenne de 10 K, un bassin d'attraction sur la surface du second état excité dans lequel le système reste confiné. Celui-ci, comme l'état fondamental de l'agrégat naphthalène<sub>2</sub>, a une structure en T caractérisée par un couplage très faible (addition

## 5. Dynamique Semi-Classique du Transfert de l'Excitation

de termes égaux et opposés) ; les déviations à cette structure rencontrées au cours de la dynamique ne suffisent pas à promouvoir le saut, étant donné le caractère quasiment interdit de la transition des monomères. Il faut d'ailleurs mettre en exergue le fait que les limites du modèle excitonique sont ici clairement atteintes : les termes de couplage à courte portée ne sont sans doute plus négligeables devant le couplage dipolaire.

Au contraire, le second état excité collectif de l'agrégat naphthalène<sub>2</sub> construit sur l'état S<sub>2</sub> des monomères à une température de 10 K présente des configurations en V et des couplages conséquents (la transition correspondante des monomères étant très permise) qui favorisent les sauts entre surfaces d'énergie potentielle, donc la désexcitation non radiative. La suite de cette étude se restreindra donc à de tels états excités.

### 5.4. Désexcitation non radiative d'agrégats colonnaires de naphthalène

L'extension de la méthode TSH à de plus gros agrégats de naphthalène excités constitue la suite logique de l'étude préliminaire précédente. Un certain nombre de problèmes apparaissent dès lors. Si les énergies d'interaction dans les agrégats aromatiques sont généralement faibles, celles des molécules périphériques d'un gros système le sont encore davantage. De tels agrégats présentent une forte propension à l'évaporation de monomères, favorisée de surcroît par la conversion de l'énergie électronique en énergie cinétique au cours du processus de désexcitation. De plus, les surfaces de potentiel des agrégats aromatiques de nombre d'agrégation supérieur ou égal à quatre sont caractérisées par un grand nombre de minima locaux quasiment isoénergétiques (voir à ce sujet l'article en deuxième partie, paragraphe 2.2), rendant le choix des conditions initiales complexe.

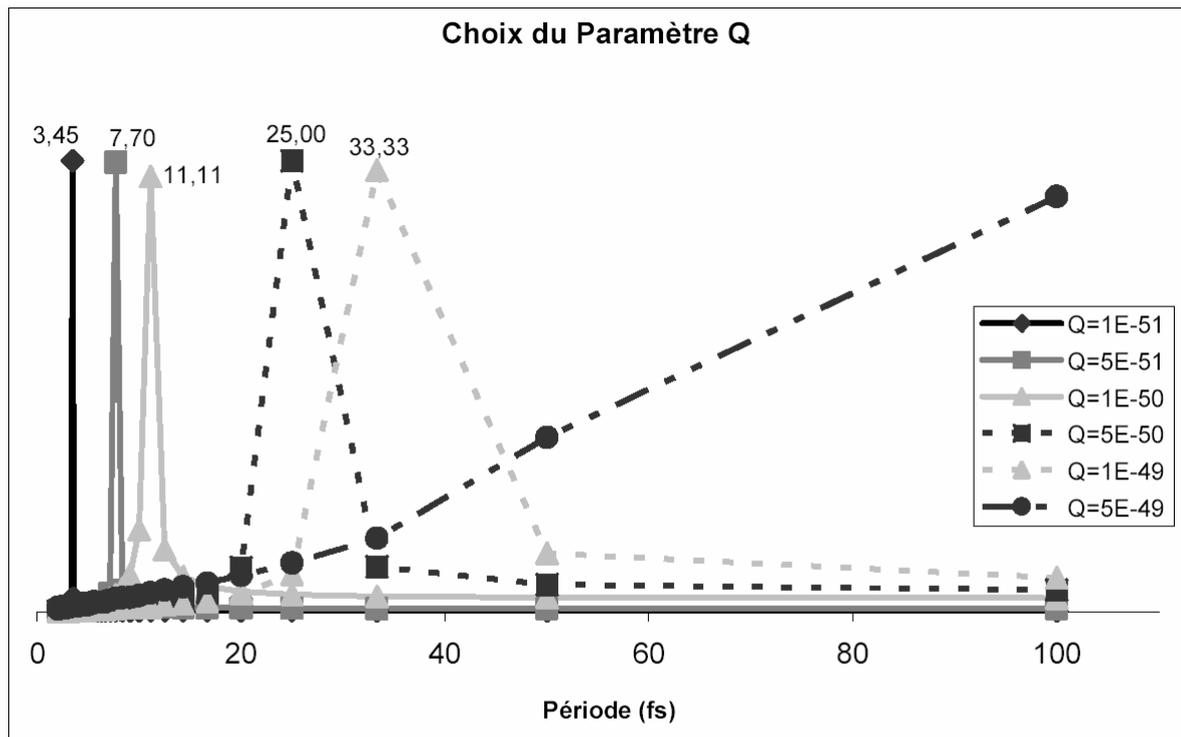
Il a donc été choisi de restreindre l'étude à des agrégats colonnaires de naphthalène, c'est-à-dire dont les centres de gravité des monomères se situent sur un axe. L'occurrence de tels phénomènes de *stacking*, fréquente dans les polymères naturels ou artificiels (cristaux liquides par exemple), justifie ce choix. L'emploi d'une contrainte de Gauss permet d'astreindre les centres de gravité des molécules à l'axe au cours de la dynamique. Un thermostat de Nosé-Hoover assure une dissipation réaliste de l'énergie cinétique générée lors de la désexcitation non radiative de l'agrégat, prévenant ainsi l'évaporation des monomères.

#### *5.4.1. Paramétrisation du thermostat de Nosé-Hoover*

Comme expliqué précédemment, le thermostat de Nosé-Hoover est responsable de l'apparition d'oscillations sans sens physique, de période  $t_0 = 2\pi\sqrt{Q/2gk_B T}$  où  $g$  est le nombre de degrés de liberté du système,  $T$  la température imposée et  $Q$  un paramètre qui règle l'efficacité du couplage avec le bain thermique. La recherche d'une valeur raisonnable pour ce paramètre a été conduite sur le système (naphthalène)<sub>5</sub> empilé dans son état électronique fondamental.

Comme spécifié en première partie de ce mémoire, la période des oscillations précédentes doit être suffisamment petite pour qu'un grand nombre de ces dernières soient inclus dans la simulation, et assez grande pour être supérieure aux temps de corrélation typiques du système à l'étude. La figure 5.8 présente le spectre en période de la température de translation de l'agrégat (naphthalène)<sub>5</sub>, obtenu par analyse de Fourier<sup>268</sup> de cette température au cours d'une dynamique de 0,5 ps. La période des oscillations constatées est en bon accord avec sa prévision théorique, d'autant plus longue que  $Q$  est grand. Des périodes de quelques dizaines de femtosecondes sont très acceptables au vu de la durée des simulations envisagées (quelques picosecondes). Si la durée de vie typique des fonctions de corrélation dynamiques n'est pas connue dans les système à l'étude, elle ne dépasse sans doute pas quelques femtosecondes ; de plus, l'importance de ce critère est à relativiser puisque l'étude de phénomènes de corrélation n'est pas envisagée ici.

## 5. Dynamique Semi-Classique du Transfert de l'Excitation



Q (J.s <sup>2</sup> )	1.10 <sup>-51</sup>	5.10 <sup>-51</sup>	1.10 <sup>-50</sup>	5.10 <sup>-50</sup>	1.10 <sup>-49</sup>	5.10 <sup>-49</sup>
t <sub>0</sub> (fs)	3,45	7,72	10,91	24,40	34,51	77,17

Figure 5.8

Le paramètre  $Q$  gouverne en outre la 'raideur' du couplage avec le bain thermique, le système étant d'autant moins contraint que la valeur de  $Q$  est grande. Comme le montre la figure 5.9, un couplage efficace avec le bain thermique (obtenu aux faibles valeurs de  $Q$ ) permet une meilleure conservation de la température de référence au cours de la simulation. Une valeur de  $Q$  de  $1.10^{-50} \text{ J.s}^2$  a donc été retenue comme meilleur compromis entre ces divers aspects, pour les mouvements de rotation et de translation. De plus, il a été choisi de conserver cette valeur pour l'ensemble des agrégats empilés étudiés, quelle que soit leur taille.

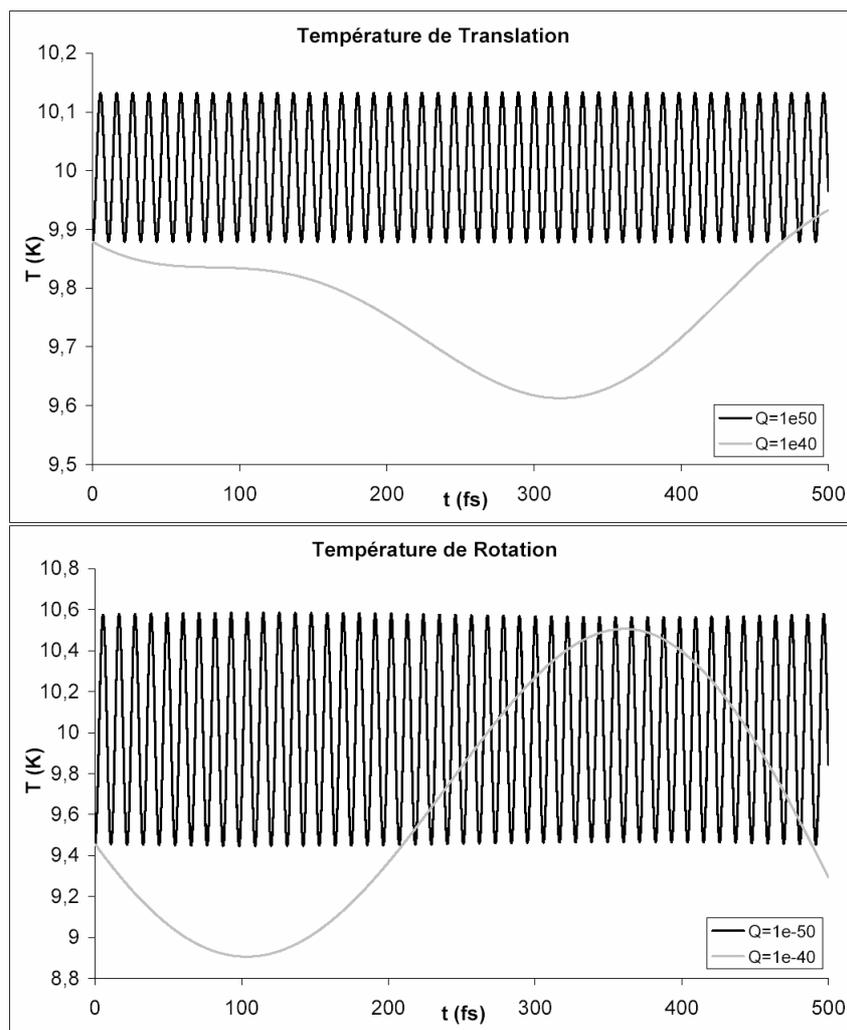


Figure 5.9

#### 5.4.2. Décamère colonnaire de naphthalène

Un agrégat colonnaire de naphthalène, comportant dix monomères, est équilibré à une température de 10 K dans son état fondamental, puis est soumis à une excitation verticale vers son dixième état excité. La désexcitation de cet agrégat vers son état électronique excité de plus basse énergie est alors suivie, le système étant maintenu à 10 K à l'aide d'un thermostat de Nosé-Hoover. La topologie du vecteur propre correspondant à la surface d'énergie potentielle en cours fournit un aperçu du degré de délocalisation spatiale de l'excitation. Il peut être relié au *participation ratio* PR (nombre de monomères concernés par l'excitation), dont l'expression est :

$$PR_k = \left( \sum_m (C_k^m)^2 \right)^{-1}$$

## 5. Dynamique Semi-Classique du Transfert de l'Excitation

où  $k$  est l'index de l'état excité du supersystème,  $m$  représente un monomère de l'agrégat et  $C_k^m$  le coefficient de l'état diabatique où  $m$  est excité dans le vecteur propre de l'état  $k$ . La figure 5.10 représente ainsi le nombre de monomères concernés par cette délocalisation, de même que l'index de l'état excité actuel, au cours d'une simulation typique pour ce système.

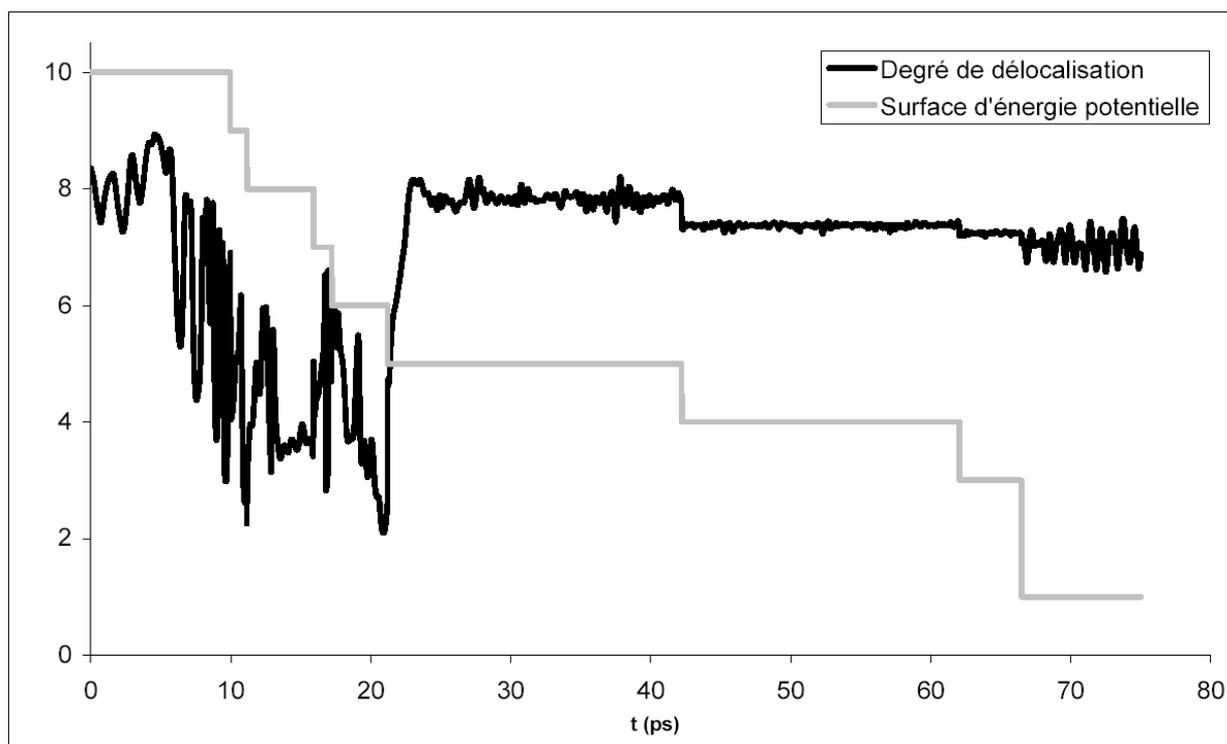


Figure 5.10

Le saut entre surface, c'est-à-dire la modification instantanée de la fonction d'onde électronique du système, s'accompagne généralement d'une modification brutale du degré de délocalisation de l'excitation. En revanche, celle-ci peut être de faible amplitude, comme c'est le cas entre les états excités d'index inférieur ou égal à cinq qui impliquent tous environ huit molécules. Il est à noter que le degré de délocalisation est très sensible aux variations structurales de l'empilement au sein d'un même état électronique (voir par exemple environ 7 ps après le début de la dynamique). En effet, lorsque deux monomères forment une configuration en T, la symétrie des distributions de charges de transition annule le couplage entre les formes diabatiques pour lesquelles l'excitation est localisée sur l'un et l'autre des

deux molécules. Or, la configuration la plus favorable de l'état électronique fondamental d'un agrégat de naphthalène dans lequel les centres de gravité des monomères sont astreints à un axe est un empilement de dimères en T. En conséquence, les états électroniques d'un tel agrégat sont localisés sur les monomères pairs ou impairs, et non sur les deux à la fois. La figure 5.11 le met en évidence sur les deux premiers états excités du décamère de naphthalène empilé, dans la géométrie optimale de son état fondamental. Il en résulte que le facteur de délocalisation desdits états est toujours inférieur à cinq, donc bien moindre que ceux obtenus par la simulation de TSH (Figure 5.10).

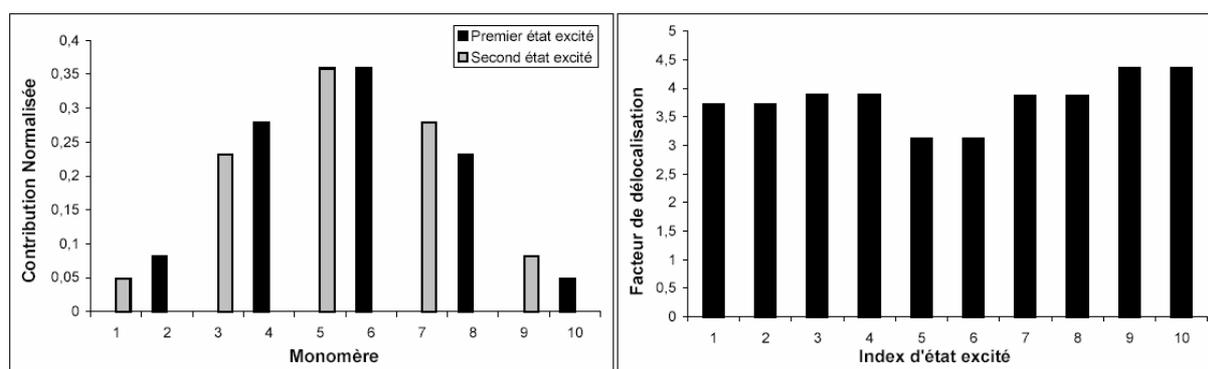


Figure 5.11

Tout écart à la structure en T, au contraire, lève la symétrie et autorise le couplage. Ainsi, l'inclusion du désordre structural par la simulation de TSH favorise la délocalisation de l'excitation. Comme dans le cas de l'ADN (paragraphe 4), le rôle du désordre non diagonal se révèle donc ici prépondérant.

### 5.5. Conclusion

L'application de la méthode TSH au transfert de l'excitation dans de petits agrégats de naphthalène, décrite ci-dessus, avait pour buts principaux le test de l'implémentation de l'algorithme ainsi qu'une compréhension plus 'appliquée' des paramètres qui y interviennent. Un certain nombre d'études complémentaires sur ces systèmes modèle pourraient d'ailleurs compléter ce tableau : influence de la température ou de la taille de l'agrégat sur la durée de vie de l'excitation, par exemple...

L'étape suivante est la mise en œuvre de la méthode sur des systèmes de tailles plus conséquentes, notamment en vue de la comparaison avec les résultats expérimentaux de transfert de l'énergie d'excitation dans les gros agrégats ou les cristaux moléculaires aromatiques<sup>269-271</sup>. Cependant, il a été constaté au cours de cette étude préliminaire que, même dans le cas de petits systèmes, un échantillonnage exhaustif des conditions initiales de simulation était très coûteux. Or, son évolution avec l'augmentation du nombre de degrés de liberté classiques et quantiques qui accompagne l'accroissement de la dimension d'un système est sans doute très défavorable. Un gain préalable en temps de calcul est donc nécessaire. Trois voies sont envisageables à cet égard. Le modèle de dispersion-répulsion, très général et assez complexe, pourrait être simplifié et reparamétré spécifiquement pour les systèmes à l'étude. Le couplage dipolaire pourrait être décrit par le biais de dipôles de transition étendus (paragraphe 2.1.3) plutôt qu'à l'aide de charges atomiques de transition, étant donnée la symétrie de la molécule de naphthalène. Enfin, l'adaptation du programme aux machines massivement parallèles, fondée sur l'affectation à chaque microprocesseur d'une trajectoire indépendante, se révélerait également très bénéfique.



## 6. Conclusion

Comme dans le cas du transfert de charge, la présente partie a montré qu'à l'aide de modèles bien adaptés aux systèmes à l'étude, il était possible de rendre compte de manière très correcte de phénomènes pourtant complexes dans des systèmes de grande taille. L'inclusion dans le modèle du désordre structural, et plus particulièrement de son influence sur les éléments non diagonaux de l'hamiltonien effectif, s'est révélé indispensable pour atteindre une caractérisation correcte des états excités des agrégats étudiés. Ce couplage entre géométrie et fonction d'onde électronique introduit naturellement le concept de dynamique semiclassique, mis en application dans cette étude par le biais de la méthode TSH.

C'est d'ailleurs sur cette méthode que devrait se focaliser toute extension future de la présente étude. Si les outils correspondants ont été mis en place et les premières applications semblent prometteuse, la méthode n'en reste pas moins riche et complexe, et laisse augurer d'intéressantes applications. Son coût en temps de calcul, essentiellement dû à la statistique qu'il est nécessaire de réaliser, la destine plutôt, dans l'immédiat, à des agrégats de petite taille. Néanmoins, la richesse des informations fournies, et le fait que les propriétés de petits systèmes sont souvent extrapolables à des objets de taille plus conséquente, lui confèrent un indéniable attrait. Au niveau pratique, toute amélioration ou extension du logiciel conçu durant cette étude devrait en respecter la structure modulaire, dont la flexibilité est garante d'un large spectre d'applications.

L'extension du modèle employé pour le calcul des éléments de la matrice excitonique semble également un domaine de recherche fructueux ; il serait notamment intéressant de pouvoir traiter des états excités triplets, couplés par des termes à courte portée. En tout état de cause, la simplicité du modèle devrait autant que possible être conservée s'il est envisagé de recourir à des simulations de dynamique classique ou semi-classique, qui peuvent se révéler coûteuses.

Enfin, l'investigation des diverses méthodes semiclassiques de la classe du TSH se révélerait sans doute très intéressante. De telles méthodes connaissent à l'heure actuelle une évolution rapide et sont de plus en plus employées en chimie computationnelle. La connaissance de leurs divers avantages et inconvénients permettrait d'effectuer un choix raisonné, en fonction du système et des propriétés à l'étude, en vue d'applications futures.



# CONCLUSION GÉNÉRALE



*Rien n'est plus facile pour un soldat que de commettre un crime de guerre, pour un bon joueur de se mettre à tricher, à un homme cultivé de redevenir une bête sauvage.*

Maurice G. Dantec, *Babylon Babies*



L'étude présentée dans ce mémoire a permis, par le biais de développements méthodologiques, d'améliorer la description des phénomènes de transfert de charge et d'excitation, de modélisation pourtant notoirement délicate. Cet objectif a pu être atteint par la conjonction de deux axes de recherche.

Tout d'abord, le développement et l'exploitation de potentiels adaptés aux systèmes à l'étude a contribué à la diminution des coûts en temps de calcul, tout en maintenant un bon niveau de qualité. Le rapport des coûts respectifs de la méthode CCSD-T, nécessaire à une description *ab initio* correcte des agrégats de benzène, et du potentiel modèle employé ici pour ces mêmes espèces en est un exemple suffisamment évocateur. Il convient également de noter que ce même potentiel, employé pour décrire les structures *valence bond* des agrégats aromatiques chargés, fournit de très bons résultats dans un domaine caractérisé par une absence quasi-totale de calculs *ab initio* convaincants. Ce potentiel a ensuite été simplifié pour l'étude du transfert d'excitation dans les agrégats de naphthalène par la méthode TSH, en y incluant uniquement les termes pertinents. L'étude de l'influence des effets de désordre structural sur les états excités de fragments d'ADN a rendu nécessaire l'élaboration d'un potentiel intermoléculaire adéquat ; l'extension du paramétrage du champ de forces AMBER aux états excités des bases de l'ADN *via* des calculs *ab initio* a permis d'atteindre cet objectif, tout en conservant la robustesse qui caractérise ce champ de forces dans la série des acides nucléiques.

En second lieu, la conjonction de divers niveaux de calcul (*ab initio* et semi-empiriques), de méthodes d'exploration de surfaces et de dynamique classique ou semi-classique, et de résultats expérimentaux a rendu possible, lors de la constitution de l'hamiltonien effectif, l'augmentation progressive de la complexité de la description des systèmes. Il en va notamment ainsi de l'inclusion graduelle des aspects dynamiques des phénomènes de transfert, depuis le modèle statique ou l'exploration des surfaces d'énergie potentielle à 0 K, jusqu'au TSH.

D'intéressantes tendances ont ainsi pu être mises en lumière. Par exemple, si le désordre structural joue un rôle important aussi bien en transfert de charge qu'en transfert d'excitation, il ne s'y manifeste pas de la même manière. Le transfert de charge, caractérisé

## Conclusion Générale

par des couplages forts et à courte portée, est particulièrement susceptible au désordre diagonal, qui explique l'influence dominante des phénomènes de polarisation dès que les tailles des systèmes augmentent. Au contraire, le transfert d'excitation procède de couplages dipolaires assez faibles et à longue portée ; c'est le désordre non diagonal qui fait montre, dans ce cas, de l'influence la plus marquée sur les propriétés des états électroniques des systèmes.

Un certain nombre de limitations et de restrictions nuisent cependant à la généralité des modèles présentés. La description du couplage entre formes *valence bond* à charge localisée par le produit d'une intégrale de recouvrement et d'un paramètre moléculaire a clairement montré ses limites. Tout d'abord, la relation linéaire supposée dans ce travail n'est plus vérifiée aux faibles recouvrements, révélant l'existence d'un terme négligeable aux forts recouvrements et non pris en compte dans le modèle. Il pourrait par exemple s'agir d'une délocalisation du 'trou' sur une orbitale moléculaire du monomère chargé autre que celle considérée dans le modèle, avec un poids statistique faible mais qui donnerait lieu à des recouvrements plus importants avec l'orbitale moléculaire *ad hoc* du deuxième monomère. En second lieu, l'affectation d'une unique constante de proportionnalité à un système n'est plus justifiée si plusieurs types d'interactions à transfert de charge de natures différentes y coexistent (pyridine<sub>2</sub> ou agrégats hétéromoléculaires). Il conviendrait donc de mieux cerner la dépendance du coefficient de proportionnalité, notamment vis-à-vis du couple donneur-accepteur et de son environnement. D'autres modèles de transfert de charge, plus spécifiques mais plus simples, peuvent également être envisagés, comme celui mis au point au laboratoire dans le système UO<sub>2</sub><sup>2+</sup> / H<sub>2</sub>O<sup>272</sup>. L'application du modèle de transfert de charge à des systèmes plus complexes (fragments d'ADN par exemple) nécessiterait la formulation et l'implémentation d'un modèle de polarisation plus flexible que celui employé dans cette étude (un modèle de type Bader, par exemple).

En ce qui concerne le transfert d'excitation, il semble indispensable d'inclure la modélisation des termes de transfert à courte distance, afin de pouvoir traiter correctement les transitions à l'état triplet, de même que celles qui sont peu permises. Par analogie avec le transfert de charge, ces termes pourraient être fonction (non obligatoirement linéaire) d'une intégrale de recouvrement entre orbitales moléculaires des monomères et de paramètres tirés de calculs *ab initio* d'états excités. Une modélisation *ab initio* rigoureuse du couplage entre formes non adiabatiques à excitation localisée nécessiterait sans doute l'emploi de calculs d'interaction de configurations en base non orthogonale<sup>273-275</sup>. Une autre extension digne

d'intérêt est de tenter de dépasser le modèle de l'exciton de Frenkel afin de rendre compte des états excités à transfert de charge intermoléculaire, dont le rôle n'est souvent pas négligeable.

L'introduction des aspects dynamiques des phénomènes de transfert se heurte au cadre conceptuel des molécules rigides. Ainsi, lors de la mise en œuvre de la méthode TSH, il a du être supposé que l'énergie électronique ne pouvait être évacuée vers les degrés de liberté intramoléculaires. De même, au cours de l'étude des mécanismes d'ionisation dans les agrégats aromatiques, la contribution des agrégats à modes de vibration intramoléculaires excités aux courbes de courant d'ions n'a pu être pris en compte. Pourtant, les systèmes étudiés ici s'accrochent assez bien d'une telle approximation, ce qui est loin d'être général... D'autres thématiques du laboratoire (complexation sélective de cations par des *n*-pyridines ou des calixarènes par exemple) bénéficieraient également de l'inclusion de quelques degrés de flexibilité intramoléculaires. Il faudrait pour cela se doter d'un potentiel intramoléculaire, et réenvisager certains termes du potentiel intermoléculaire dans cette optique. La mise au point d'un modèle compatible avec la rupture de liaisons chimiques, certes plus complexe, serait amplement rentabilisée par les études de dissociation de molécules, sous l'effet de l'ionisation ou de l'excitation, qu'elle rendrait possible par la méthode TSH (particulièrement adaptée à de tels systèmes de petite taille).

Les perspectives d'extension de ce travail sont donc nombreuses et, méthodologiquement parlant, riches et intéressantes. Il peut d'ailleurs s'agir de méthodes générales applicables à d'autres problématiques que celle des phénomènes de transfert, ce qui constitue une motivation supplémentaire pour l'exploration de ces pistes de recherche.



RÉFÉRENCES  
BIBLIOGRAPHIQUES

## Références Bibliographiques

## Références Bibliographiques

- (1) Orrit, M. *Single Mol.* **2002**, 3, 255.
- (2) Orrit, M. *J. Chem. Phys.* **2002**, 117, 10938.
- (3) Condon, E. U. *Phys. Rev.* **1928**, 32, 858.
- (4) Franck, J. *Trans. Faraday Soc.* **1925**, 21, 536.
- (5) Born, M.; Huang, K. *Dynamical Theory of Crystal Lattices*; Clarendon: Oxford, 1954.
- (6) Born, M.; Oppenheimer, R. *Ann. Phys.* **1927**, 84, 457.
- (7) Hund, F. *Zeit. Phys.* **1927**, 43, 805.
- (8) Stückelberg, E. C. G. *Helv. Chim. Acta* **1932**, 5, 369.
- (9) Zener, C. *Proc. R. Soc. A* **1932**, 137, 696.
- (10) Landau, L. D. *Phys. Z. Sowetun.* **1932**, 1, 88.
- (11) Gurney, R. W. *Proc. R. Soc. A* **1931**, 134, 137.
- (12) Davydov, A. S. *Zh. Eksp. Teor. Fiz.* **1948**, 18, 913.
- (13) Kubo, R. *Phys. Rev.* **1952**, 86, 929.
- (14) Duysens, L. M. N. *Prog. Biophys.* **1964**, 14, 1.
- (15) Libby, W. F. *J. Phys. Chem.* **1952**, 56, 63.
- (16) Marcus, R. A. *J. Chem. Phys.* **1956**, 24, 966.
- (17) Marcus, R. A. *Ann. Rev. Phys. Chem.* **1964**, 15, 155.
- (18) Hush, N. S. *Trans. Faraday Soc.* **1961**, 57, 557.
- (19) Hush, N. S. *J. Chem. Phys.* **1958**, 28, 962.
- (20) Levich, V. G.; Dogonadze, R. R. *Dokl. Akad. Nauk. SSSR, Ser. Fiz. Khim.* **1960**, 133, 1368.
- (21) Levich, V. G.; Dogonadze, R. R. *Collect. Czech. Chem. Commun.* **1961**, 26, 193.
- (22) Levich, V. G. *Adv. Electr. Electr. Eng.* **1966**, 4, 249.
- (23) Siatkowski, R. E.; Carter, F. L. The Chemical Elements of Logic. In *Molecular Electronic Devices*; Carter, F. L., Siatkowski, R. E., Wohltjen, H., Eds.; North-Holland: Amsterdam, 1988; pp 321.
- (24) Stern, A.; Stern, V. Intelligent Processor / Memory Elements and Systems. In *European Patent Bulletin*, 1986; Vol. 1986101.
- (25) Stambler, I. *High Tech.* **1984**, 36.

## Références Bibliographiques

- (26) Przybylski, A. Organic Neuron-like Switching and Memory Elements. In *Molecular Electronic Devices*; Carter, F. L., Siatkowski, R. E., Wohltjen, H., Eds.; North-Holland: Amsterdam, 1988; pp 389.
- (27) Kuhn, H. *Naturwiss.* **1967**, *54*, 429.
- (28) Förster, T. In *Modern Quantum Chemistry*; Sinanoglu, O., Ed.; Academic Press: New York, 1965.
- (29) Bradforth, S. E.; Jiminez, R.; Van Mourik, F.; Van Grondelle, R.; Fleming, G. R. *J. Phys. Chem.* **1995**, *99*, 16179.
- (30) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.
- (31) The International Human Genome Mapping Consortium. *Nature* **2001**, *409*, 934.
- (32) The Celera Genomics Sequencing Team. *Science* **2001**, *291*, 1304.
- (33) Campbell, K. H. S.; McWhir, J.; Ritchie, W. A.; Wilmut, I. *Nature* **1996**, *380*, 64.
- (34) Meier, P.; Finch, A.; Evan, G. *Nature* **2000**, *407*, 796.
- (35) De Laurenzi, V.; Melino, G. *Nature* **2000**, *406*, 135.
- (36) Green, D. R.; Beere, H. M. *Nature* **2000**, *405*, 28.
- (37) Barbic, A.; Zimmer, D. P.; Crothers, D. M. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 2369.
- (38) Thorpe, J. H.; Teixeira, S. C. M.; Gale, B. C.; Cardin, C. J. *Nucleic Acids Res.* **2003**, *31*, 844.
- (39) MacKerell, A. D.; Banavali, N. K. *J. Comput. Chem.* **2000**, *21*, 105.
- (40) Langley, D. R. *J. Biomol. Struct. Dyn.* **1998**, *16*, 487.
- (41) Foloppe, N.; MacKerell, A. D. *Abstr. Pap. Am. Chem. Soc.* **1997**, *214*, 236.
- (42) Flatters, D.; Zakrzewska, K.; Lavery, R. *J. Comput. Chem.* **1997**, *18*, 1043.
- (43) Pinak, M. *Cancer Epidemiol. Biomarkers Prev.* **2002**, *11*, A138.
- (44) Florian, J.; Goodman, M. F.; Warshel, A. *Biopolymers* **2003**, *68*, 286.
- (45) Varnai, P.; Djuranovic, D.; Lavery, R.; Hartmann, B. *Nucleic Acids Res.* **2002**, *30*, 5398.
- (46) Mazur, A. K. *J. Am. Chem. Soc.* **2002**, *124*, 14707.
- (47) Fuxreiter, M.; Luo, M.; Jedlovszky, P.; Simon, I.; Osman, R. *J. Mol. Biol.* **2002**, *323*, 823.
- (48) Bhaumik, S. R.; Chary, K. V. R. *J. Biomol. Struct. Dyn.* **2002**, *20*, 199.
- (49) Kurita, N.; Inoue, H.; Sekino, H. *Chem. Phys. Lett.* **2003**, *370*, 161.
- (50) Kryachko, E. S.; Sabin, J. R. *Int. J. Quantum Chem.* **2003**, *91*, 695.
- (51) Kryachko, E. S.; Nguyen, M. T. *J. Phys. Chem. A* **2002**, *106*, 9319.

- (52) Hobza, P.; Sponer, J. *J. Am. Chem. Soc.* **2002**, *124*, 11802.
- (53) Lien, T.; Ault, K.; Pezeshk, A. *Abstr. Pap. Am. Chem. Soc.* **2002**, *223*, 658.
- (54) Jacobson, J.; Rice, K.; Pezeshk, A. *Abstr. Pap. Am. Chem. Soc.* **2001**, *221*, 336.
- (55) Brocklehurst, B. *Radiat. Res.* **2001**, *155*, 637.
- (56) Pezeshk, A.; Jacobson, M.; Jacobson, J. *Abstr. Pap. Am. Chem. Soc.* **2000**, *219*, 629.
- (57) Razskazovskiy, Y.; Debije, M. G.; Bernhard, W. A. *Radiat. Res.* **2000**, *153*, 436.
- (58) Fielden, E. M.; Michael, B. D.; Prise, K. M. *Radiat. Res.* **1997**, *148*, 481.
- (59) Shikazono, N.; Tanaka, A.; Kitayama, S.; Watanabe, H.; Tano, S. *Radiat. Environ. Biophys.* **2002**, *41*, 159.
- (60) O'Neill, P. The Role of Hydration and Radiation Quality in the Induction of DNA-Damage - Chemical Aspects. In *Life Sciences and Space Research XXV(2)*, 1994; Vol. 14; pp 221.
- (61) Phillips, R. M.; Ward, T. H. *Anticancer Res.* **2001**, *21*, 1795.
- (62) Al-Nabulsi, I.; Wheeler, K. T. *Radiat. Res.* **1999**, *151*, 188.
- (63) Sharon, C. E.; Musich, P. R. *Faseb J.* **2002**, *16*, A527.
- (64) Gilbert, B. C.; Silvester, S.; Walton, P. H.; Whitwood, A. C. *J. Chem. Soc.-Perkin Trans. 2* **1999**, 1891.
- (65) Adam, W.; Cadet, J.; Dallacqua, F.; Epe, B.; Ramaiah, D.; Sahamoller, C. R. *Angew. Chem.-Int. Edit. Engl.* **1995**, *34*, 107.
- (66) Cadet, J.; Bellon, S.; Berger, M.; Bourdat, A. G.; Douki, T.; Duarte, V.; Frelon, S.; Gasparutto, D.; Muller, E.; Ravanat, J. L.; Sauvaigo, S. *Biol. Chem.* **2002**, *383*, 933.
- (67) Cadet, J.; Berger, M.; Douki, T.; Morin, B.; Raoul, S.; Ravanat, J. L.; Spinelli, S. *Biol. Chem.* **1997**, *378*, 1275.
- (68) Ziegler, A.; Jonason, A. S.; Leffell, D. J.; Simon, J. A.; Sharma, H. W.; Kimmelman, J.; Remington, L.; Jacks, T.; Brash, D. E. *Nature* **1994**, *372*, 773.
- (69) Brash, D. E.; Haseltine, W. A. *Nature* **1982**, *298*, 189.
- (70) Eley, D. D.; Spivey, D. I. *Trans. Faraday Soc.* **1962**, *58*, 411.
- (71) Steenken, S. *Chem. Rev.* **1987**, *89*, 503.
- (72) Steenken, S. *Biol. Chem.* **1997**, *378*, 1293.
- (73) Stemp, E. D. A.; Arkin, M. R.; Barton, J. K. *J. Am. Chem. Soc.* **1997**, *119*, 2921.
- (74) Hall, D. B.; Barton, J. K. *J. Am. Chem. Soc.* **1997**, *119*, 5045.
- (75) Breslin, D. T.; Caury, J. E.; Anderson, J. R.; McFail-Isom, L.; Kan, Y.; Williams, L. D.; Bottomley, L. A.; Schuster, G. B. *J. Am. Chem. Soc.* **1997**, *119*, 5043.
- (76) Guallar, V.; Douhal, A.; Moreno, M.; Lluch, J. M. *J. Phys. Chem. A* **1999**, *103*, 6251.

## Références Bibliographiques

- (77) Demple, B.; Harrison, L. *Ann. Rev. Biochem.* **1994**, *63*, 915.
- (78) Loft, S.; Poulsen, H. E. *J. Mol. Med.* **1996**, *74*, 297.
- (79) Schoonveld, W. A.; Vrijmoeth, J.; Klapwijk, T. M. *Appl. Phys. Lett.* **1998**, *73*, 3884.
- (80) Nelson, S. F.; Lin, Y. Y.; Gundlach, D. J.; Jackson, T. N. *Appl. Phys. Lett.* **1998**, *72*, 1854.
- (81) van de Craats, A. M.; Siebbeles, L. D. A.; Bleyl, I.; Haarer, D.; Berlin, Y. A.; Zharikov, A. A.; Warman, J. M. *J. Phys. Chem. B* **1998**, *102*, 9625.
- (82) Berlin, Y. A.; Burin, A. L.; Ratner, M. A. *Superlattices Microstruct.* **2000**, *28*, 241.
- (83) Lisdat, F.; Ge, B.; Scheller, F. W. *Electrochem. Commun.* **1999**, *1*, 65.
- (84) Winfree, E.; Liu, F. R.; Wenzler, L. A.; Seeman, N. C. *Nature* **1998**, *394*, 539.
- (85) Marshall, A.; Hodgson, J. *Nat. Biotechnol.* **1998**, *16*, 27.
- (86) Mirkin, C. A.; Letsinger, R. L.; Mucic, R. C.; Storhoff, J. J. *Nature* **1996**, *382*, 607.
- (87) Lewis, F. D.; Letsinger, R. L.; Wasielewski, M. R. *Acc. Chem. Res.* **2001**, *34*, 159.
- (88) Giese, B. *Acc. Chem. Res.* **2000**, *33*, 631.
- (89) Schuster, G. B. *Accounts Chem. Res.* **2000**, *33*, 253.
- (90) Grinstaff, M. W. *Angew. Chem.-Int. Edit. Engl.* **1999**, *38*, 3629.
- (91) Diederichsen, U. *Angew. Chem.-Int. Edit. Engl.* **1997**, *36*, 2317.
- (92) Jortner, J.; Bixon, M.; Voityuk, A. A.; Rosch, N. *J. Phys. Chem. A* **2002**, *106*, 7599.
- (93) Bixon, M.; Jortner, J. *Chem. Phys.* **2002**, *281*, 393.
- (94) Bixon, M.; Jortner, J. *J. Am. Chem. Soc.* **2001**, *123*, 12556.
- (95) Bixon, M.; Jortner, J. *J. Phys. Chem. A* **2001**, *105*, 10322.
- (96) Voityuk, A. A.; Jortner, J.; Bixon, M.; Rosch, N. *J. Chem. Phys.* **2001**, *114*, 5614.
- (97) Voityuk, A. A.; Rosch, N.; Bixon, M.; Jortner, J. *J. Phys. Chem. B* **2000**, *104*, 9740.
- (98) Bixon, M.; Giese, B.; Wessely, S.; Langenbacher, T.; Michel-Beyerle, M. E.; Jortner, J. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 11713.
- (99) Jortner, J.; Bixon, M.; Langenbacher, T.; Michel-Beyerle, M. E. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 12759.
- (100) Voityuk, A. A.; Siriwong, K.; Rosch, N. *Phys. Chem. Chem. Phys.* **2001**, *3*, 5421.
- (101) Georghiou, S.; Phillips, G. R.; Ge, G. *Biopolymers* **1992**, *32*, 1417.
- (102) Georghiou, S.; Zhu, S.; Weidner, R.; Huang, C.-R.; Ge, G. *J. Biomol. Struct. Dyn.* **1990**, *84*, 289.
- (103) Ge, G.; Georghiou, S. *Photochem. Photobiol.* **1991**, *54*, 301.
- (104) Huang, C.-R.; Georghiou, S. *Photochem. Photobiol.* **1992**, *56*, 95.
- (105) Guéron, M.; Eisinger, J.; Schulman, R. G. *J. Chem. Phys.* **1967**, *47*, 4077.

- (106) Rayner, D. M.; Szabo, A. G.; Loutfy, R. O.; Yip, R. W. *J. Phys. Chem.* **1980**, *84*, 289.
- (107) Miyata, T.; Yomosa, S. *J. Phys. Soc. Jpn.* **1969**, *27*, 720.
- (108) Cadet, J.; Vigny, P. In *Bioorganic Photochemistry*; Morrison, H., Ed.; Wiley: New York, 1990; Vol. 1; pp 35.
- (109) Bensasson, R. V.; Land, E. J.; Truscott, T. G. *Flash Photolysis and Pulse Radiolysis*; Pergamon: Oxford, 1983.
- (110) Lin, N. Y. *Radiat. Phys. Chem.* **1996**, *47*, 209.
- (111) Nikogosan, D. N. *Int. J. Radiat. Biol.* **1990**, *57*, 233.
- (112) Brenner, V. Modélisation des forces intermoléculaires et caractérisation des hypersurfaces d'énergie potentielle. Applications à de petits agrégats de van der Waals, Université Paris XI, UFR scientifique d'Orsay, 1993.
- (113) Derepas, A. L. Polarisation et transfert de charge dans les systèmes cation métallique - molécules : développement de potentiels modèle, Université Paris XI, UFR scientifique d'Orsay, 2001.
- (114) Bertolus, M. Etude de petits agrégats mixtes par la méthode de la fonctionnelle de la densité et par des potentiels modèle, Paris XI, UFR scientifique d'Orsay, 1998.
- (115) Gaigeot, M.-P. Etude théorique par dynamique moléculaire de propriétés de petits agrégats de Van der Waals, Toulouse, Université Paul Sabatier, 1997.
- (116) Germain, A. Localisation des orbitales moléculaires sigma: vers une meilleure caractérisation théorique des états électroniques excités de grandes molécules conjuguées, Université Paris XI, UFR scientifique d'Orsay, 1996.
- (117) Markovitsi, D.; Marguet, S.; Bondkowski, J.; Kumar, S. *J. Phys. Chem. B* **2001**, *105*, 1299.
- (118) Markovitsi, D.; Marguet, S.; Gallos, L. K.; Sigal, H.; Millie, P.; Argyrakis, P.; Ringsdorf, H.; Kumar, S. *Chem. Phys. Lett.* **1999**, *306*, 163.
- (119) Lemaistre, J. P.; Ecoffet, C.; Markovitsi, D. *J. Lumines.* **1995**, *66-7*, 327.
- (120) Rivail, J.-L. *Elements de chimie quantique à l'usage des chimistes*; CNRS Editions: Paris, 1999.
- (121) Evangelisti, S.; Daudey, J. P.; Malrieu, J. P. *Chem. Phys.* **1983**, *75*, 91.
- (122) Huron, B.; Malrieu, J. P.; Rancurel, P. *J. Chem. Phys.* **1973**, *58*, 5745.
- (123) Klopman, G.; Evans, R. C. In *Modern Theoretical Chemistry*; Segal, G. A., Ed.; Plenum Press: New York, 1977; Vol. 7.
- (124) Momicchioli, F.; Baraldi, I.; Bruni, M. C. *Chem. Phys.* **1983**, *82*, 229.
- (125) Pople, J. A.; Beveridge, D. L.; Dobosh, P. A. **1967**, *47*, 2026.

## Références Bibliographiques

- (126) Szczepanski, J.; Vala, M.; Talbi, D.; Parisel, O.; Ellinger, Y. *J. Chem. Phys.* **1993**, *98*, 4494.
- (127) Momicchioli, F.; Baraldi, I.; Berthier, G. *Chem. Phys.* **1988**, *123*, 103.
- (128) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
- (129) Heitler, H.; London, F. *Zeit. Phys.* **1927**, *44*, 455.
- (130) Hiberty, P. C.; Cooper, D. L. *Theochem - J. Mol. Struct.* **1988**, *169*, 437.
- (131) Levine, I. N. *Quantum Chemistry*, 5 ed.; Prentice Hall: Upper Saddle River (NJ), 2000.
- (132) Claverie, P. *Intermolecular Interactions: from Diatomics to Biopolymers*; Wiley: New York, 1978.
- (133) Stone, A. J. *Theory of Intermolecular Forces*; Clarendon Press: Oxford, 1997.
- (134) Vigné-Maeder, F.; Claverie, P. *J. Chem. Phys.* **1988**, *88*, 4934.
- (135) Derepas, A. L.; Soudan, J. M.; Brenner, V.; Millie, P. *J. Comput. Chem.* **2002**, *23*, 1013.
- (136) Feynman; Leighton; Sands. Electromagnétisme. In *Le Cours de Physique de Feynman*; InterÉditions: Paris, 1979; Vol. 2; pp 196.
- (137) Lefevre, R. J. W. *Dipole Moments*, 3 ed.; Methuen: London, 1953.
- (138) Lefèvre, R. J. W. *Advances in Physical Organic Chemistry*; Academic Press: New York, 1965; Vol. 3.
- (139) Mulder, F.; Van Dijk, G.; Huiszoon, C. *Mol. Phys.* **1979**, *38*, 577.
- (140) Kitaigorodski, A. I. *Tetrahedron* **1961**, *14*, 975.
- (141) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.
- (142) Stogryn, D. E. *Mol. Phys.* **1971**, *22*, 81.
- (143) Caillet, J.; Claverie, P. *Acta Cryst.* **1975**, *A31*, 448.
- (144) Langlet, J.; Claverie, P.; Caron, F.; Boeue, C. *Int. J. Quantum Chem.* **1981**, *20*, 299.
- (145) Gresh, N. *J. Comput. Chem.* **1995**, *16*, 856.
- (146) Davydov, A. S. *Theory of Molecular Excitons*; Plenum Press: New York, 1971.
- (147) Engdahl, A.; Nelander, B. *J. Phys. Chem.* **1985**, *89*, 2860.
- (148) Feller, W. *An Introduction to Probability Theory and Its Applications*; Wiley: New York, 1950; Vol. 1.
- (149) Seneta, E. *Non-negative Matrices and Markov Chains, Second Edition*; Springer-Verlag: New York, 1981.
- (150) Heerman, R. *Computer Simulation Methods*; Springer-Verlag: Berlin, 1990.

- (151) Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P. *Science* **1983**, *220*, 671.
- (152) Metropolis, N.; Rosenbluth, A.; Rosenbluth, M.; Teller, A.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.
- (153) Goldberg, D. E. *Genetic Algorithms in Search, Optimization, and Machine Learning*; Addison-Wesley: New York, 1989.
- (154) Holland, J. H. *Adaptation in Natural and Artificial Systems*; The University of Michigan Press: Ann Harbor, 1975.
- (155) Garel, T.; Orland, H.; Velikson, B. *J. Phys. A* **1990**, *23*, L621.
- (156) Garel, T.; Niel, J. C.; Orland, H.; Velikson, B. *J. Chim. Phys.* **1991**, *88*, 2473.
- (157) Shanno, D. F. *Math. Comput.* **1970**, *24*, 647.
- (158) Goldfarb, D. *Math. Comput.* **1970**, *24*, 23.
- (159) Fletcher, R. *Comput. J.* **1970**, *13*, 317.
- (160) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes (FORTRAN Version)*; Cambridge University Press: Cambridge, 1989.
- (161) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes in Fortran 90*; Cambridge University Press: Cambridge, 1996.
- (162) Aarts, E.; Korst, J. *Simulated Annealing and Boltzmann Machines - A Stochastic Approach to Combinatorial Optimization and Neural Computing*; John Wiley & Sons: New York, 1989.
- (163) José, J. V.; Saletan, E. J. *Classical Dynamics: A Contemporary Approach*; Cambridge University Press: Cambridge, 1998.
- (164) Verlet, L. *Phys. Rev.* **1967**, *159*, 98.
- (165) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford Science Publications: Oxford, 1989.
- (166) Evans, D. J. *Mol. Phys.* **1977**, *34*, 317.
- (167) Evans, D. J.; Murad, S. *Mol. Phys.* **1977**, *34*, 327.
- (168) Svanberg, M. *Mol. Phys.* **1997**, *92*, 1085.
- (169) Fincham, D. *Mol. Simul.* **1992**, *8*, 165.
- (170) McQuarrie, D. A. *Statistical Mechanics*; Harper Collins: New York, 1976.
- (171) Berne, B. J.; Boon, J. P.; Rice, S. A. *J. Chem. Phys.* **1966**, *45*, 1086.
- (172) Cooley, J. W.; Lewis, P. A. W. *IEEE Trans. Aud. Electr.* **1969**, *17*, 77.
- (173) Cochran, W. T.; Cooley, J. W. *IEEE Trans. Aud. Electr.* **1967**, *15*, 45.
- (174) Harris, F. J. *Proc. IEEE* **1978**, *66*, 51.

## Références Bibliographiques

- (175) Foulquier, N. "Etude des spectres de vibrations intermoléculaires dans les agrégats de van der Waals par dynamique moléculaire," Université de Paris-Sud, 1997.
- (176) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
- (177) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.
- (178) Nosé, S. *Mol. Phys.* **1984**, *52*, 255.
- (179) Nosé, S. *Molecular Dynamics Simulations at Constant Temperature and Pressure*. In *Computer Simulations in Materials Science*; Meyer, M., Pontikis, V., Eds.; Kluwer Academic Publishers: Amsterdam, 1991; pp 21.
- (180) Evans, D. J.; Morriss, G. P. *Statistical Mechanics of Nonequilibrium Liquids*; Academic Press: London, 1990.
- (181) Tully, J. C. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1991**, *25*, 299.
- (182) Tully, J. C. *J. Chem. Phys.* **1990**, *93*, 1061.
- (183) Tully, J. C.; Preston, R. K. *J. Chem. Phys.* **1971**, *55*, 562.
- (184) Hammes-Schiffer, S.; Tully, J. C. *J. Chem. Phys.* **1994**, *101*, 4657.
- (185) Billing, G. D. In *Encyclopedia of Computational Chemistry*; von Schleyer, P. R., Ed.; Wiley: New York, 1998; pp 1587.
- (186) Drukker, K. *J. Comput. Phys.* **1999**, *153*, 225.
- (187) Jungwirth, P.; Gerber, R. B. *Chem. Rev.* **1999**, *99*, 1583.
- (188) Makri, N. *Annu. Rev. Phys. Chem.* **1999**, *50*, 167.
- (189) Teixidor, L. "Dynamique d'un atome de calcium excité sur un agrégat d'argon," Université Paul Sabatier, 2003.
- (190) Cohen-Tannoudji, C.; Diu, B.; Laloë, F. *Mécanique Quantique*; Hermann: Paris, 1977.
- (191) Jortner, J.; Bixon, M. In *Protein Structure: Molecular and Electronic Reactivity*; Austin, R., Buhks, E., Chance, B., DeVault, D., Dutton, P. L., Frauenfelder, H., Goldanskii, V. I., Eds.; Springer: New York, 1985; pp 277.
- (192) Kalosakas, G.; Rasmussen, K. O.; Bishop, A. R. *J. Chem. Phys.* **2003**, *118*, 3731.
- (193) Castet, F.; Ducasse, L.; Fritsch, A. *Chem. Phys.* **1998**, *232*, 37.
- (194) Castet, F.; Fritsch, A.; Ducasse, L. *J. Phys. I* **1996**, *6*, 583.
- (195) Pople, J. A. *Trans. Faraday Soc.* **1953**, *49*, 1375.
- (196) Pariser, R.; Parr, R. G. *J. Chem. Phys.* **1953**, *21*, 466.
- (197) Brunaud, G.; Castet, F.; Fritsch, A.; Ducasse, L. *Phys. Chem. Chem. Phys.* **2002**, *4*, 6072.

- (198) Brunaud, G.; Castet, F.; Fritsch, A.; Ducasse, L. *Phys. Chem. Chem. Phys.* **2003**, *5*, 2104.
- (199) Brunaud, G.; Castet, F.; Fritsch, A.; Kreissler, M.; Ducasse, L. *J. Phys. Chem. B* **2001**, *105*, 12665.
- (200) Troisi, A.; Orlandi, G. *Chem. Phys. Lett.* **2001**, *344*, 509.
- (201) Troisi, A.; Orlandi, G. *J. Phys. Chem. B* **2002**, *106*, 2093.
- (202) Ladik, J. J.; Ye, Y. *J. Phys. Status Solidi B-Basic Res.* **1998**, *205*, 3.
- (203) Priyadarshy, S.; Risser, S. M.; Beratan, D. N. *J. Phys. Chem.* **1996**, *100*, 17678.
- (204) Voityuk, A. A.; Rosch, N. *J. Phys. Chem. B* **2002**, *106*, 3013.
- (205) Voityuk, A. A.; Rosch, N. *J. Chem. Phys.* **2002**, *117*, 5607.
- (206) Lewis, J. P.; Cheatham, T. E.; Starikov, E. B.; Wang, H.; Sankey, O. F. *J. Phys. Chem. B* **2003**, *107*, 2581.
- (207) Gresh, N.; Pullman, A.; Claverie, P. *Int. J. Quantum Chem.* **1985**, *28*, 757.
- (208) Stone, A. J. *Chem. Phys. Lett.* **1993**, *211*, 101.
- (209) Hinchliffe, A.; Soscùn, H. J. M. *Theochem - J. Mol. Struct.* **1993**, *300*, 1.
- (210) Archibong, E. F.; Thakkar, A. *J. Mol. Phys.* **1994**, *81*, 557.
- (211) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (212) Fraga, S.; Karwowski, J.; Saxena, K. M. S. *Handbook of Atomic Data*; Elsevier: Amsterdam, 1976.
- (213) Ambrose, D.; Baulch, D. L.; Berger, L. I.; Berger, M. J.; Burns, G. W.; Covington, A. K.; Dolezalek, H.; Evans, H. T.; Fokin, L. R.; Frederikse, H. P. R.; Gschneidner, K. A.; Gurvich, L. V.; Hammond, C. R.; Hampson, R. F.; Holden, N. E.; Jenkins, H. D. B.; Kehiaian, H. V.; Kerr, J. A.; Krupke, W. F.; Kushitsu, K.; Latham, J.; Lias, S. G.; Luckenbach, R.; Maki, A.; Mangum, B. W.; Miller, T. M.; Perry, D. L.; Phillips, S. L.; Rao, C. N. R.; Raychaudhuri, A. K.; Tammet, H.; Taylor, B. N.; Burns, D. T.; Trippe, T. G.; Uman, M. A.; Vanysek, P.; Weaver, J. H.; Wexler, A.; Wiese, W. L.; Williams, G. P.; Wohlfarth, C. *CRC Handbook of Chemistry and Physics*; CRC Press: Boca Raton, 1996.
- (214) Pacios, L. F.; Christiansen, P. A. *J. Chem. Phys.* **1985**, *82*, 2664.
- (215) Dehez, F.; Chipot, C.; Millot, C.; Angyan, J. G. *Chem. Phys. Lett.* **2001**, *338*, 180.
- (216) Bader, R. F. W. *Atoms in Molecules - A Quantum Theory*; Oxford University Press: London, 1990.
- (217) Bouvier, B.; Brenner, V.; Millie, P.; Soudan, J. M. *J. Phys. Chem. A* **2002**, *106*, 10326.
- (218) Hobza, P.; Selze, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 18790.
- (219) McWeeny, R.; del Re, G. *Theor. Chim. Acta* **1968**, *10*, 13.

## Références Bibliographiques

- (220) Gilbert, T. L. *Molecular Orbitals in Chemistry, Physics and Biology*; Academic Press: London, 1964.
- (221) Dupuis, M.; Marquez, A.; Davidson, E. R. HONDO 95.3 from CHEM-Station; IBM Corporation: Kingston, NY, 1995.
- (222) Kennard, O.; Hunter, W. N. In *Nucleic Acids, Landolt-Börnstein New Series vol. VII/1a*; Saenger, W., Ed.; Springer: Berlin, 1989; pp 1.
- (223) Olson, W. K.; Bansal, M.; Burley, S. K.; Dickerson, R. E.; Gerstein, M.; Harvey, S. C.; Heinemann, U.; Lu, X. J.; Neidle, S.; Shakked, Z.; Sklenar, H.; Suzuki, M.; Tung, C. S.; Westhof, E.; Wolberger, C.; Berman, H. M. *J. Mol. Biol.* **2001**, *313*, 229.
- (224) Kennard, O.; Hunter, W. N. In *Nucleic Acids, Landolt-Börnstein New Series vol. VII/1a*; Saenger, W., Ed.; Springer: Berlin, 1989; pp 255.
- (225) Arnott, S.; Selsing, E. *J. Mol. Biol.* **1974**, *88*.
- (226) Viswamitra, M. A.; Shakked, Z.; Jones, P. G.; Sheldrick, G. M.; Salisbury, S. A.; Kennard, O. *Biopolymers* **1982**, *21*, 513.
- (227) Norinder, U. *Theochem-J. Mol. Struct.* **1987**, *151*, 259.
- (228) Katritzky, A. R.; Karelson, M. *J. Am. Chem. Soc.* **1991**, *113*, 1561.
- (229) Person, W. B.; Szczepaniak, K.; Szczesniak, M.; Kwiatkowski, J. S.; Hernandez, L.; Czerminsky, R. *Theochem-J. Mol. Struct.* **1989**, *194*, 239.
- (230) Leao, M. B. C.; Longo, R. L.; Pavao, A. C. *Theochem-J. Mol. Struct.* **1999**, *490*, 145.
- (231) Marian, C. M.; Schneider, F.; Kleinschmidt, M.; Tatchen, J. *Eur. Phys. J. D* **2002**, *20*, 357.
- (232) Poulter, C. D.; Frederick, G. D. *Tetrahedron Lett.* **1975**, *26*, 2171.
- (233) Szczesniak, M.; Nowak, M. J.; Rostkowska, H.; Szczepaniak, K.; Person, W. B.; Shugar, D. *J. Am. Chem. Soc.* **1983**, *106*, 5969.
- (234) Holmen, A.; Broo, A. *Int. J. Quantum Chem.* **1995**, 113.
- (235) Salter, L. M.; Chaban, G. M. *J. Phys. Chem. A* **2002**, *106*, 4251.
- (236) Dreyfus, M.; Dodin, G.; Bensaude, O.; Dubois, J. E. *J. Am. Chem. Soc.* **1975**, *97*, 2369.
- (237) Wolfenden, R. V. *J. Mol. Biol.* **1969**, *40*, 307.
- (238) Rashbah, E. I.; Sturge, M. D. *Excitons*; North-Holland: Amsterdam, 1982.
- (239) Frenkel, J. *Phys. Rev.* **1931**, *37*, 1276.
- (240) Agranovich, V. M.; Galanin, M. D. *Electronic Excitation Energy Transfer in Condensed Matter*. In *Modern Problems in Condensed Matter Sciences*; North-Holland: Amsterdam, 1982.

- (241) May, V.; Kühn, O. *Charge and Energy Transfer Dynamics in Molecular Systems - A Theoretical Introduction*; Wiley-VCH: Berlin, 2000.
- (242) Germain, A. DEA de Physicochimie Moléculaire, 1993.
- (243) Czikkleky, V.; Fosterling, H. D.; Kuhn, H. *Chem. Phys. Lett.* **1970**, *6*, 207.
- (244) Longuet-Higgins, H. C. *Proc. R. Soc. A* **1956**, *235*, 537.
- (245) Callis, P. R. *Ann. Rev. Phys. Chem.* **1983**, *34*, 329.
- (246) Volosov, A.; Woody, R. W. *J. Chem. Phys.* **1992**, *96*, 4845.
- (247) Novros, J. S.; Clark, L. B. *J. Phys. Chem.* **1986**, *90*, 5666.
- (248) Miles, D. W.; Robins, M. J.; Robins, R. K.; Winkley, M. W.; Eyring, H. *J. Am. Chem. Soc.* **1969**, *91*, 825.
- (249) Miles, D. W.; Robins, M. J.; Robins, R. K.; Winkley, M. W.; Eyring, H. *J. Am. Chem. Soc.* **1969**, *91*, 831.
- (250) Volosov, A. *Int. J. Quantum Chem.* **1989**, *36*, 473.
- (251) Germain, A.; Millie, P. *Chem. Phys.* **1997**, *219*, 265.
- (252) Marguet, S.; Germain, A.; Millie, P. *Chem. Phys.* **1996**, *208*, 351.
- (253) Despres, A.; Lejeune, V.; Migirdicyan, E.; Admasu, A.; Platz, M. S.; Berthier, G.; Parisel, O.; Flament, J. P.; Baraldi, I.; Momicchioli, F. *J. Phys. Chem.* **1993**, *97*, 13358.
- (254) Marguet, S.; Mialocq, J. C.; Millie, P.; Berthier, G.; Momicchioli, F. *Chem. Phys.* **1992**, *160*, 265.
- (255) Wang, J. C. P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
- (256) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Metz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weimer, P. K.; Kollman, P. A. AMBER 6; University of California: San Francisco, 1999.
- (257) Levy, B.; Enescu, M. *Theochem-J. Mol. Struct.* **1998**, *432*, 235.
- (258) Francl, M. M.; Carey, C.; Chirlian, L. E.; Gange, D. M. *J. Comput. Chem.* **1996**, *17*, 367.
- (259) Bayly, C. L.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.
- (260) Spackman, M. A. *J. Comput. Chem.* **1996**, *17*, 1.
- (261) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
- (262) Werner, H.-J.; Knowles, P. J.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Celani, P.; Cooper, D. L.; Deegan, M. J. O.; J., D. A.; Eckert, F.; Hampel, C.; Hetzer, G.; Korona, T.;

## Références Bibliographiques

Lindh, R.; Lloyd, A. W.; McNicholas, S. J.; Manby, F. R.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Rauhut, G.; Schütz, M.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T. *MOLPRO 2002.3*, 2002.

(263) Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.

(264) Woon, D. E.; Dunning Jr., T. H. *J. Chem. Phys.* **1993**, *98*, 1358.

(265) Delhommelle, J. Etablissement de potentiels d'interactions pour la simulation moléculaire. Application à la prédiction des équilibres liquide-vapeur de mélanges binaires alcane-molécule multipolaire, Paris XI, UFR scientifique d'Orsay, 2000.

(266) DeVoe, H.; Tinoco, I. *J. Mol. Biol.* **1962**, *4*, 518.

(267) D'Orsogna, M. R.; Rudnick, J. *Phys. Rev. E* **2002**, *66*, art. no.

(268) Wolfram, S. *The Mathematica Book, 4th edition*; Wolfram Research: Cambridge, 1999.

(269) Sato, N.; Ito, S.; Sugiura, K.; Yamamoto, M. *J. Phys. Chem. A* **1999**, *103*, 3402.

(270) Schlosser, D. W.; Philpott, M. R. *J. Chem. Phys.* **1982**, *77*, 1969.

(271) Argyrakis, P.; Blumen, A.; Kopelman, R.; Zumofen, G. *J. Phys. Chem.* **1984**, *88*, 1973.

(272) Clavaguera-Sarrio, C. Etude théorique de composés moléculaires d'actinides : propriétés électroniques, vibrationnelles et dynamiques, Toulouse, Université Paul Sabatier, 2003.

(273) Chaudhury, P.; Bhattacharyya, S. P. *Int. J. Quantum Chem.* **1999**, *74*, 153.

(274) Ayala, P. Y.; Schlegel, H. B. *J. Chem. Phys.* **1998**, *108*, 7560.

(275) Archirel, P. *Chem. Phys.* **1995**, *193*, 55.











## RÉSUMÉ

---

Les phénomènes de transfert de charge et d'excitation se manifestent dans des domaines aussi variés que la photosynthèse, l'électronique moléculaire, ou les dommages causés aux acides nucléiques. Pour autant, leur caractérisation théorique est notoirement délicate, de par la nature des interactions mises en œuvre et la taille des systèmes concernés. S'appuyant sur les méthodes à hamiltonien effectif, cette étude propose une approche théorique raisonnée de ces phénomènes. Les interactions responsables des phénomènes de transfert sont identifiées, puis modélisées à l'aide de potentiels modèle spécifiques mis au point sur la base de calculs *ab initio* de bonne qualité. L'élaboration de ces potentiels est présentée, tant au plan méthodologique qu'au plan pratique.

Le phénomène de résonance de la charge dans les agrégats aromatiques ionisés constitue la première application de ces méthodes. Si l'influence du terme à transfert de charge est dominante pour les agrégats de petite taille, il rentre en compétition avec les effets de polarisation dans de plus grands systèmes. L'exploration globale des surfaces d'énergie potentielle de ces agrégats (neutres et chargés) permet la détermination de leurs structures optimales et des propriétés électroniques associées. Les potentiels d'ionisation expérimentaux de ces agrégats sont en outre prédits par des simulations de dynamique moléculaire.

Le problème du transfert d'excitation à l'état singulet dans de petits fragments d'ADN est ensuite abordé. Il est montré que les états excités de tels fragments peuvent concerner de larges pans de la séquence de bases, favorisant la migration de l'énergie le long de la double hélice. L'influence du désordre structural sur ces états est par ailleurs évaluée. Enfin, le couplage entre degrés de liberté classiques et quantiques est introduit, à l'aide de la méthode *trajectory surface hopping*, et appliquée à l'étude de la désexcitation non radiative d'agrégats aromatiques excités.

## ABSTRACT

---

Charge and excitation transfer phenomena play a predominant role in varied physical and chemical problems such as photosynthesis, molecular electronics, or the damage to nucleic acids. However, due to the subtlety of the interactions involved and the extent of the systems in which they occur, these phenomena are complex to characterize theoretically. In this work, the interactions are studied, and then modelled, using a combination of an effective Hamiltonian methodology and model potentials derived from accurate *ab initio* calculations. The design of these potentials is presented, from a methodological and practical standpoint.

The model is applied, in the first place, to the study of charge resonance phenomena in aromatic cluster ions. Although charge transfer effects dominate in small clusters, they compete with polarization effects in larger ones. Global exploration of the potential energy surfaces of the neutral and ionic clusters provides the optimal structures for these species, as well as the associated electronic properties. The experimental ionization potentials of these species are correctly predicted by molecular dynamics simulations.

In a second part, the study of singlet excitation transfer in small DNA fragments is undertaken. It is shown that the electronic excited states of such fragments may involve large tracts inside the base pair sequence, promoting energy migration along the double helix. The influence of structural disorder on the topology of the excited states is also evaluated. Finally, the coupling between classical and quantum mechanical degrees of freedom is implemented using a trajectory surface hopping formalism, applied to the non-radiative de-excitation of excited aromatic clusters.