Theoretical study of chemical reactivity: from nano-objects to follow the entropy via metadynamics simulations

Mircea Oltean

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Chimie Physique Moléculaire et Structurale**

Arrêté ministériel : 7 août 2006

Présentée par

Mircea OLTEAN

Thèse dirigée par **Anne MILET**

préparée au sein du **Département de Chimie Moléculaire UMR UJF-CNRS 5250, Laboratoire du Chimie Théorique**
dans **l'École Doctorale Chimie et Sciences de Vivant**

Etude théorique de la réactivité chimique : des nano-objets au suivi de l'entropie par métadynamique

Thèse soutenue publiquement le **9 décembre 2010**,
devant le jury composé de :

Pascale, MALDIVI
Directeur du Recherche CEA Grenoble, Président
Stéphane, HUMBEL
Professeur Université Paul Cézanne Aix-Marseille III, Rapporteur
Gérald, MONARD
Professeur Université Henry Poincaré Nancy I, Rapporteur
Vasile, CHIŞ
Professeur Université Babeș-Bolyai Cluj-Napoca, Membre
Anne, MILET
Professeur Université Joseph Fourier Grenoble I, Membre

Merci !

Je remercie très sincèrement ma directrice de thèse, Anne Milet, pour m'avoir donné l'opportunité de travailler sur ce sujet passionnant, pour avoir su m'encadrer avec efficacité tout en me laissant l'initiative, et surtout pour la confiance accordée pendant ma thèse. . .

J'exprime mes vifs remerciements aux membres du Jury et, en particulier aux rapporteurs, pour avoir accepté de juger ce travail.

Je remercie également le Département de Chimie Moléculaire pour la bourse the thèse accordée. Je remercie tous les membres de ce laboratoire pour m'avoir accueilli au sein de cette structure. Je tiens à remercier en particulier le service informatique, pour le soutien technique, et plus largement le CECIC pour les moyens informatiques.

Je tiens à remercier vivement tous ceux qui ont su me faire confiance, Carine Michel, Rosa Bulo, Hélène Jamet, Marie-Louise Dheu-Andries et tous les reste des mes collaborateurs qui ont rendu ce travail passionnant.

Un grand merci à Marius, Adriana, Maylis, Christian, et tout l'ensemble de la chimie théorique du DCM qui ont su rendre le travail en laboratoire si agréable.

Et pour finir, merci à tous ceux qui m'ont soutenu pendant ces années grenobloises: ma famille et surtout mes parents.

TABLE OF CONTENTS

RESUME

La méthode théorique la plus utilisée pour étudier la réactivité chimique est la recherche des points stationnaires de la surface d'énergie potentielle. Mais la taille croissante des systèmes étudiés augmente la flexibilité du système et des simulations incluant l'aspect dynamique devient alors nécessaire.

Dans le présent manuscrit, nous étudions diverses problématiques chimiques. Nous avons ainsi modélisé des interrupteurs et rotors moléculaires. Nous avons testé pour ces systèmes ainsi que pour l'interaction cryptophane-petites molécules, où les interactions faibles jouent un rôle important, la validité des DFT-D. D'un point de vue de la réactivité chimique, nous avons étudié une réaction de cycloaddition catalysée par un complexe de Pt (II). Finalement nous avons déterminé les grandeurs thermodynamiques et en particulier l'entropie de réactions péricycliques : une réaction intramoléculaire Diels-Alder et la réaction de Claisen. L'ensemble de ces phénomènes a été étudié par métadynamique et/ou calculs statiques.

La première étude nous a permis d'évaluer les interactions, qui sont présentes dans un interrupteur et un rotor moléculaire. Sur la base de ces interactions nous avons proposé des nouveaux substituant et fait le lien entre la nature du substituant et les propriétés observées. Puis, nous avons implémenté l'approche de Grimme pour tenir compte des interactions de type London dans le logiciel Gaussian09 pour l'atome de Pt. La dernière partie du manuscrit propose deux implémentations : l'une pour le suivi des effets entropiques en utilisant l'énergie comme variable collective dans la métadynamique, l'autre concernant la méthode du bias-exchange-métadynamique

ABSTRACT

The most theoretical method used to study the chemical reactivity is the search for stationary points of potential energy surface. But the increasing size of systems studied increases system flexibility and simulation's including the dynamic aspect becomes necessary.

In this manuscript, we study various chemical problems. We have modeled the molecular switches and rotors. We tested for these systems as well as the interaction cryptophane-small molecules, where weak interactions play an important role, the validity of the DFT-D. From a viewpoint of chemical reactivity, we studied a cycloaddition reaction catalyzed by a complex of Pt (II). Finally we determined the thermodynamic quantities and in particular the entropy of pericyclic reactions: an intramolecular reaction Dies-Alder and Claisen reaction. All of these phenomena have been studied by metadynamics and / or static calculations.

The first study allowed us to assess the interactions that are present in a molecular switch and a rotor. Based on these interactions we have proposed new substituent and the link between the substituent and the observed properties. Then we implemented the Grimme approach to take account of London dispersion interactions for Pt atom in the software Gaussian09. The last part of the manuscript proposed two implementations: one for monitoring the entropic effect by using the energy as a collective variable in the metadynamics simulations, the other on the method of bias- exchange metadynamics.

INTRODUCTION GENERALE

La compréhension des processus chimiques nécessite souvent des simulations effectuées à différents niveaux de la théorie. Les calculs de chimie quantique représentent aujourd'hui un outil puissant capable de répondre aux nombreuses questions liées à différentes propriétés moléculaires. Par exemple, les approches de chimie quantique sont utilisées pour déterminer le mécanisme d'une réaction donnée, les états de transition moléculaires impliqués, l'affectation et l'interprétation des caractéristiques subtiles d'un spectre particulier. Pour une réponse appropriée, le niveau de la modélisation doit être suffisamment fiable pour reproduire fidèlement la réalité chimique. Mais, au-delà des difficultés de la modélisation, les théoriciens ont à jeter un nouvel éclairage pour expliquer et rationaliser les données expérimentales. Dans toutes les recherches que j'ai faites au cours de mon doctorat, j'ai poursuivi ces objectifs: comprendre et appliquer les modèles utilisés dans le domaine de la chimie quantique, la compréhension des processus qui régissent la réactivité chimique et, si possible, améliorer les modèles utilisés en chimie quantique.

Aujourd'hui, la méthode théorique la plus simple pour étudier la réactivité chimique est la recherche des points fixes de la surface d'énergie potentielle, PES[1,2]. Mais avec la taille croissante des systèmes étudiés, augmentent aussi la flexibilité des systèmes et la pertinence de leur description par les minima de la PES. Par exemple: l'eau peut elle être décrite à la température ambiante par un minimum? De même, une molécule avec sa sphère de solvatation peut elle être décrite de manière efficace et appropriée par des procédures d'optimisation de la géométrie? Le développement actuel de la dynamique moléculaire *ab initio* semble une solution attrayante à ces questions. En outre, la dynamique peut donner beaucoup d'informations statistiques, y compris les quantités thermodynamiques énergétiques, à travers les calculs de la trajectoire. À ce jour, ces méthodes ont été appliquées avec succès principalement dans le domaine des phases condensées, mais, elles ne sont guère appliquées dans la réactivité chimique, en raison des barrières élevées (quelques kcal.mol-1). En effet, les réactions chimiques apparaissent comme des événements rares et n'ont aucune chance d'être simulées dans un temps de simulation de quelques pico-secondes. Pire encore, les réactions chimiques impliquent la formation de liaisons et de rupture de la liaison et donc l'énergie de ces systèmes doit être calculée à un niveau quantique. Par exemple, une réaction chimique avec une énergie d'activation de 20 kcal mol^{-1} à 30 ° Celsius (90 ° F) pourra

atteindre un rendement de 50%, lors d'une simulation de 0,1 ms = 10^{11} femtosecondes. Au cours d'une simulation de dynamique moléculaire, le pas de temps de la simulation doit être suffisamment petit pour assurer la conservation de l'énergie. Donc, la dynamique *ab initio* Born-Oppenheimer, avec un pas de 1fs est une valeur réaliste. Si 3 minutes de temps de calcul sont nécessaires pour chaque étape (évaluation de l'énergie et de forces), $3x10^{11}$ minutes seront nécessaires, ce qui signifie environ 570 776 années des temps de calcul! Cette estimation de base présente clairement le problème de l'observation des événements rares dans la dynamique moléculaire classique.

Pour résoudre ce problème, plusieurs solutions ont été proposées, comme par exemple le « Transition Path Sampling » et de ses variantes, développé par Dellago, Bohlius, et al.[5,6], le « Milsestoning »[7], « Taboo search »[8], le « MD multicanoniques »[9] etc. Malheureusement, certains d'entre eux nécessitent soit le calcul d'un grand nombre de trajectoires, soit une bonne connaissance des coordonnées de la réaction, soit un grand nombre d'échantillonnages sur les réactifs et les produits pour calculer la connexion correspondant au chemin de réaction comme dans la méthode « Transition Path Sampling ».

Alessandro Laio et Michele Parrinello ont introduit une méthode novatrice nommée «Metadynamique». L'idée principale de cette méthode est d'empêcher le système de revisiter les mêmes «lieux» afin d'accélérer les événements rares et il a été utilisé avec succès dans de nombreux cas[10,11]. Dès son apparition en 2002, cette méthode a été développée en permanence et à ce moment a de nouvelles extensions telles que: « multiples walkers »[12], « Parallel tempering metadynamics »[13] et « Bias-exchange metadynamics »[14]. Toutes ces extensions ont été proposées avec l'objectif principal d'améliorer la précision et le temps de simulation pour une métadynamique ordinaire.

Dans le présent travail, nous appliquons la metadynamique et les modèles quantiques classiques pour caractériser les différents problèmes dans le domaine de la chimie organique et inorganique. Nous présentons également les améliorations apportées par notre groupe à la méthode metadynamique.

La première partie présente une introduction générale dans le domaine de la chimie quantique. Nous allons présenter les méthodes que nous avons utilisées dans notre recherche, en particulier les capacités des méthodes adaptées aux problèmes notamment étudiés: *ab initio* de dynamique moléculaire, théorie de la fonctionnelle de densité (DFT) et DFT avec des corrections pour la dispersion (DFT-D), méthode semi-empirique PM6 et aussi une

courte description du code CP2K accompagné de son module « Quickstep », metadynamique et ses extensions.

Dans le deuxième chapitre nous allons présenter l'interrupteur moléculaire et les rotors moléculaires traités par les outils de la chimie quantique. Ce chapitre mettra en évidence les capacités et le défi de la chimie quantique pour étudier ces systèmes nano moléculaire. Cela représente le travail effectué en collaboration avec un groupe expérimental, de l'équipe CIRE du DCM. Nous utilisons la méthode de chimie quantique pour caractériser ces types de nano machines moléculaires et prédire les substituants qui peuvent être utilisés dans la synthèse, afin d'obtenir des molécules ayant des propriétés spécifiques.

Dans le troisième chapitre, nous proposons de trouver l'énergie de liaison d'un système non covalent entre un système comme le Cryptophane complexé avec différentes petites molécules. Pour la deuxième partie de ce chapitre, nous présentons les développements réalisés dans le domaine de la DFT-D par notre groupe en particulier dans les paramétrages et la mise en œuvre des paramètres de l'atome de platine utilisant la méthode de Grimme[15] et implémenté dans le programme Gaussian09. À la fin de ce chapitre, nous avons étudié la réaction d'un complexe Pt(II) utilisé comme catalytiseur dans une réaction organique de cycloaddition [2+1] en phase gazeuse.

Dans le dernier chapitre en utilisant les développements réalisés dans notre équipe, dans le domaine de la métadynamique, nous calculons les données thermodynamiques pour une réaction Diels Alder intramoléculaire et une réaction de Claisen. En outre, nous développons une application «Entropy.py» pour calculer ces valeurs, à partir de l'énergie libre de surface en fonction de l'énergie potentielle et d'une coordonnée géométrique de la réaction. Ici je vais vous présenter aussi la stratégie de mise en œuvre de cette nouvelle variable collective dans la nouvelle version de logiciel CP2K et le "bias_exe.py", application qui est utilisée pour effectuer la « Bias-Exchange Metadynamics. »

GENERAL INTRODUCTION

The understanding of the chemical processes often requires simulations performed at different levels of theory. Quantum chemical calculations represent now a powerful tool able to answer many questions related to different molecular properties. For instance, quantum chemistry approaches are used to determine the mechanism of a given reaction, the molecular transition states involved, the assignment and interpretation of subtle features of a particular spectrum. For an appropriate answer, the level of modeling has to be reliable enough to reproduce accurately the chemical reality. But, beyond the difficulties of modeling, theoreticians have to bring new insight to explain and rationalize experimental data. From all the research I have made during my PhD, I have pursued these goals: understanding and applying the models used in quantum chemistry field, understanding the processes that govern chemical reactivity and, if possible, improving the models used in quantum chemistry.

Nowadays, the most straightforward theoretical method to investigate the chemical reactivity is the search for stationary points of the potential energy surface, PES.[1,2]. But with the ever-growing size of the systems of interest, their consequent increasing flexibility questions the pertinence of their description by minima of the PES. For example: can the water at room temperature be described by a minimum? Similarly, can a molecule with its first solvation sphere be described efficiently and appropriately by geometry optimization procedures? The actual development of the molecular *ab initio* dynamics appears an appealing solution to these questions. Moreover, dynamics can give a lot of statistical information, including energetic thermodynamic quantities, through the computations of the trajectories. To date, these methods were applied with great success mainly in the field of condensed phases[3,4] but, they were hardly applied in chemical reactivity, because of the high barriers involved (few $kcal.mol^{-1}$). Indeed, chemical reactions appear as "rare events" and have no chance to be simulated in the simulation's time of a couple of pico-seconds. Even worse, chemical reactions imply bond's formation and bond's breaking and so the energy of these systems has to be computed at a quantum level. For example, a chemical process with an activation energy of $20kcal.mol^{-1}$ at 30° Celsius (90°F) will need, to reach a 50% yield, at least 0.1ms $=10^{11}$ femtoseconds. During a molecular dynamics simulation, the time steps of the simulation should be small enough to ensure the conservation of energy… so using *ab initio* Born-Oppenheimer dynamics, steps of 1fs is a fair value. If 3 minutes of computation time are needed for each step (evaluation of the energy and forces, so depending on the machine

and the investigated system, 3 minutes is a good average value), 3.10^{11} minutes will be needed, which means about 570 776 years of computational time! This basic estimation clearly exhibits the problem of observing rare events in conventional molecular dynamics.

To solve this problem, several solutions have been proposed, as for example the Transition Path Sampling and its variants, developed par Dellago, Bohlius, et al.[5,6], the milestoning[7], taboo search[8], the multicanonical MD[9] … . Unfortunately, some of them require the computation of a large number of trajectories, or they rely on a good knowledge of the reaction coordinates or they need data on the reactants and products to compute the correspondent connecting reaction path as in the Transition Path Sampling method. Alessandro Laio and Michele Parrinello have introduced an innovative method: the metadynamics. The main idea of this method is to prevent the system from revisiting the same "places" in order to accelerate rare events and it has been used with great success in many cases[10,11]. From its appearance in 2002 this method has been developed continuously and at this point has new extensions such as: *Multiple walkers*[12], *Parallel tempering metadynamics*[13] and *Bias exchange metadynamics*[14]. All these extensions were proposed with the main goal to improve the accuracy and simulation time for an ordinary metadynamics.

In the present work we apply the metadynamics and classical quantum models to characterized different problems in the organic and inorganic chemistry field. We also present the improvements made by our group to the metadynamics method.

The first chapter presents a general introduction in the field of quantum chemistry. We will present the methods that we have used in our research, especially the capabilities of the methods suitable for the particular problems studied: a*b-initio* molecular dynamics, density functional theory (DFT) and DFT with long range corrections (DFT-D), semiempirical method PM6 and also a short description of the CP2K code together with Quickstep program, Metadynamics and their extensions.

In the second chapter we will present (reveal) the molecular switch and molecular rotors treated by quantum chemistry tools. This chapter will highlight the capabilities and the challenge of the quantum chemistry in the field of nano-molecular design. This represents the work made in collaboration with an experimental group, of CIRE from the DCM department. We use the quantum chemistry approach to characterize these types of nanomachines and to predict the substituents that can be used in the synthesis, in order to have materials with a specific property.

In the third chapter we propose to find the binding energy of a non-covalent bonding system between a host-guest system such as Cryptophane complexed with different small molecules. To the second part of this chapter we present the developments made in the field of the DFT-D by our group especially in the parameterizations and implementation part of the parameters for the platinum atom using Grimme's method[15] implemented in the Gaussian09 program. At the end of this chapter we studied a reactivity pathway of a Pt(II)-complex used as catalytic compound in an organic reactions of [2+1] cycloaddition in gas phase.

In the fourth chapter using the developments made in our lab, to the metadynamics simulations, we compute the thermodynamic data for a Dies-Alders intermolecular reaction and a Claisen reaction. Also, we develop an application “EntroPy” to compute these values, starting from the free energy surface as a function of the potential energy and one ordinary geometrical reaction coordinate. Here I will present also the strategy of implementation of this new collective variable in the new version of the CP2K and the “bias_exe.py”, application that is used to perform bias exchange metadynamics.

Bibliography to general introduction

(1) Bruin, T. J. M. T. D.; Milet, A.; Robert, F.; Gimbert, Y.; Greene, A. E. Theoretical Study of the Regiochemistry-Determining Step of the Pauson−Khand Reaction. *Journal of the American Chemical Society***2001**, *123*, 7184-7185.

(2) Dedieu A.; Hutscka F.; Milet A. Dans *Transition State Modeling for Catalysis*; ACS Symposium Series; American Chemical Society, 1999; Vol. 721, p. 100-113.

(3) Kuo, I. W.; Mundy, C. J.; McGrath, M. J.; Siepmann, J. I.; VandeVondele, J.; Sprik, M.; Hutter, J.; Chen, B.; Klein, M. L.; Mohamed, F.; Krack, M.; Parrinello, M. Liquid Water from First Principles: Investigation of Different Sampling Approaches. *The Journal of Physical Chemistry B***2004**, *108*, 12990-12998.

(4) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Oxford University Press, 1989.

(5) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. Transition path sampling and the calculation of rate constants. *J. Chem. Phys.***1998**, *108*, 1964.

(6) Dellago, C.; Bolhuis, P.; Geissler, P. Dans *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1*; 2006; p. 349-391.

(7) Faradjian, A. K.; Elber, R. Computing time scales from reaction coordinates by milestoning. *J. Chem. Phys.***2004**, *120*, 10880.

(8) Cvijovicacute, D.; Klinowski, J. Taboo Search: An Approach to the Multiple Minima Problem. *Science***1995**, *267*, 664-666.

(9) Nakajima, N.; Higo, J.; Kidera, A.; Nakamura, H. Flexible docking of a ligand peptide to a receptor protein by multicanonical molecular dynamics simulation. *Chemical Physics Letters***1997**, *278*, 297-301.

(10) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proceedings of the National Academy of Sciences of the United States of America***2002**, *99*, 12562-12566.

(11) Micheletti, C.; Laio, A.; Parrinello, M. Reconstructing the Density of States by History-Dependent Metadynamics. *Phys. Rev. Lett.***2004**, *92*, 170601.

(12) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics†. *The Journal of Physical Chemistry B***2006**, *110*, 3533-3539.

(13) Bussi, G.; Gervasio, F. L.; Laio, A.; Parrinello, M. Free-Energy Landscape for β Hairpin Folding from Combined Parallel Tempering and Metadynamics. *Journal of the American Chemical Society***2006**, *128*, 13435-13441.

(14) Piana, S.; Laio, A. A Bias-Exchange Approach to Protein Folding. *The Journal of Physical Chemistry B***2007**, *111*, 4553-4559.

(15) Stefan Grimme Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.***2006**, *27*, 1787-1799.

I Theoretical Background

Chapter 1 METHODOLOGY

Résume

Dans ce chapitre d'introduction, je présenterai les méthodes et les stratégies utilisées au cours du ma thèse.
Ce chapitre va exposai les thèmes suivant:

- La dynamique moléculaire de type Born-Oppenheimer ou *ab-initio*
- Les méthodes DFT et DFT avec des corrections pour la dispersion (DFT-D)
- Méthodes semiempirique (PM6)
- Théorie de la méthode GPW avec une courte description du programme de Quickstep qui fait partie du code CP2K.
- Métadynamique et les extensions de la métadynamique « Bias-Exchange » et les implémentations du « Bias-Exchange » dans un programme externes écrite en python qui utilisent le module « Quickstep » pour effectuer la métadynamique

1.1 Introduction

In this introductory chapter, I will present the methods and the strategies used during my thesis.

This chapter will organized as follows the next topics:

- Born-Oppenheimer or *ab-initio* Molecular Dynamics
- DFT methods and DFT with long range corrections (DFT-D)
- Semiempirical methods (PM6)
- Theory of the GPW method with short description of the Quickstep program which is part of the CP2K code. Basis sets in GPW method and methods to perform wave function optimization.
- Metadynamics and Bias-Exchange extension of metadynamics application. Implementations of the Bias-Exchange metadynamics in an external python program which use the Quickstep modulus to perform metadynamics.

1.2 Born-Oppenheimer Molecular Dynamics (BOMD)

Born Oppenheimer approximation appears from the fact that electrons are much lighter that nuclei, roughly 1600 times. Because of that electrons are much faster comparatively with the nuclei (ions), so the nuclei feel only the average force of the electrons. Such considerations make that the movement of the nuclei and the movement of the electrons can be decoupled. In other terms we can treat separately the movement of the ions and the movement of electrons. This is a great advantage in molecular dynamic simulations[1a]. We can propagate the nuclei with equations of motion, equation (1.1), and on the other hand we can calculate the energy and the gradient of the energy. To determining the minimum energy of the system at this potential V(R) (position of the nuclei) we must solve the Schrodinger equation, eq. (1.2) using the Hamiltonian at the level of theory that is relevant for our system.

$$M_I \ddot{R}_I = -\nabla_I \min_{\psi_0} \langle \psi_0 | H | \psi_0 \rangle \qquad (1.1)$$

[1a] The length and time scale of typical *ab-initio* molecular dynamic simulations are currently given by approximately 10 to 1000 atoms and 1 to 100 ps

$$E_0\psi_0 = H\psi_0 \quad (1.2)$$

In the equation of motion (1) and in Schrödinger equation (2): M_I are the masses of the ions, R_I are the positions of the nuclei in space, ∇_I (the gradient) represents the derivative with respect to the $R_I$of the potential $V(R_I) = \min_{\psi_0}\langle\psi_0|H|\psi_0\rangle$, $\psi_0 = \psi_0(x_i)$ is the wave function of the electrons, H is the Hamiltonian and E_0 is the ground-state energy.

The wave function ψ_0 has to be normalized and must satisfied the Pauli principle (electrons are fermions). So she must be antisymmetric with respect to the exchange of the variables $x_i = (r_i, \sigma_i)$, where the r_i are the spatial coordinates of the electron "i" and σ_i are the spin coordinate of the electron "i".

For a molecular system with N atoms (ions) and an "*n*" number of electrons the non-relativistic Hamiltonian can be written in atomic units as:

$$H = T + V_{Ne} + V_{ee} + V_{NN} = \sum_i^n -\frac{1}{2}\cdot\nabla_i^2 - \sum_i^n\sum_I^N \frac{Z_I}{|R_I - r_i|} + \sum_{i<J}^n \frac{1}{|r_i - r_j|} + \sum_{I<J}^N \frac{Z_I Z_J}{|R_I - R_J|} \quad (1.3)$$

where the T is the kinetic energy of the electrons, V_{Ne} represents the nucleus electron interactions, V_{ee} the electron electron interaction and V_{NN} is a constant which represent the nuclear nuclear interaction.

Because we can decouple the problem in two distinct parts, electronic part and ion part, electronic wave function doesn't have to be in the ground state. We can have the wave function in an exited state or in the ground states. Nevertheless in general case, with BOMD we treat systems that have electronic wave function in the ground state.

To have an application, which performs BOMD, it is easy to interconnect two distinct codes. One to calculate the energy and the gradient and the second one to propagate the nuclei at the next step in potential V(R) calculated with the first code. This was the standard procedure for BOMD applications in the past, but now the majority of quantum chemical packages incorporated both applications in one standalone application (framework). Dalton, DMol, Gamess, Gaussian, DeMon, Gromacs, Cp2K … are examples of codes that perform BOMD. Each of these codes has their own characteristics and is suitable for BOMD calculations at different levels of theory.

The most expensive part in a *ab-initio*, BO MD simulations is to calculate the $V(R_I)$ and forces ($-\nabla_I V(R_I)$) because these quantities are scaled in time with $O(N^4)$ for a system with N atoms treated at high level of theory (i.e. DFT) against O(N) for systems treated at classical potential. So to be able to use this method we must be able to have applications that can reduce the time for computing this quantities (speed up the SCF), or to have applications which run in parallel or to develop methods that can accelerate the dynamics.

1.3 Density Functional Theory (DFT)

The concept of Density Functional has the origin in the 1920s by the work of E. Fermi and L. H. Tomas [1,2] where the authors introduce the idea of expressing the energy of a system as a function of the total electron density. Later in 1951 J.C. Slater[3] developed the Hartree – Fock method using this future concept of Density Functional. Later this method was known as Xα method, which is the predecessor of the DFT theory.

Hohenberg and Kohn[4] postulate two theorems based on which the Energy of a system in the ground state at an external potential $V_{ext}(r)$ with an ***n*** numbers of electrons can be determined in a unique way in principle using just electronic density.

$$\rho(r) = |\psi_0(r)|^2 \quad (1.4)$$

The first theorem, which demonstrates the legitimacy of this formalism, says:

"*The external potential $V_{ext}(r)$ is (within a constant) and hence the total energy $E_0$is a unique functional form of $\rho(r)$; we say that the full many particle ground state is a unique functional of $\rho(r)$*" This theorem (lemma) is mathematical rigorous and can be easy demonstrated by *reductivo ad absurdum* method.

The second theorem presents how it is possible to use this formalism to find the Energy E_0 of the ground state:

"*The ground state energy can be obtained variationally: the density that minimizes the total energy is the exact ground state density*."

$$E_v(\rho(r)) = \int \rho(r)v(r)dr + F_{HK}(\rho(r)) \quad (1.5)$$

where the $F_{HK}(\rho(r))$ represents the universal functional of the density. The main problem is that we don't know this form of the functional because it is quasi impossible to represent the kinetic energy of the system in this form.

In 1965 W. Kohn and L. Sham[5] proposed a practical way to solve the Hohenberg-Kohn theorem for a set of interacting electrons starting from a virtual system of non-interacting electrons. The central premise in the Kohn-Sham approach is that the kinetic energy functional of a system can be split into two parts: one part that can be calculated exactly considering electrons as non interacting particles and one other part with a small correction term accounting for the kinetic energy interaction. These premises were developed with the introduction of atomic orbital's formalism, which became the fundamental foundation for the current applications of DFT in the computational chemistry field. The electronic energy of the ground state of a system with ***n*** electrons and ***N*** nuclei in Kohn-Sham formalism can be written as:

$$E(\rho) = -\frac{1}{2}\sum_{i=1}^{n}\int \psi_i^*(r_1)\nabla_i^2\psi_i(r_1)dr_1 - \sum_{I=1}^{N}\int \frac{Z_I}{r_{Ii}}\rho(r_1)dr_1 + \frac{1}{2}\int\int \frac{\rho(r_1)\rho(r_2)}{r_{12}}dr_1dr_2 + E^{XC}(\rho) \quad (1.6)$$

where ψ_i are the Kohn Sham orbital's with i=1…n. The first term represents the kinetic energy of noninteracting electrons, the second term accounts for the nuclear electron interactions, and the third term corresponds to the Coulomb repulsion term between the total charge distributions at r_1 and r_2. The last one is known as the Exchange –Correlation term, which represents the correction to the kinetic energy arising from the interacting nature of the electrons, and all nonclassic corrections to the electron-electron repulsion energy.

This formulation of Kohn – Sham is adequate because the $E^{XC}(\rho)$ is much smaller that $F_{HK}(\rho(r))$, and better approximations exist for it. From a computational point of view the density $\rho(r)$ of the system is computed from the orbitals ψ_i; because from this spin orbital's it is easy to compute the corresponding density:

$$\rho(r) = \sum_i |\psi_i|^2 \quad (1.7)$$

1.4 Long range correction to DFT

Kohn–Sham density functional theory (DFT) is nowadays the most widely used quantum chemical method for electronic structure calculations in chemistry and physics. This is due to

the fact that this method presents the great advantages of being very efficient, in terms of accuracy versus time of calculation. But the main inconvenience in almost all current density exchange-correlation functionals ($E^{XC}(\rho)$) is the lack of proper attractive long-range interactions. For systems that present non-covalent inter- and intra-molecular interactions, for example in the field of supermolecular chemistry and biochemistry, this term is very important to describe correctly the energetic stability of the systems. This attractive term is part of the electron correlation functional and the form of this functional can be derived from basic physical principles. In a very basic picture we can see this interaction like interactions between two dipole moments that is proportional with $\sim\frac{1}{r^6}$.

1.4.1 The functional form of the long-range corrections

The functional form of this long-range corrections term presented herein follows the work of G.C Maitland et al.[6] and the work of M. E. Foster et al.[7] Let's imagined now the simplest molecular system of two non-polar molecules separated by a distance **r**as shown in Figure 1.1. The electron of each individual molecule oscillates around each molecule with an ω_0 (angular frequency) inducing an instant dipole moment ($\mu_a = q \cdot r_a(t)$) around the monomer.

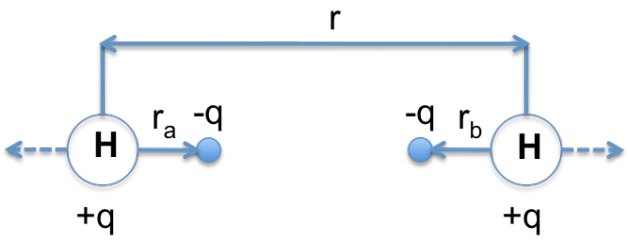


Figure 1.1 Two non-polar molecules separated by a distance r. Instantaneous dipole moment around each monomer is present in the approximations of $r >> r_a, r_b$

The Schrodinger equation of an individual monomer is:

$$\frac{1}{m}\frac{\partial^2\psi_a}{\partial r_a^2} + \frac{2}{\hbar^2}\left(E_a - \frac{1}{2}kr_a^2\right)\psi_a = 0 \quad (1.8)$$

Solving equation (1.8) we found the oscillator energy. The solutions of this equation are given in equation (1.9) where the n_a and n_b are the principal quantum numbers.

$$E_a = \left(n_a + \frac{1}{2}\right)\hbar\omega_0 ; E_b = \left(n_b + \frac{1}{2}\right)\hbar\omega_0 \quad (1.9)$$

For long distance separations between two monomers ($r \to \infty$) we can say that there are no interactions between the two molecules and the energy of the system is simply the sum of the each individual monomer. So, the energy of the ground state in this case is: $E(\infty) = E_a + E_b = \hbar\omega_0$

Now let's imagine a second case with the two monomers, which are interacting. The Schrodinger equation became equation (1.10) where the last term in the parentheses represents the potential energy between the dipoles. Dipoles, which are locally formed around each monomer emphasized in Figure 1.1.

$$\frac{1}{m}\frac{\partial^2\psi}{\partial r_a^2} + \frac{1}{m}\frac{\partial^2\psi}{\partial r_b^2} + \frac{2}{\hbar^2}\left(E - \frac{1}{2}kr_a^2 - \frac{1}{2}kr_b^2 + \frac{2r_a r_b q^2}{4\pi\varepsilon_0 r^3}\right)\psi = 0 \quad (1.10)$$

If we use the variable R_1, R_2, k_1 and k_2 (see equation 1.11) we can rewrite equation 1.10 and we will obtain equation 1.12, which is the Schrödinger equation for two non-interacting oscillators.

$$R_1 = \frac{r_a + r_b}{\sqrt{2}}; R_2 = \frac{r_a - r_b}{\sqrt{2}}; k_1 = k - \frac{2q^2}{4\pi\varepsilon_0 r^3}; k_2 = k + \frac{2q^2}{4\pi\varepsilon_0 r^3} \quad (1.11)$$

$$\frac{1}{m}\frac{\partial^2\psi}{\partial R_1^2} + \frac{1}{m}\frac{\partial^2\psi}{\partial R_2^2} + \frac{2}{\hbar^2}\left(E - \frac{1}{2}k_1 R_1^2 - \frac{1}{2}k_2 R_2^2\right)\psi = 0 \quad (1.12)$$

Solving this equation (1.12) we find E(r) (equation 1.13) :

$$E(r) = \left(n_1 + \frac{1}{2}\right)\hbar\omega_1 + \left(n_2 + \frac{1}{2}\right)\hbar\omega_2 \quad (1.13)$$

For the ground state, E(r) has this expression:

$$E(r) = \frac{1}{2}\hbar(\omega_1 + \omega_2) \quad (1.14)$$

with ω_1 and ω_2 defined like in expression 1.15

$$\omega_1 = \omega_0\left[1 - \frac{2q^2}{4\pi\varepsilon_0 r^3 k}\right]^{1/2}; \omega_2 = \omega_0\left[1 + \frac{2q^2}{4\pi\varepsilon_0 r^3 k}\right]^{1/2} \quad (1.15)$$

If we substitute this frequencies in eq. 1.14 and use the binomial expansion theorem (eq. 1.16)

$$(x+y)^n = \sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k \text{ with } \binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (1.16)$$

we obtain the energy of the system in interaction. The energy in the ground state is:

$$E(r) = \hbar\omega_0 - \frac{q^4\hbar\omega_0}{2(4\pi\varepsilon_0)^2 r^6 k^2} + \ldots \quad (1.17)$$

In the expansion of the binomial solution we keep only the leading term, which is the first term of the expansion. If we subtract the energy when the two molecules are completely separated from the energy when the system is in interaction we obtain the dispersion energy:

$$E_{DISP} = E(r) - E(\infty) = -\frac{q^4\hbar\omega_0}{2(4\pi\varepsilon_0)^2 r^6 k^2} + \ldots \quad (1.18)$$

This energy is proportional with $\frac{1}{r^6}$ which demonstrates that the long-range corrections are dominated by this term. And the negative sign of this expression represents that this term is attractive, so the stabilization energy increases.

1.4.2 Classifications of the long-range corrections

Because of their non-local (long-range) character, the dispersion interactions are accurately accounted only by correlated wavefunction methods (post Hartree-Fock methods) like MP2 and CCSD(T). But these methods are very costly from a computational point of view and for large systems are impracticable to use. Thus the KS-DFT remains an open field and a challenge for the quantum chemistry community.

Stefan Grimme *et al.*[8] propose a classification of the long-range correction by the method used to construct the exchange-correlations functionals.

- *Conventional functional in the generalized gradient approximations (GGA) including hybrid or meta-hybrids.*

- *Special correlations functionals or orbital-based DFT methods.*
- *DFT/molecular mechanics – based hybrid schemes.*

The research group of Grimme has developed semiempirical corrections to the existing GGA type functional. The same idea was originally applied to HF theory for the correlation term in the 1970[9,10] but was forgotten over 30 years. In this kind of approach to correct Kohn-Sham energy the dispersion energy term is simply added.

$$E = E_{KS-DFT} + E_{DISP} \quad (1.19)$$

where $E_{KS\text{-}DFT}$ is the usual Kohn-Sham energy computed with the chosen density functional and E_{DISP} is an empirically correction term which has this form:

$$E_{DISP} = -s_6 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{C_6^{ij}}{R_{ij}^6} \cdot f_{dump}\left(R_{ij}^6\right) \quad (1.20)$$

where:

- S_6 represents the global scaling factor and depends only on the density functional used. For examples for BLYP, B97, PBE functionals this factor is 1.2, 1.25 and 0.75 respectively.
- $C_6^{ij} = \sqrt{C_6^i C_6^j}$ denotes the dispersion coefficient for atom pair ij, and is defined by the geometric mean of atomic C_6 coefficients.
- R_{ij}^6 denotes the interatomic distances between atoms pairs ij
- $f_{dump}\left(R_{ij}^6\right) = \frac{1}{1 + \exp\left(-d\left(\frac{R_{ij}}{R_{vdw}} - 1\right)\right)}$ is the damping function, which attenuates the $\frac{1}{r^6}$ long range interactions at short range separations between the atoms pairs ij. $R_{vdw} = R_0^i + R_0^j$ is the sum of Van der Waals radii of atoms pairs ij. "d" is an unitless parameter set to 20 in the case of Grimme approach, which determines the steepness of the dumping function. As this parameter increases, the function $f_{dump}$is more abruptly shutting off at low values of r. The depth of the potential well depends on the "d" and S_6 coefficients. When "d" increases together or in parallel with the S_6

coefficient, the depth of the potential well goes higher. However, the S_6 coefficient influences more this depth of the potential well comparatively with "d" coefficient.

The C_6 parameters are extracted from a DFT/PBE0 calculation of atomic ionization potentials and static dipole polarizabilities.

$$C_6^i = 0.05 N I_p^i \alpha^i \quad (1.21)$$

where the N represents the number of maximum electrons number in each row of the periodic table, e.g. 2, 10, 18, 36 and 54 for the atoms from rows 1-5 of the periodic table. I_p^i is the ionization potential of atom i and α^i is static polarizabilities of the atom i given in a.u.

The atomic Van der Waals radii (R_0^i) values are derived from the radius of the $0.01 a_0^{-3}$ electron density contour from ROHF/TZV computations of the atoms in the ground state.

The main disadvantage of this method is the empirical character of the coefficients determination and the fact that these values are fixed in the context of the molecular structures. Indeed these values should be derived specifically from molecular properties as it is done in recent work of Becke and Johnson[11]. For elements from group I and II and for transition metals, the differences between the coefficients of a free atom and the atoms from a molecule are so large that this kind of approach are so far to describe correctly the interactions. To overcome this issue, Grimme does an average of the coefficients for Scandium to Zinc atoms and from Yttrium to Cadmium atoms. Grimme postulates that this simple average procedure should be sufficient if the number of metal atoms in the molecule remains much smaller than the total number of atoms in the molecule.

Even if this method has some disadvantages, it is a state of the art in description of the dispersive energy of the Noncovalent Interactions systems. The results of the Interaction energy, for this kind of systems, computed at this level DFT-D are comparable with the interaction energy computed at high level of theory, but, with a huge difference of computational resources in favor of the DFT-D method.[12]

1.4.3 DiLabio Method (Dispersion correcting potentials)

The method of DiLabio *et al.*[13,14,15] is an empirical method to calculate the long range dispersion for Van der Waals systems using optimized, atom-centered non-local potentials that are normally used in the context of pseudopotentials for core-electrons. Dispersion

correcting potentials (DCPs) combines a long-range, weakly attractive potential with a short-range, weakly repulsive potential. Using this pseudopotential for the atoms, the valence electrons are forced to move in an environment that is closer to the real Van der Waals potential and in such manner we can model this kind of interactions.

The principal advantage of DCPs is that they can be used in conjunction with any computational chemistry program that can accommodate the use of effective core potentials without any modifications of the code. All of the properties can be computed with this method, geometry optimizations, vibration calculations, solvation, electronic excitations, etc… and the computational cost of this method is not different from a conventional DFT calculation.

One big limitation of this kind of approach (DCPs) is that is dependent on the pseudo-potential parameters for each combination of functional and basis set. The authors of this method argued that for functionals from the same family (e.g. B971 and PBE, which are both B86-based functionals) the DCPs tends to be almost the same and for that in certain cases the DCPs parameters are transferable.

1.5 Semiempirical methods

Semiempirical quantum chemistry methods are based on the *ab-initio* formalism, but make many approximations, which include parameters from experimental data. For example the two electron integrals are parametrizedand their results are fitted by a set of parameters, to reproduce results that best agree with experimental data or with data obtain at post Hartree-Fock level of theory. These empirical methods can be classified according to the approximations used: the methods restricted only to pi electrons such as the Pariser-Parr-Pople method[16,17], or to all valence electrons. These last methods, which are restricted to all valence electrons, can be also classified in several subgroups:

- Methods such as CNDO/2, INDO and NDDO. The implementations aimed to fit, not experiment, but *ab initio* minimum basis set results.
- MINDO, MNDO, AM1, PM3, RM1 and SAM1. Here the objective is to use parameters to fit experimental heats of formation, dipole moments, ionization potentials, and geometries. These methods have origins in the work of the group of Michael Dewar and J. J. P. Stewart, Reviews in Computational Chemistry, Volume 1, Eds. K. B. Lipkowitz and D. B. Boyd, VCH, New York, 45, (1990)

- ZINDO, methods whose primary aim is to calculate excited states and hence predict electronic spectra.M. Zerner, Reviews in Computational Chemistry, Volume 2, Eds. K. B. Lipkowitz and D. B. Boyd, VCH, New York, 313, (1991)

Semiempirical calculations are much faster than their *ab-initio* counterparts. The results obtained at semiempirical level are highly dependent by the sets of molecules used in parameterizations. This inconvenient is well known in the field of organometalic chemistry, where the metal ion can be found in many different oxidative states and different number of coordination's, which impose a limit in these semiempirical methods. Because of this fact, semiempirical calculations have been most successful in the description of organic compounds, where only a few elements are extensively used, (ex: O, N, C and H atoms).

1.5.1 PM6 method

The new PM6 method, NDDO method, is the newest one from its family. This method presents some modifications for the core-core interaction term and for the parameters optimization procedure comparatively with the previous ones PM3[18], (PM4 and PM5, which are unpublished), according to the author specifications.[19]

References data for molecular geometries used to parameterize this method (PM6) are taken from WebBook,[20] from thermochemistry compendia, from Cambridge Structural Database[21] and from *ab-initio* calculations. For *ab-initio* reference data, the HF and B3LYP level of theory were used combined with 6-31G(d) basis set for elements in the periodic table up to argon. For systems involving heavier elements, the B88-PW91 functional was used in conjunction with DZVP basis set.

The parameterization method 6 (PM6) has the advantage to be parameterized for the 70 elements from periodic table, which made from this method a good method applied to biochemistry systems or organic systems.

1.5.2 PM6 augmented with corrections for hydrogen bonding and for dispersion

Because of its construction and parameterization for 70 elements, the semiempirical quantum chemical PM6 method is superior to other similar methods. Despite their advantages the PM6 method fails to describe accurately noncovalent interactions specifically for the dispersion energy and for H-bonding.

Pavel Hobza et al.[22,23] propose an extension to the PM6 method to overcome this problem of non covalent interactions, augmenting the ordinary PM6 method with corrections for dispersion and for H-bonding. These corrections are made using the same approach like Grimme, by adding an empirical term for dispersion and for H-bonding. The parameters used for the correctional form of the dispersion was, for the method proposed by Hobza, taken for his previously experience with this kind of problem applied for the dispersion correction to the HF and DFT methods [24,25,26,27]. Only the parameters used in the damping function, which are the scaling radii s_r and the exponent α of the damping function, are optimized to reproduce interaction energies of dispersion of a benchmark data set. (S22 benchmark data set, with only dispersion bonded complexes)[28]. These optimizations led to $s_r = 1.07$ and $\alpha = 11$ with an average absolute value for the error function (difference between the values of the benchmark data set and the PM6-D method) equals to $e_{avg} = 0.4$ kcal/mol.

Improving hydrogen bonding interactions is not as straightforward as correcting dispersion interactions because the electrostatic term, mainly responsible for the description of H-bonding, is included in the original parameterization. To avoid the double count of this term and inspired by dispersion correction approach, Hobza et al. introduce a specific correction that would affect only hydrogen bonds, added on top of an unmodified semiempirical calculation (PM6 in this case). Figure 1.2 presents a general view of a hydrogen bonding system.

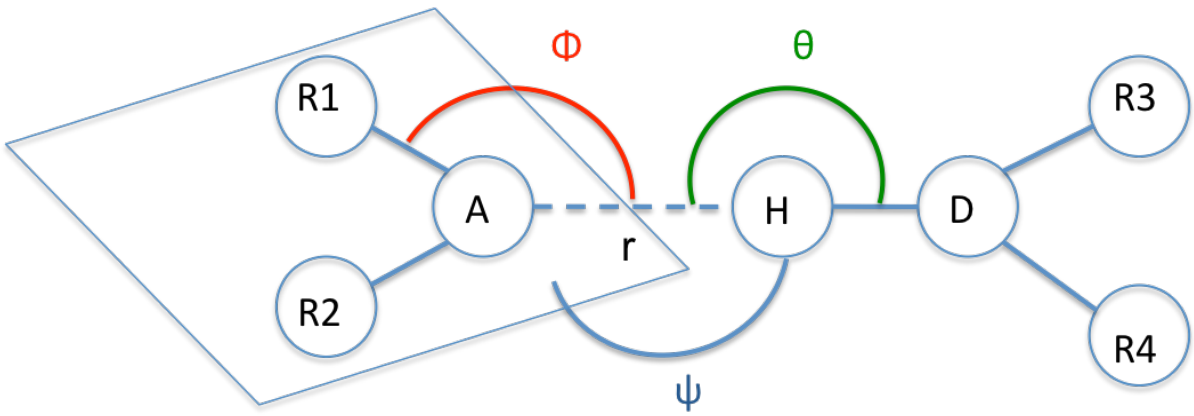


Figure 1.2General presentation for a system involved in a hydrogen bonding interaction.

The sterical arrangement of this system is fully described by the four parameters, which are:

- r the H-bond distance

- Θ angle between Acceptor … Hydrogen and Donor atoms
- ϕ angle between R_1 – A ---H atoms
- ψimproper angle between R_1R_2A plane and the H atom

The first two mentioned coordinates represent the distances between the two parts of the system, which are bind by an H-bonding and define also the directionality of this binding. The second two mentioned coordinates define the relative positions of the acceptor atom system part (the spatial arrangement of the acceptor lone pair), which is important to prevent nonphysical contributions to the H-bond interactions energies. As a result of these considerations, empirical H-bonding correction takes the following form:

$$E_{H-bond} = \left[a \frac{q_A \cdot q_H}{r^b} + c \cdot d^r \right] \cdot \cos(\theta)\cos(\phi)\cos(\psi) \quad (1.22)$$

This equation was analyzed for the importance of different terms and parameterized for acceptor atom types in a stepwise optimization process. The benchmark set was composed by 26 complexes (S22 set and four singly hydrogen bonding systems)[23]. The values of these parameters: a, b, c, d are divided in two types of parameters. One which are not dependent on the atom character (parameters: b, c, d) and the second part which includes only parameter "a", which are dependent of the acceptor atom character involved in H-bonding. The values in next table are taken from (ref 23).

Parameter	Element	Value
a	N	1.48
	O	1.56
	O_{acid}	1.55
	Opeptide	0.96
	Owater	0.76
	S	0.85
b	Not dependent	3.0
c	*ibidem*	0.65
d	*ibidem*	5.0

With this kind of approach an important problem of the PM6 semiempirical method was solved for systems that present weak interactions like for H-bondings. But a major drawback remains, for H-bonding systems, where the acceptor atom-type changes, such as for example proton transfer reactions.

1.6 CP2K suit of programs

CP2K is a suite of free available libraries of programs hosted at http://cp2k.berlios.de/ and is written in Fortran 95. The goals of this suite of programs are to perform various kinds of atomic and molecular simulations. In the last couple of years this program has been extensively developed and improved. The fact that it is continuously under development is an inconvenience and no official version is released at this moment. Because of that the manual of this collections of programs are very briefly presented. Despite this little inconveniences cp2k and more precisely the Quickstep modulus, which performs DFT calculations, is a robust and fast program able to tackle big systems. The Quickstep modulus is an hybrid Gaussian and plane waves method, which solves for big systems the SCF equations in a time that scales approximately linearly with the system size. For other codes, where for example only Gaussian basis functions are used for SCF equations, the time to solve this kind of equations increases as the basis sets and the time scale is proportional to $O(N^3)$ where N represents the number of basis functions. A second advantage of this package is that is very well parallelized and nearly scales linearly with the number of processors used.

The QS program uses an atom-centered Gaussian-type basis to describe the wave function, but also uses an auxiliary plane wave basis to describe the density. With the density represented through plane waves on a regular grid, the efficiency of Fast Fourier Transform (FFT) is exploited to solve the Poisson equations and to obtain the Hartree energy for a molecular system. The efficiency of this kind of approach is well known[29,30,31,32,33]. Also using FFT is an advantage to compute systems that are under periodic boundary conditions such as liquids and solids[34,35,36,37,38]. Besides the many advantages of CP2K collection of programs, the methods like Orbital Transformation [39] and Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases[40] increase the capabilities of this program. Indeed is very important to perform BOMD in a manner that is efficient and because of all the reasons given previously CP2K/Quickstep is an excellent candidate to perform this task.

1.7 Metadynamics

Despite the success of molecular dynamics for simulations in the fields of material science, biology, astrophysics and chemistry, this simulations method has some limitations that reduce the scope of the method and restrain its applicability. For examples, in many circumstances the free energy surface (FES) presents several local minima separated by large barriers. To overcome these barriers with an ordinary MD simulation is almost impossible because of the so called "rare events". When systems cross the barrier from one minimum to another, a huge amount of simulation time and a very favorable circumstance for the system is need, which are not the case for most of the chemical reactions. Thus, in the last years many methods have been proposed to lift these restrictions and to be able to overcome these problems[41-54]. In 2002 Alessandro Laio and Michele Parinello proposed a method, which they called Metadynamics[55], that combine ideas of coarse-grained dynamics[56,57] with those of adaptive bias potentials methods[58,59], obtaining a procedure that allows the system to escape from local minima in the FES and in the same time to have a quantitative estimation of the barrier height between the minima's.

A molecular dynamics simulation, which acts under a history-dependent external potential V_G , constructed as a sum of Gaussians along the trajectory in a subspace of real space of the system was named Metadynamics or hills method, named after the form of the Gaussian that form the history-dependent potential.

The aim of the metadynamics is to explore efficiently the Free Energy Surface (FES) discouraging the system from revisiting the same "places" during the dynamics run. To achieve this goal, a repulsive Gaussian is added at every visited "place", gradually filling the initial potential well and then pushing the system in the next well trough the lowest transition path, as it is shown in the following graph.

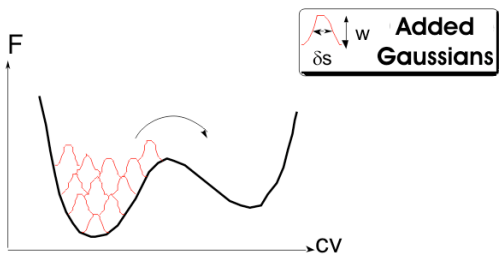


Figure 1.3Free energy surface model

Form a theoretical point of view, the essence of the metadynamics is firstly to define a relevant subset of the 3N-6 space, standing for the 3N-6 degrees of freedom of a system of N atoms. This subset consists in few collective variables (CV) s_i, i =1…d, d is the total number of CVs, generally 2 or 3. The collective variables can be any function of the atoms position such as the distances, angles, dihedral angle, coordination number, dipole moment, etc. Choosing these relevant CV for a simulation remains a delicate issue because it is done in an intuitive manner and depends how this CV can describe efficiently the reactions under study. In practice, to perform a relevant metadynamics run, one has to choose appropriate Gaussians height w, width δs, periodicity τ_G and to select a pertinent set of collective variables. Indeed, a bad choice can hinder the reaction or, even harder to detect, induce an overestimated activation barrier. This problem is illustrated by the plot of a schematic free energy surface in function of two collective variables cv1 and cv2 reported in the next graph. From a subset reduced to cv1, one would compute the purple reaction path and would overestimate the corresponding barrier. Both collective variables cv1 and cv2 are clearly needed to fully characterize the reactive path, here in orange, and then to compute the proper activation free energy.

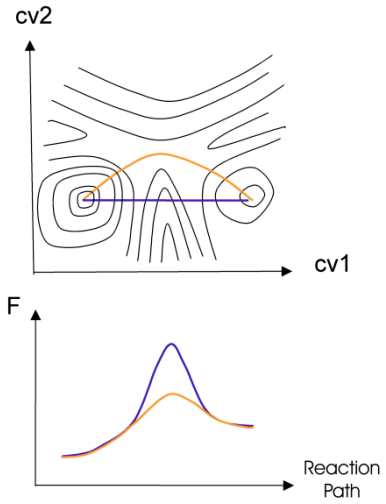


Figure 1.4 Reaction path model

Since the history dependent potential iteratively compensate the underling free energy, a system, which evolved under metadynamics, will tend to escape from any free energy minimum through the lowest free energy saddle point.

Another great advantage of this method is that we can estimate the free energy profile after metadynamic simulations. The surface of the free energy is reconstructed recursively. The authors of Metadynamics exemplify this by an intuitive example of a man, which deposits the sand in a poll at every time τ_G in the place where he is located at that time. Assuming the man is walking randomly in the poll after a while, $n \cdot \tau_G$, the sand in the poll will be sufficiently deposited and the man can escape from the poll with no difficulty to overcome the well of the poll. The negative part of the surface of the sand deposited in the poll, after the poll is full of sand and our hero escapes from them, represents exactly estimation of the surface of the pool. To make a parallel between the example of the authors and the real systems the sand is replaced by the small repulsive Gaussians and the poll is replaced by the PES in the CV subspace of the system.

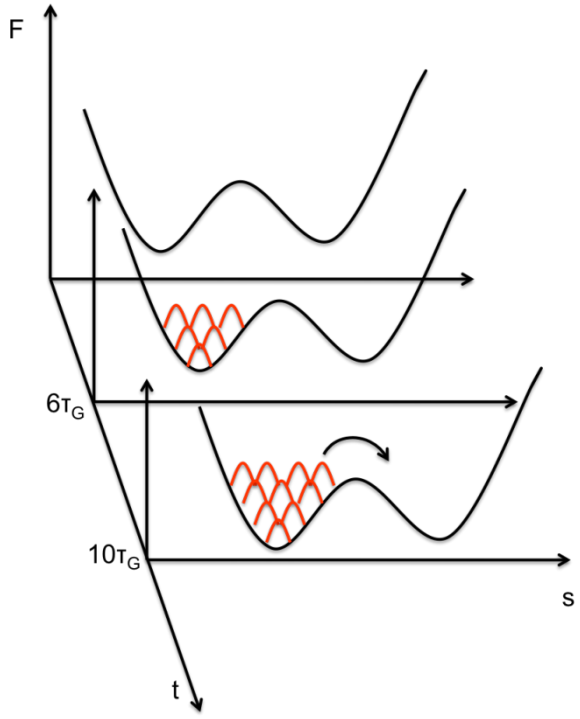


Figure 1.5 Schematic representation of the Metadynamics principle. At every τ_G period is placed a Gaussian in the CVs subspace.

The sum of the repulsive Gaussians added with a τ_G periodicity represents the history dependent potential V_G and has this expression:

$$V_G\left(s(r),t\right) = w \sum_{\substack{t'<t \\ t'=\tau_G,2\tau_G,3\tau_G\ldots}} \exp\left(-\frac{\left(s_t(r)-s_{t'}(r')\right)^2}{2(\delta s)^2}\right) \quad (1.23)$$

The parameters that enter to the definitions of the V_G are:

- w the Gaussian height
- δs the Gaussian width
- τ_G the frequency at which the Gaussians are added

These parameters influence the accuracy and efficiency of the free energy reconstruction. And the basic assumption of metadynamics is that the V_G defined in eq(1.23) after a sufficiently long time provides an estimation of the free energy[55]:

$$\lim_{t\to\infty} V_G\left(s(r),t\right) \approx -F(s) \qquad (1.24)$$

This method can be implemented almost straightforwardly in any code that performs molecular dynamics. In the routine that computes the interatomic forces, one should add a routine that performs the following operations:

- Compute the value of each CV's at the time t and stored in an array s_i^t where i=1…d represents the dimensionality of the FES.
- For every atom i of position r_i add to the ordinary thermodynamic force f_i the metadynamics force:

$$f_i = f_i - \frac{\partial V_G(s(r),t)}{\partial r} = f_i - \frac{wdt}{\tau_G}\sum_{t'<t}\sum_{j=1}^{d}\left(\frac{s_j^t(r)-s_j^{t'}(r')}{(\delta s)^2}\right)\exp\left(-\frac{\left(s_t(r)-s_{t'}(r')\right)^2}{2(\delta s)^2}\right)\frac{\partial s_j(r)}{\partial r_i} \qquad (1.25)$$

  where the last term in this equation represents the derivatives of the CVs with respect to the atomic positions. For that it requires for every collective variables to implement explicit derivative of these quantities.

To conclude, the metadynamics method has numerous and great advantages: various fields of application, its capacity to find the lowest transition state if a pertinent set of CVs has been chosen, without any assumption on the obtained product. Furthermore, its accuracy to reconstruct the free energy surface is fully determined by the system and the metadynamics parameters. The most crucial point is the choice of the collective variables subset.

1.7.1 Extensions

When we have a system, which are described simultaneously by more than two or three CVs, the time spend to fill the free energy surface in many dimensions will be computationally too expensive. Indeed it scales exponentially with the number of CV's used. The metadynamics will thus lose his purpose to accelerate the dynamics simulations. In the last couple of years to overcome this drawback, some extensions of the ordinary metadynamics simulations, were proposed. These are: Multiples walker[60], Parallel tempering metadynamics[61] and Bias exchange metadynamics[62] and all these approaches are based on running metadynamics simultaneously on several replica of the system, gathering the information from all the replicas in order to improve the estimation statistics of the free energy.

1.7.2 Multiples walkers

It is based on running multiple interacting simulations, walkers, all contributing simultaneously to construct the same metadynamics bias shared by all the walkers. Since all walkers contribute simultaneously to the history dependent potential, this leads to an algorithm that scales linearly with the number of walkers which is a great advantage for systems where the potential energy surface is reconstructed in a multi dimensional space. The history dependent potential has this form:

$$V_G\left(s(r),t\right) = w\sum_{t'<t}^{wal\,\mathrm{ker}} \sum_{i=1}^{NR} \exp\left(-\frac{\left(s_i^t(r) - s_i^{t'}(r)\right)^2}{2(\delta s)^2}\right) \quad (1.26)$$

where $s_i^t(r)$ is the trajectory of the walker "i" at a time t.

Computational point of view the only thing that the walkers must share is the list of added Gaussians that are gathered in the same file.

1.7.3 Parallel tempering metadynamics

Methods based on the choice of a set of CVs were extensively used to study protein folding. A serious problem of this method was observed and the result can be affected by the choice of the CVs and large hysteresis can be observed if a relevant CV is omitted. Bussi et al[61] propose an approach that combines Parallel Tempering and Metdaynamics. *Parallel Tempering Metadynamics* (PTM) is an approach that performs multiple Metadynamics simulations at different temperatures, periodically allowing exchange between the replicas

according with a Metropolis algorithm. For that this method has the capability to explore low probability regions of the system in a Metadynamics simulation and allows sampling of the degrees of freedom not explicitly included by the used CVs.

The PTM algorithm is quite similar to Metadynamics scheme. If we use a number of replica that are run in parallel at different temperatures after a time τ we try to exchange the coordinates of two replicas at adjacent temperatures, in the spirit of the replica exchange methods. The acceptance ratio is P and is defined by:

$P = \min\{1, f(\beta,U,V)\}$ where the $f(\beta,U,V)$ is the functional form of the replica i and j and depends by the inverse temperatures β_i and β_j, by the V_i and V_j the adaptive bias potential of the replicas and by internal energy U of replica i and j.

If the acceptance ratio P is accepted at the time τ then the coordinates are exchanged and momentum are rescaled as:

$$q'_i = q_j; p'_i = \sqrt{\frac{\beta_j}{\beta_i}} p_j$$

$$q'_j = q_i; p'_j = \sqrt{\frac{\beta_j}{\beta_i}} p_i \qquad (1.27)$$

The biases for the different temperatures tend to compensate the corresponding free energies, and the CVs start diffusing freely on a barrierless landscape.

1.7.4 Bias exchange metadynamics

Extending PTM S. Piana and A. Laio[62] introduce a powerful methodology, which allows the parallel reconstruction of the free energy of a system in a virtually unlimited number of collective variables. This approach is named *Bias Exchange Metadynamics* where multiple metadynamics simulations are performed of the same system and to the same temperature biasing each replica with a time-dependent potential constructed in a different set of collective variables.

Using Bias Exchange metadynamics we are able to biased simultaneously a large number of different collective variables and all the residual barriers orthogonal to the reaction

coordinates can be crossed in the available simulation time. The result of these simulations is not the free energy in several dimensions, but several low dimensional projections of the free energy surface along the collective variables. From this property of multidimensional nature of the method the system is capable of exploring a complex free energy landscape with high efficiency.

Because this algorithm is based on the ordinary metadynamics algorithm is easy to implement in any code of metadynamics. Following this receipt we write a Python code, which performs BE-MD interface with the CP2K code, for ordinary metadynamics. The Bias_exe.py code is fully written in Python 2.5 language and we use a modular programming technique.

Relative Bibliography to Chapter 1

(1) Fermi, E. Un Metodo Statistico per la Determinazione di alcune Prioprietà dell'Atomo. *Rend. Accad. Naz. Lincei***1927**, *6*, 602-607.

(2) L. H. Thomas The Calculation of Atomic Fields. *Mathematical Proceedings of the Cambridge Philosophical Society***1927**, *23*, 542-548.

(3) Slater, J. C. A Simplification of the Hartree-Fock Method. *Phys. Rev.***1951**, *81*, 385.

(4) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.***1964**, *136*, B864.

(5) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.***1965**, *140*, A1133.

(6) Maitland, G.; Rigby, M.; Smith, E.; Wakeham, W.; Henderson, D. Intermolecular Forces: Their Origin and Determination. *Oxford U.P., New York, 1981***1983**.

(7) Foster, M. E.; Sohlberg, K. Empirically corrected DFT and semi-empirical methods for non-bonding interactions. *Phys. Chem. Chem. Phys.***2010**, *12*, 307-322.

(8) Grimme, S.; Antony, J.; Schwabe, T.; Muck-Lichtenfeld, C. Density functional theory with dispersion corrections for supramolecular structures, aggregates, and complexes of (bio)organic molecules. *Org. Biomol. Chem.***2007**, *5*, 741.

(9) Hepburn, J.; Scoles, G.; Penco, R. A simple but reliable method for the prediction of intermolecular potentials. *Chemical Physics Letters***1975**, *36*, 451-456.

(10) Ahlrichs, R.; Penco, R.; Scoles, G. Intermolecular forces in simple systems. *Chem. Phys.***1977**, *19*, 119-130.

(11) Becke, A. D.; Johnson, E. R. A density-functional model of the dispersion interaction. *J. Chem. Phys.***2005**, *123*, 154101.

(12) Stefan Grimme Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.***2006**, *27*, 1787-1799.

(13) Johnson, E. R.; Mackie, I. D.; DiLabio, G. A. Dispersion interactions in density-functional theory. *J. Phys. Org. Chem.***2009**, *22*, 1127-1135.

(14) Mackie, I. D.; DiLabio, G. A. Interactions in Large, Polyaromatic Hydrocarbon Dimers: Application of Density Functional Theory with Dispersion Corrections. *J. Phys. Chem. A***2008**, *112*, 10968-10976.

(15) Mackie, I. D.; McClure, S. A.; DiLabio, G. A. Binding in Thiophene and Benzothiophene Dimers Investigated By Density Functional Theory with Dispersion-Correcting Potentials. *The Journal of Physical Chemistry A***2009**, *113*, 5476-5484.

(16) Pariser, R.; Parr, R. G. A Semi-Empirical Theory of the Electronic Spectra and Electronic Structure of Complex Unsaturated Molecules. I. *J. Chem. Phys.***1953**, *21*, 466.

(17) Pople, J. A. Electron interaction in unsaturated hydrocarbons. *Trans. Faraday Soc.***1953**, *49*, 1375-1385.

(18) Stewart, J. J. P. Optimization of parameters for semiempirical methods IV: extension of MNDO,

AM1, and PM3 to more main group elements. *J. Mol. Model.***2004**, *10*, 155-164.

(19) Stewart, J. J. P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J Mol Model***2007**, *13*, 1173-1213.

(20) H.Y. Afeefy; J.F. Liebman; S.E. Stein Neutral Thermochemical Data. *Linstrom PJ, Mallard, WG, Eds. NIST Chemistry WebBook, NIST Standard Reference Number 69 http://webbook.nist.gov National Institute of Standards and Technology, Gaithersburg MD, 20899***2003**.

(21) Allen, F. H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr B Struct Sci***2002**, *58*, 380-388.

(22) Ř ezáč , J.; Fanfrlík, J.; Salahub, D.; Hobza, P. Semiempirical Quantum Chemical PM6 Method Augmented by Dispersion and H-Bonding Correction Terms Reliably Describes Various Types of Noncovalent Complexes. *Journal of Chemical Theory and Computation***2009**, *5*, 1749-1760.

(23) Korth, M.; Pitoň ák, M.; Ř ezáč , J.; Hobza, P. A Transferable H-Bonding Correction for Semiempirical Quantum-Chemical Methods. *Journal of Chemical Theory and Computation***2010**, *6*, 344-352.

(24) Hobza, P.; Mulder, F.; Sandorfy, C. Quantum chemical and statistical thermodynamic investigations of anesthetic activity. 1. The interaction between chloroform, fluoroform, cyclopropane and an O-H.cntdot..cntdot..cntdot.O hydrogen bond. *J. Am. Chem. Soc.***1981**, *103*, 1360-1366.

(25) Hobza, P.; Mulder, F.; Sandorfy, C. Quantum chemical and statistical thermodynamic investigations of anesthetic activity. 2. The interaction between chloroform, fluoroform, and an NH...O:C hydrogen bond. *J. Am. Chem. Soc.***1982**, *104*, 925-928.

(26) Hobza, P.; Sandorfy, C. Quantum chemical and statistical thermodynamic investigations of anesthetic activity. 3. The interaction between CH4, CH3Cl, CH2Cl2, CHCl3, CCl4, and an O--H ... O hydrogen bond. *Can. J. Chem.***1984**, *62*, 606-609.

(27) Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. R. Density functional theory augmented with an empirical dispersion term. Interaction energies and geometries of 80 noncovalent complexes compared with ab initio quantum mechanics calculations. *J. Comput. Chem.***2007**, *28*, 555-569.

(28) Petr Jurecka; Jiri Sponer; Jiri Cerny; Pavel Hobza Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.***2006**, *8*, 1985.

(29) Fü sti-Molnár, L.; Pulay, P. The Fourier transform Coulomb method: Efficient and accurate calculation of the Coulomb operator in a Gaussian basis. *J. Chem. Phys.***2002**, *117*, 7827.

(30) Fü sti-Molnár, L. New developments in the Fourier transform Coulomb method: Efficient and accurate localization of the filtered core functions and implementation of the Coulomb energy forces. *J. Chem. Phys.***2003**, *119*, 11080.

(31) Füsti-Molnár, L.; Pulay, P. Gaussian-based first-principles calculations on large systems using the Fourier Transform Coulomb method. *Journal of Molecular Structure: THEOCHEM***2003**, *666-667*, 25-30.

(32) Baker, J.; Füsti-Molnar, L.; Pulay, P. Parallel Density Functional Theory Energies using the Fourier Transform Coulomb Method†. *The Journal of Physical Chemistry A***2004**, *108*, 3040-3047.

(33) Fü sti-Molnar, L.; Pulay, P. Accurate molecular integrals and energies using combined plane wave and Gaussian basis sets in molecular electronic structure theory. *J. Chem. Phys.***2002**, *116*, 7795.

(34) Krack, M.; Parrinello, M. All-electron ab-initio molecular dynamics. *Phys. Chem. Chem. Phys.***2000**, *2*, 2105-2112.

(35) Hura, G.; Russo, D.; Glaeser, R. M.; Head-Gordon, T.; Krack, M.; Parrinello, M. Water structure as a function of temperature from X-ray scattering experiments and ab initio molecular dynamics. *Phys. Chem. Chem. Phys.***2003**, *5*, 1981.

(36) Kuo, I. W.; Mundy, C. J. An ab Initio Molecular Dynamics Study of the Aqueous Liquid-Vapor Interface. *Science***2004**, *303*, 658-660.

(37) Kuo, I. W.; Mundy, C. J.; McGrath, M. J.; Siepmann, J. I.; VandeVondele, J.; Sprik, M.; Hutter, J.; Chen, B.; Klein, M. L.; Mohamed, F.; Krack, M.; Parrinello, M. Liquid Water from First Principles: Investigation of Different Sampling Approaches. *The Journal of Physical Chemistry B***2004**, *108*, 12990-12998.

(38) VandeVondele, J.; Mohamed, F.; Krack, M.; Hutter, J.; Sprik, M.; Parrinello, M. The influence of temperature and density functional models in ab initio molecular dynamics simulation of liquid water. *J. Chem. Phys.***2005**, *122*, 014515.

(39) VandeVondele, J.; Hutter, J. An efficient orbital transformation method for electronic structure calculations. *J. Chem. Phys.***2003**, *118*, 4365.

(40) VandeVondele, J.; Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.***2007**, *127*, 114105.

(41) Sprik, M.; Ciccotti, G. Free energy from constrained molecular dynamics. *J. Chem. Phys.***1998**, *109*, 7737.

(42) Patey, G. N.; Valleau, J. P. A Monte Carlo method for obtaining the interionic potential of mean force in ionic solution. *J. Chem. Phys.***1975**, *63*, 2334.

(43) Rodriguez-Gomez, D.; Darve, E.; Pohorille, A. Assessing the efficiency of free energy calculation methods. *J. Chem. Phys.***2004**, *120*, 3563.

(44) Darve, E.; Pohorille, A. Calculating free energies using average force. *J. Chem. Phys.***2001**, *115*, 9169.

(45) Bash, P.; Singh, U.; Brown, F.; Langridge, R.; Kollman, P. Calculation of the relative change in binding free energy of a protein-inhibitor complex. *Science***1987**, *235*, 574-576.

(46) Carter, E.; Ciccotti, G.; Hynes, J. T.; Kapral, R. Constrained reaction coordinate dynamics for the simulation of rare events. *Chem. Phys. Lett.***1989**, *156*, 472-477.

(47) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. Multidimensional free-energy calculations using the weighted histogram analysis method. *J. Comput. Chem.***1995**, *16*, 1339-1350.

(48) Ferrenberg, A. M.; Swendsen, R. H. New Monte Carlo technique for studying phase transitions. *Phys.*

*Rev. Lett.***1988**, *61*, 2635.

(49) Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.***1997**, *78*, 2690.

(50) Crooks, G. Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems. *J. Stat. Phys.***1998**, *90*, 1481-1487-1487.

(51) Rosso, L.; Mináry, P.; Zhu, Z.; Tuckerman, M. E. On the use of the adiabatic molecular dynamics technique in the calculation of free energy profiles. *J. Chem. Phys.***2002**, *116*, 4389.

(52) Grubmüller, H. Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Phys. Rev. E***1995**, *52*, 2893.

(53) Gullingsrud, J. R.; Braun, R.; Schulten, K. Reconstructing Potentials of Mean Force through Time Series Analysis of Steered Molecular Dynamics Simulations. *J. Comput. Phys.***1999**, *151*, 190-211.

(54) Roux, B. The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.***1995**, *91*, 275-282.

(55) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proceedings of the National Academy of Sciences of the United States of America***2002**, *99*, 12562-12566.

(56) Theodoropoulos, C.; Qian, Y.; Kevrekidis, I. G. "Coarse" stability and bifurcation analysis using time-steppers: A reaction-diffusion example. *Proceedings of the National Academy of Sciences of the United States of America***2000**, *97*, 9840 -9843.

(57) Kevrekidis, I. G.; Gear, C. W.; Hummer, G. Equation-free: The computer-aided analysis of complex multiscale systems. *AIChE J.***2004**, *50*, 1346-1355.

(58) Huber, T.; Torda, A. E.; Gunsteren, W. F. Local elevation: A method for improving the searching properties of molecular dynamics simulation. *J Computer-Aided Mol Des***1994**, *8*, 695-708.

(59) Wang, F.; Landau, D. P. Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States. *Phys. Rev. Lett.***2001**, *86*, 2050.

(60) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics†. *The Journal of Physical Chemistry B***2006**, *110*, 3533-3539.

(61) Bussi, G.; Gervasio, F. L.; Laio, A.; Parrinello, M. Free-Energy Landscape for β Hairpin Folding from Combined Parallel Tempering and Metadynamics. *Journal of the American Chemical Society***2006**, *128*, 13435-13441.

(62) Piana, S.; Laio, A. A Bias-Exchange Approach to Protein Folding. *The Journal of Physical Chemistry B***2007**, *111*, 4553-4559.

II Studied systems and developments

Chapter 2 NANO-MOLECULAR DEVICES CHARACTERIZED BY QUANTUM CHEMICAL APPROACH

Résume

Dans le deuxième chapitre de ce manuscrit, nous nous concentrerons sur les machines moléculaires étudiés à partir d'un point de vue théorique. Cette partie a été faite en collaboration avec un groupe expérimental, CIRE, du notre département, DCM qui a synthétisé ces structures. En particulier, les molécules étudiées dans ce chapitre ont été synthétisé par Adriana Iordache, Christophe Bucher et Eric Saint-Aman. Le domaine des nano-matériaux présente un défi pour la chimie théorique. En effet, la taille des nano-objets est assez grand pour un traitement quantique de haut niveau, mais d'autre part pas assez grand pour un traitement classique. En outre, les propriétés de ces objets sont souvent liées à processus des associations et des dissociations, qui implique interactions «faible» difficiles à être traite avec DFT[1-3]. Mais nous verrons que le nouveau développement dans la DFT, DFT-D[1], peut en effet améliorer la description de ces systèmes et la prédiction des propriétés.

2.1 Abstract

In the second chapter of this manuscript, we will focus on the molecular machines designed for nano-materials applicability studied from a computational point of view. This part was made in collaboration with an experimentally group, CIRE, in our department DCM, which synthesized these structures. In particular, Adriana Iordache, Christophe Bucher and Eric Saint-Aman have designed the molecules studied in this chapter. The quantum chemistry has a challenge task in the field of nano-materials. Indeed, the size of the nano objects is quite big for a high level quantum treatment but on the other hand not big enough for a classical treatment. Moreover the properties of these objects are often linked with associations and dissociations processes, which involved "weak" interactions difficult to treat at the DFT level[1-3] . But we will see that the newly developed DFT, DFT-D[1], can indeed improve the description of these systems and the prediction of the properties.

This chapter presents:

- Viologene-based redox switchable receptors.
- Nature and importance of the interactions in the substituent group of 1,1'-dimethyl-4,4'-bipyridinium units. The role of this substituent group in stabilization energy of the complex.
- Molecular rotors based on Ferocene-bis-pyridinium units.

2.2 Viologene-based redox switchable receptors studied by quantum chemistry

2.2.1 Introduction

The design of artificial receptors capable of recognizing/sensing anionic species in polar media still remains a challenging task, with many potential applications in biology, medicine or environmental science[4-6]. The anion-binding properties of polyammonium derivatives are particularly well recognized, and one finds numerous examples of saturated polyammonium compounds efficiently complexing anionic species in highly competitive media[7]. Their unusual anion-binding properties arise from favorable combinations of structurating effects, electrostatic forces and hydrogen bonds but the lack of intrinsic electrochemical or photochemical read-out signals usually requires computational calculations. Ammoniums

deriving from conjugated heterocycles such as pyridine, imidazole, pyrimidine, pyrazine or purine[5],while keeping their efficient anion-binding capabilities, may conversely exhibit well defined photochemical and electrochemical signatures which could be advantageously used as transduction signals of anion-binding processes. Among this class of molecules, viologens have attracted considerable interest due to their wide range of properties finding potential applications in analytical chemistry, in solar energy conversion or molecular electronics[8].

In the present part of this study, we are interested in the study of 1,1'-dimethyl-4,4'-bipyridinium dimer linked by an propyle linker (abbreviated $[\mathbf{2}^{2+}]_{Dim}$ in reduced form and $[\mathbf{2}^{4+}]_{Dim}$ for non reduced form) in order to obtain information about the structure at different oxidation states. The main goal of this study is to gain insights into the geometry of the π-dimerized bis-viologen species $[\mathbf{2}^{2+}]_{Dim}$ and into the dynamical process. Theoretical calculations were undertaken using various levels of calculation, namely DFT, DFT-D, DFT-CP and MP2.

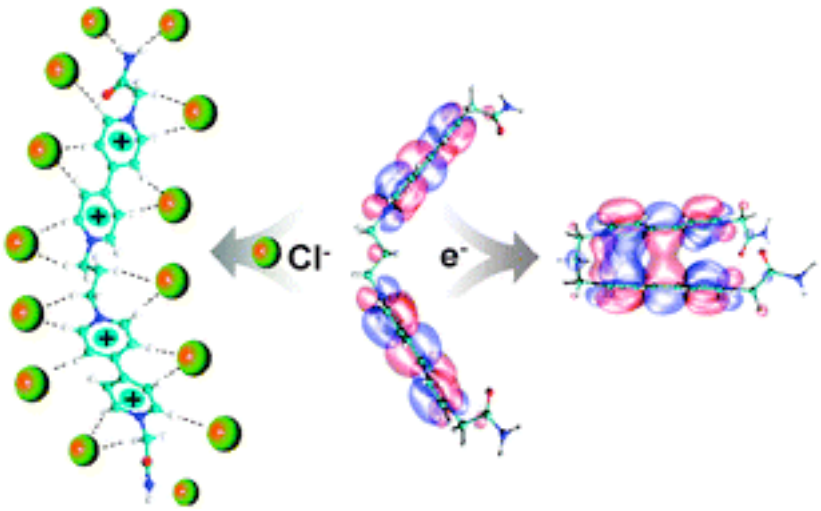


Figure 2.1Model of receptors proposed to recognizing/sensing Chloride anions.

The π-dimer $[\mathbf{2}^{2+}]_{Dim}$ is expected to develop intramolecular interactions mostly driven by dispersion energy. This contribution is computationally demanding and if MP2 is well known to give a proper description of the dispersion energy, it is also well known that a large basis set is mandatory. In order to include this dispersion energy at a lower cost, we also used the DFT-D[1] and DFT-CP[9] levels of calculations. These levels have already been used for large molecular systems[10,11] with an important contribution of the dispersion energy. For comparison purpose, DFT calculations using the BLYP and B3LYP functionals were also

performed.

2.2.2 Computational details

The quantum chemistry calculations were carried out with both the CP2KQuickStep[12] program and the Gaussian03[13]program. With this latter code, we used the BLYP[14,15], B3LYP[16], B3LYP-CP[17] (Corrected Potential) functional and MP2[18]with the 6-31+G(d,p)basis sets.

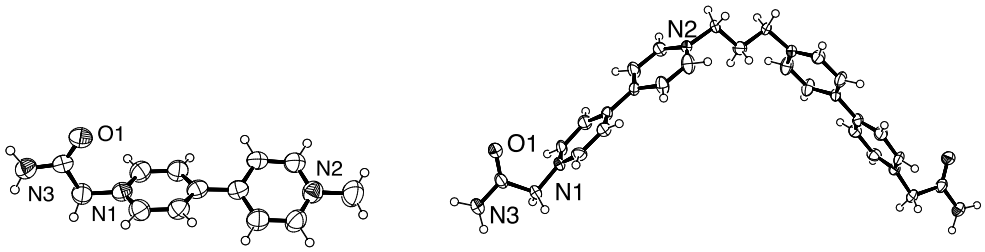
Solvent effects have been evaluated using the polarizable continuum model (PCM)[19-22]in which the cavity is created via a series of overlapping spheres.

The *ab initio* Born-Oppenheimer dynamics calculations were performed using the CP2K-QuickStep programat the DFT level with the BLYP - D[1]functional. The basis set used was a double-ζvalence set[23] of Gaussian orbitalsin conjunction with the Goedecker-Teter-Hutter pseudopotentials[24-26]. The auxiliary PW basis set was defined by a cubic box of 25 $Å^3$and by a density cutoff of 360 Ry.

Metadynamics has been used to overcome the problem of observing rare events in conventional molecular dynamics and determining the energy barrier of our system. A time step of 1 fs is used for the dynamics, and the deposition rate of the hills is every τ_G=20 fs. The height (w) and with (δs) of the Gaussians-shaped potential hills added are 0.6 kcal/mol and 1 Bohr respectively.

2.2.3 Non-reduced 1,1'-dimethyl-4,4'-bipyridinium molecule

We focused primarily on the non-reduced starting material $[\mathbf{2}^{4+}]_{Dim}$, whose structure could be obtained through geometry optimization. In agreement with the existence of electrostatic repulsion forces driving doubly charged viologens away from each other, our calculations lead invariantly to an open structure. These species are also characterized at the solid state by X-ray diffraction analysis in the presence of chloride or iodide anions (see Figure 2.2)[27].



$\mathbf{1}^{2+}$ $[\mathbf{2}^{4+}]_{dim}$

Figure 2.2Ortep views of 1^{2+} and $[2^{4+}]_{dim}$structures

The agreement of our calculations with these X-Ray data is good (Table2.1). We found that the amplitude of the torsion angle γ C7-C8-C11-C12 (Table2.1), defining the tilt angle between pyridines in each viologen fragment, depends on the counter anion's nature. We also found that the torsional angle γ observed for the optimized structure of $[\mathbf{2}^{4+}]_{dim}$ in absence of counter-anions is close to the one measured experimentally (~31°) on the solid state structure of $\mathbf{1}\cdot(PF_6)_2$**.** The same angle estimated theoretically with $\mathbf{2}^{4+}$ in interaction with two iodine anions is conversely matching the angle measured (~13°) from the X-Ray structure of the bis-viologen derivative [**2**(DMSO)$(I)_4$]. These differences point out that the tilt angle is higher with non interacting anions as PF_6 but the CT process established with iodide significantly reduces it. Moreover the comparison between $[\mathbf{2}^{4+}]_{dim}$ and the simple *N*-methylated derivative $\mathbf{1}^{2+}$ demonstrates that the amide terminal group has no influence on the geometry of the viologen units [28]. The singlet state of the bis viologen $[\mathbf{2}^{4+}]_{dim}$ was found to be the ground state of the molecule since the singlet state is 61.8 or 58.6 kcal/mol lower than the triplet state at the B3LYP/6-31+G** and the B3LYP-CP/6-31+G**, respectively.

2.2.4 Doubly reduced 1,1'-dimethyl-4,4'-bipyridinium species

The theoretical description of the doubly reduced $[\mathbf{2}^{2+}]$ species is trickier. Both the open $[\mathbf{2}^{2+}]_{open}$ and the closed π-dimer $[\mathbf{2}^{2+}]_{dim}$ forms can be characterized and frequency calculations confirm that both of them are minima on the PES (Potential Energy Surface). Moreover the singlet and triplet spin states have been investigated for each form.[2]

[2] Broken symmetry calculations have also been attempted with or without nonsymmetrical starting geometry, but the symmetrical singlet state was always found to be the converged state.

Table2.1 Comparison between structural parameters defining 2^{4+}obtained by theoretical calculation and X-ray diffraction analyses. Distances and angles are given in Å and degrees, respectively**.**

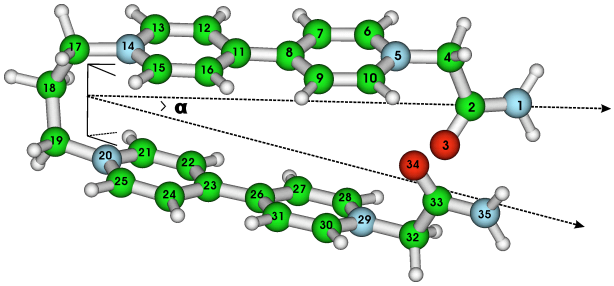
	X-ray Diffraction data		B3LYP 6-31+G**	B3LYP-CP 6-31+G**	B3LYP 6-31+G**	PBE0 6-311G** (ref.[28])
	1$(PF_6)_2$	[**2** (DMSO)$(I)_4$]	$[\mathbf{2}^{4+}]_{open}$	$[\mathbf{2}^{4+}]_{open}$	$[\mathbf{2}^{4+}]_{open}\cdot(I^-)_2$	$\mathbf{4}^{2+}$
N1-C2	1.315	1.347	1.339	1.338	1.351	-
C2-O3	1.220	1.227	1.224	1.225	1.221	-
C2-C4	1.518	1.503	1.564	1.559	1.551	-
C4-N5	1.460	1.449	1.477	1.470	1.470	1.491
N5-C6	1.349	1.329	1.354	1.351	1.364	1.351
C6-C7	1.381	1.385	1.389	1.387	1.371	1.382
C7-C8	1.376	1.367	1.401	1.396	1.418	1.401
C8-C11	1.481	1.493	1.491	1.480	1.487	1.488
C11-C12	1.404	1.346	1.409	1.404	1.426	-
C12-C13	1.382	1.387	1.384	1.381	1.369	-
C13-N14	1.318	1.335	1.359	1.355	1.366	-
N14-C17	1.478	1.474	1.498	1.490	1.475	-
C17-C18-C19	-	112.4	108.1	107.8	110.9	-
C7-C8-C11-C12	31.0	13.0	38.6	38.1	11.1	43.0

In the optimized open structure $[\mathbf{2}^{2+}]_{open}$, we found that the torsion angle γ measured between the two pyridine fragments, (C7-C8-C11-C12) in each viologen was around 5.0 degrees whatever the spin states being considered. This small value, defining a nearly coplanar arrangement, is in accordance with our assumption that the tilt angle γ strongly decreases when the electron density on the viologen increases, as seen by electrochemical reduction or through the formation of a CT complex with electron rich iodide anions. [29]

The singlet and triplet states of the closed π-dimer structures $[\mathbf{2}^{2+}]_{dim}$show significant structural differences. The singlet state closed-form $^{S}[\mathbf{2}^{2+}]_{dim}$exhibits a geometry wherein both viologen units are almost perfectly parallel (Table2.2, $\alpha \sim 4\text{-}5°$), whereas a significantly more open structure was found in the triplet state $^{T}[\mathbf{2}^{2+}]_{dim}(\alpha \sim 15°)$ (see Table2.2). Due to the

important size of the viologen units, this 10 degrees deviation measured between the triplet and singlet states leads to a difference of 1 Å in the N5-N29 distance and to a smaller shift of 0.2 Å for the N14-N20 one (Table2.2).

Table2.2 Geometrical parameters calculated for the reduced π-dimer $[2^{2+}]_{dim}$ conformations in the singlet ($^{S}[2^{2+}]_{dim}$) or triplet state ($^{T}[2^{2+}]_{dim}$) using the B3LYP/6-31+G or B3LYP+CP/6-31+G** calculation levels. The distances are given in Å; angles are given in degrees. The optimized structure shown corresponds to the closed conformation found for the triplet state $^{T}[2^{2+}]_{dim}$.**



	B3LYP/6-31+G**		B3LYP-CP/6-31+G**	
	$^{S}[2^{2+}]_{dim}$	$^{T}[2^{2+}]_{dim}$	$^{S}[2^{2+}]_{dim}$	$^{T}[2^{2+}]_{dim}$
C4-C5	1.464	1.460	1.457	1.455
N5-C6	1.369	1.373	1.363	1.366
C6-C7	1.369	1.367	1.365	1.365
C7-C8	1.431	1.433	1.424	1.426
C8-C11	1.436	1.433	1.426	1.426
C23-C26	1.436	1.433	1.427	1.425
N14-N20	3.007	3.296	2.891	3.041
N5-N29	4.153	5.556	3.678	4.589
C7-C8-C11-C12	5.3	6.1	0.9	3.3
α	5.1	18.0	5.2	13.1

The molecular orbital analysis of the singlet state of $\mathbf{2}^{2+}$ in the closed and open conformations are shown on Figure 2.3. The LUMO (Lowest Unoccupied Molecular Orbital) of the open starting material $\mathbf{2}^{4+}$ becomes the HOMO (Highest Occupied Molecular Orbital) of the doubly reduced $\mathbf{2}^{2+}$ species with a bonding character associated to the C8-C11 (C23-C26) bonds. The same bonding character was observed for the C7-C6 (C9-C10) and C13-C12 (C15-C16) bonds, whereas an antibonding character is associated to the C7-C8 (C8-C9), N5-C6 (N5-C10) and N14-C13 (N14-C15) bonds. The electron doping process has thus significant consequences on the geometry of the viologen units as shown by the differences observed between the bond distances tabulated in Table 2.1 and Table 2.2. Reducing $\mathbf{2}^{4+}$ into $\mathbf{2}^{2+}$ led's especially to the significant bond lengths increases for N5-C6 (N5-C10), C7-C8 (C9-C8) and to the concomitant decreases of the C8-C11, C6-C7 (C9-C10) bond distances. As seen onFigure 2.3, the HOMO of $\mathbf{2}^{2+}$ is also clearly responsible for the association between both viologen fragment forming the closed π-dimer $[\mathbf{2}^{2+}]_{dim}$ species.

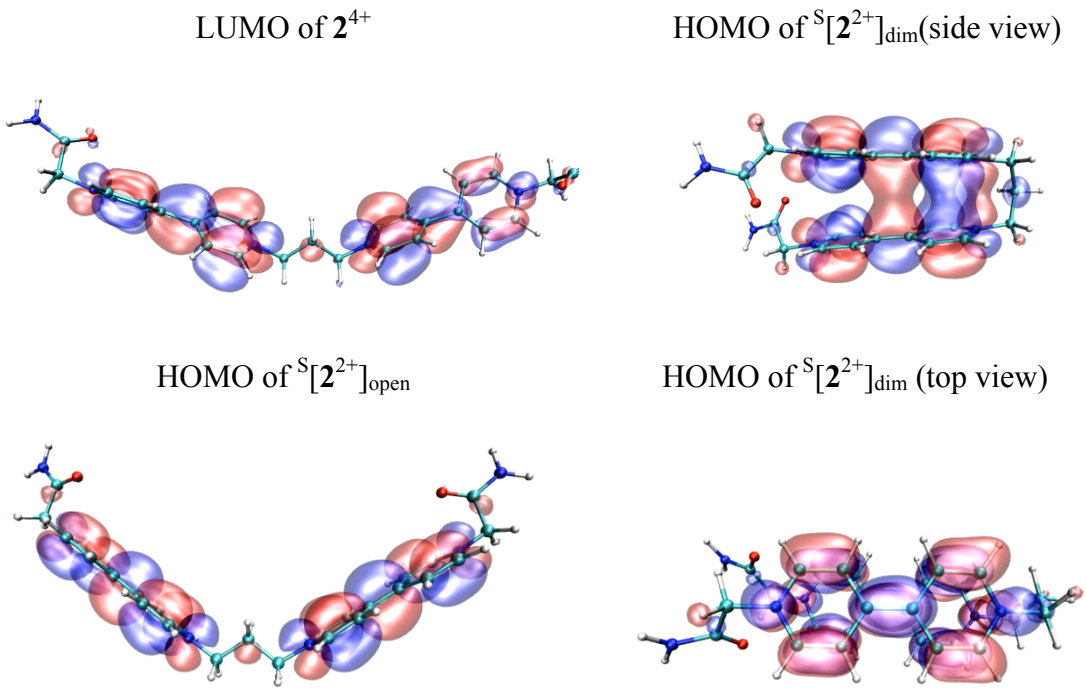


Figure 2.3Selected molecular orbitals obtained for the singlet states of $\mathbf{2}^{4+}$ and $\mathbf{2}^{2+}$ in closed ($^{S}[\mathbf{2}^{2+}]_{dim}$) and open ($^{S}[\mathbf{2}^{2+}]_{open}$) conformations. The electron density isosurface is 0.02a.u.

From an energetically point of view, the difference in energy between the open and π-dimer forms of the doubly reduced $\mathbf{2}^{2+}$ species proved to be highly sensitive to the chosen level of calculations (Table2.3). A previously reported ESR investigation conducted with a simple bisalkylated-viologen[30] V^{2+}, established that the singlet state of the reduced form $^{S}[V^{\bullet+}]$ is more stable at low temperature (around 100K), whereas the triplet state $^{T}[V^{\bullet+}]$ was only detected at room temperature. From these experiments, it was postulated that the singlet state π-dimer $^{S}[(V)_2{}^{2+}]_{Dim}$ is more stable at low temperature. On the contrary, the spectrum recorded at room temperature confirmed the presence of unpaired electrons attributed to $[V]^{\bullet+}$ with a weak contribution corresponding to the triplet dimer $^{T}[(V_2)^{2+}]_{dim}$. Similar behaviours were observed with various viologen-based radicals[31] becoming ESR silent at low temperatures. Such temperature dependant loss of ESR signal was also frequently interpreted as a diagnostic signature of the π-dimerization process assuming the formation of bis-radical spin-paired dimers at low temperatures and of non-associated radicals at RT (room temperature). In the case of $\mathbf{2}^{4+}$, the situation is much more complex since the dimerization is an intramolecular process wherein the opening of the singlet state dimerized closed structure $^{S}[\mathbf{2}^{2+}]_{dim}$ could lead to electron unpairing to produce the $^{T}[\mathbf{2}^{2+}]_{open}$ or could keep its spin state through the opening process to generate an ESR silent $^{S}[\mathbf{2}^{2+}]_{open}$ species.

At the widely used B3LYP/6-31+G** calculation level, the closed conformation in the singlet state $^{S}[\mathbf{2}^{2+}]_{dim}$ is slightly more stable than its corresponding open singlet structure $^{S}[\mathbf{2}^{2+}]_{open}$ with a stabilization of 0.9 kcal/mol. As seen in values reported in Table2.3, the inclusion of dispersion energy, through various methods involving dispersion terms or corrected potentials, clearly favors the closed conformation. The inclusion of water as solvent through PCM models[19,20,21,22] also turned out to favor the closed form, although at a lesser extent. On the contrary, the BLYP functional proved to favor the open structure with energy differences of 6.6 kcal/mol and 8.9 kcal/mol with the 6-31+G** and DZVP-GTH basis sets, respectively. But the inclusion of dispersion energy at the BLYP-D/DZVP-GTHlevelalso led to the stabilization of the π-dimer form with an energy difference of 12.1 kcal/mol.

Table2.3Relative energy differences between the singlet (S) and triplet (T) states of the closed and open conformation calculated for 2^{2+}. The closed $[2^{2+}]_{dim}$conformation is favored when $\Delta E < 0$

Method	ΔE [kcal/mol]	Spin State
B3LYP/6-31+G**	-0.9	S
B3LYP/6-31+G**	+6.6	T
B3LYP-CP/6-31+G**	-21.7	S
B3LYP-CP/6-31+G**	-3.2	T
B3LYP-CP/6-31+G** + water	-25.0	S
B3LYP/6-31+G** + water	-2.8	S
B3LYP/6-31+G** + acetonitrile	-3.4	S
MP2/6-31+G**//B3LYP-CP/6-31+G**	-9.5	S
PBE/DZVP-GTH	-3.4	S
PBE-D/DZVP-GTH	-23.0	S
PBE/DZVP-GTH	+7.2	T
PBE-D/DZVP-GTH	-1.8	T
BLYP/6-31+G(d,p)	+6.6	S
BLYP/DZVP-GTH	+8.9	S
BLYP-D/DZVP-GTH	-12.1	S

As expected, the triplet state systematically favored the open structure. For example, the open structure was found to be more stable by 6.6 kcal/mol and 7.2 kcal/mol at the B3LYP/6-31+G** and PBE/DZVP levels of calculation, respectively. The inclusion of dispersion contributions with corrected potentials also led to the stabilization of the closed conformation by 3.2 kcal/mol and 1.8 kcal/mol with B3LYP-CP/6-31G** and PBE-D/DZVP, respectively. Finally, it should be noted that at the B3LYP/6-31+G** level, the triplet state turned out to be more stable in both the open and closed conformations whereas at the B3LYP-CP/6-31+G** level, the singlet state of the closed conformation was the more stable, in agreement with the ESR experimental data.

2.2.5 Metadynamics results

Using metadynamics[32,33,34] with the N-N distances as collective variables, we have been able to estimate the free energy barrier required to switch the closed singlet state $[\mathbf{2}^{2+}]_{Dim}$into the open non associated structure at around 8 kcal/mol at 300K at the BLYP-D/ DZVP-GTH calculation level.

We considered distances R1 and R2 as collective variables. The definition of R1 and R2 can be seen in Figure 2.4. Velocity rescaling algorithm was used to enforce the system to a given temperature T of 300 Kelvin. The hydrogen atoms are changed to deuterium in conjunction with integration step of 1 fs.

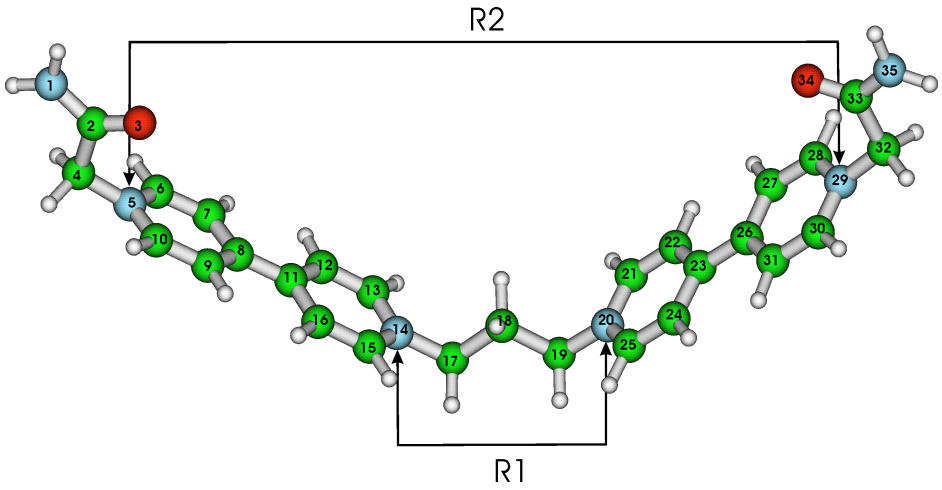


Figure 2.4Collective variable R1 and R2, which represents the distances between the nitrogen atoms.

The duration of our simulation is approximately 5 ps of Metadynamics and after 3 ps the system starts to opening. This is clearly emphasized by the variation of the collective variable R1 showed in Figure 2.5.

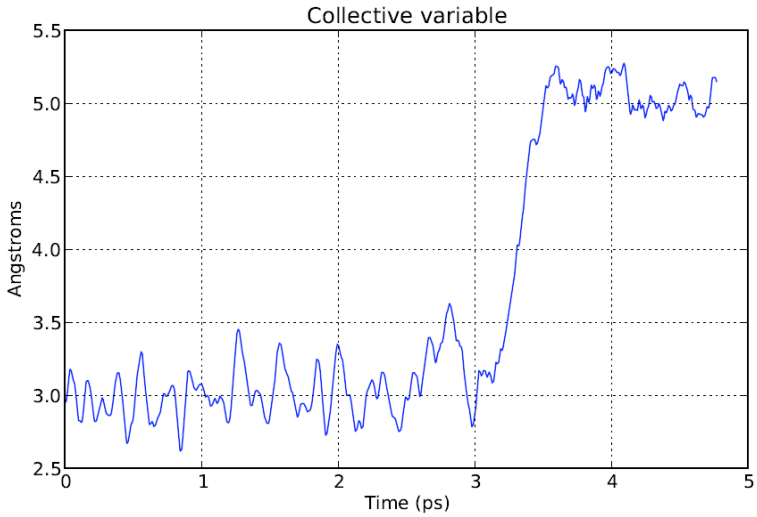


Figure 2.5Timeevolution of collective variable R1

The free energy surface F(S(R)) is then reconstructed on a 100 × 100 grid in the (1.052Å < R1 < 6.843Å) × (1.362Å < R2<16.560Å) region with an integration step of 0.111Å and 0.290Å respectively see Figure 2.6.

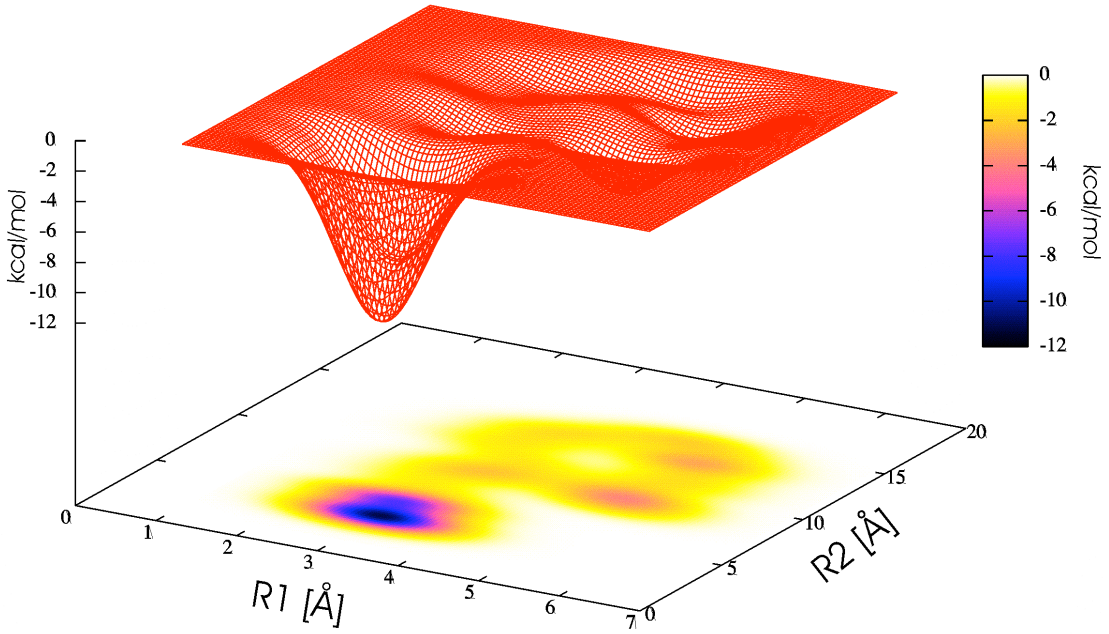


Figure 2.6 Free energy surface reconstructions from Metadynamics run at T=300K with the bottom representations of the Free Energy Surface depth

From the reconstructed FES, we can clearly see a well-defined potential well around 3Å for the R1 and R2 with a depth of 11 kcal/molcorresponding to the closed conformation. In a second region of the FES around R1=[3.5Å-6Å], R2=[5Å-15Å], we can see more minima potential wells with a relative depth of 4 kcal/mol. They correspond to the open conformation and the presence of several minima indicates the flexibility of the system in this conformation.

2.3 Nature and importance of the interactions in the substituent group of 1,1'-dimethyl-4,4'-bipyridinium units for stabilization energy of the complex.

In the precedent part of this chapter we characterized the 1,1'-dimethyl-4,4'-bipyridinium dimer linked by a propyle linker with an amide as terminal group of bipyridinium units. These kind of systems by the fact that are positively charged, electron deficient, are well known for their key role in binding and/or signalizing elements in artificial molecular receptors[35,36]. But another important property of this class of object is their capability, under reduction, to form the dimer and in our case to switch from open conformation to closed conformations. When reduced the two conformations, open and closed, can be characterized by calculations but it is clear that the terminal function used play a role in the relative stability of the conformations. This relative stability can be important to use the bis-viologen

as nano-objects. Thus using quantum chemical approach, we characterized these systems replacing the amide group by different substitutes in order to propose new class of units, which can be used to tuning the desired properties of these new compounds.

From a general point of view, we consider units linked to the Nitrogen atom in the bis-viologene molecular unit. We used the abbreviation $^{S}[(Me\text{-}Me)^{2+}]_{dim}$ if the ligand units are Methyl for both ends of the 1,1'-dimethyl-4,4'-bipyridinium dimer. Figure 2.7 presents this model used in this study where the unit A and B are the units presented in Table2.4.

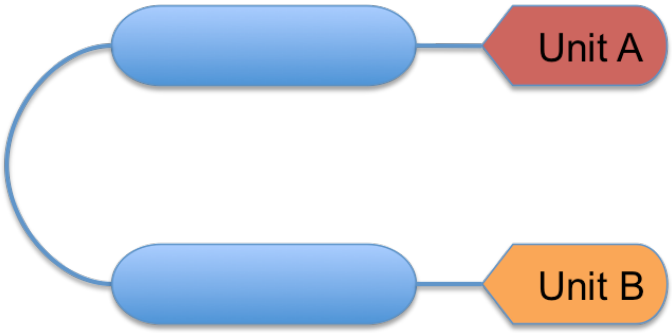
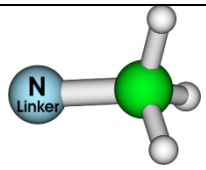
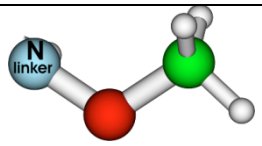
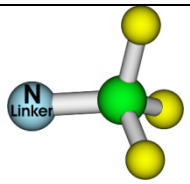
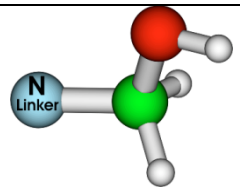
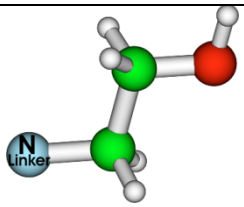
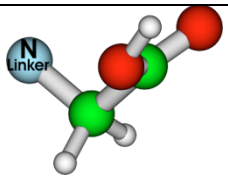
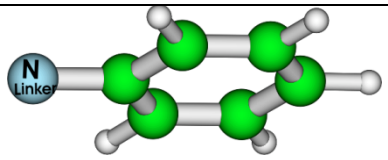
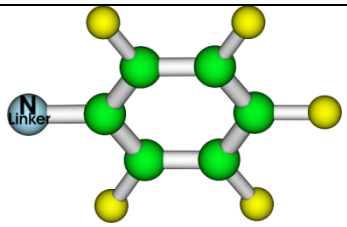


Figure 2.7Scheme, which highlights the model used.

Table2.4 Schematically representations of the terminal units linked by the 1,1'-dimethyl-4,4'-bipyridinium dimer**.**

Terminal unit name (abbreviation)	Representation
N-methyl (Me)	
N-methyl ether (OMe)	
N-tri-fluro methyl (CF3)	

N-methanol (CH2OH)	
N-ethanol (2CH2OH)	
N-carboxylic acid (COOH)	
N-phenyl (Ph)	
N-penta-fluoro phenyl (PhF)	

We used different substituents, which cover different type of interactions: no obvious interaction with substituent like methyl or even interaction potentially preventing the interaction like with ether methyl substituent, H-bonded systems with methanol, ethanol (ethanol substituent is easier to synthetise) and of course ethanoic acid. With the phenyl and penta-fluoro substituents, the interaction based on weak interactions like dispersion will be investigated.

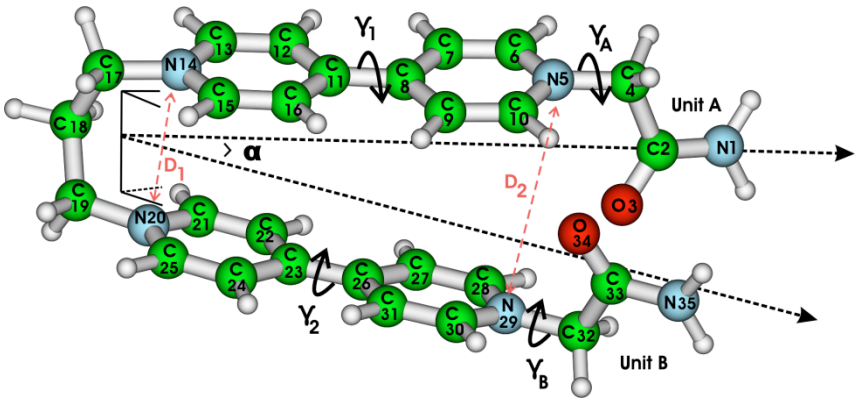
The theoretical description of doubly reduced species of 1,1'-dimethyl-4,4'-bipyridinium dimer linked by an propyle linker with different terminals groups follows the same procedure that the one described previously when we studied the system with amide group as terminal group for the 1,1'-dimethyl-4,4'-bipyridinium dimer. Our calculations are restricted to the

singlet state since Adriana Iordache has shown using ESR techniques at 100 K and 298 K that $[2^{2+}]_{dim}$ exhibit no signal, which is consistent with a singlet state.

Each substituent group was chosen by the criteria of the interactions that can exhibit this kind of systems (e.g. H-Bonding and/or dispersion interactions) to be able to tuning the kind of interaction neededfor specific properties, which need or not to have a more or less stable closed form.

From a geometrical point of view, the open form of the reduced π-dimer was quite similar independently on the substituent used. Likewise, the geometry of the open form is not dependent on the fact that the method used included or not the long range corrections.

Table2.5. Geometrical parameters calculated for the reduced π-dimer $^{S}[(UnitA\text{-}UnitB)^{2+}]_{dim}$ models with different methods. For all these methods 6-31+G** basis set was used. The distances are given in Å; angles are given in degrees. Definitions of the geometrical parameters are highlight in the figure**.**



Substituent/Method	D_1	D_2	α	γ_1	γ_2	γ_A	γ_B
$^{S}[\mathbf{2}^{2+}]_{dim}$ / B3LYP	3.007	4.153	5.1	5.3	3.5	126.4	126.8
$^{S}[\mathbf{2}^{2+}]_{dim}$ / B3LYP-D	2.946	3.747	5.2	1.3	0.4	127.0	122.0
$^{S}[(CF3\text{-}CH2OH)^{2+}]_{dim}$ / B3LYP	3.023	4.626	0.6	3.2	0.1	-	136.1
$^{S}[(Ome\text{-}Ome)]_{dim}$ / B3LYP	3.038	8.373	0.2	1.4	1.9	93.7	94.4
$^{S}[(CH2OH\text{-}CH2OH)]_{dim}$ / B3LYP	3.015	4.376	1.0	0.2	1.2	109.7	109.7
$^{S}[(2CH2OH\text{-}2CH2OH)]_{dim}$ / B3LYP	3.010	6.072	0.6	3.9	0.3	96.5	98.1
$^{S}[(CF3\text{-}CF3)]_{dim}$ / B3LYP	3.065	8.163	0.3	0.4	0.8	-	-

$^{S}[(PhF\text{-}Ph)]_{dim}$ / B3LYP	3.033	8.079	0.3	0.7	3.0	64.9	49.9
$^{S}[(PhF\text{-}Ph)]_{dim}$ / B3LYP-D	2.925	3.454	0.0	4.4	1.1	135.4	140.7
$^{S}[(Ph\text{-}Ph)]_{dim}$/ B3LYP	3.023	8.388	0.0	7.1	4.5	47.0	49.0
$^{S}[(Ph\text{-}Ph)]_{dim}$/ CAM-B3LYP	2.923	4.567	0.0	4.5	4.5	46.7	46.7
$^{S}[(Ph\text{-}Ph)]_{dim}$/ B3LYP-D	2.912	3.891	6.1	2.9	2.2	34.0	44.0
$^{S}[(COOH\text{-}COOH)]_{dim}$ / B3LYP	3.027	6.353	0.0	3.0	0.0	84.4	84.4
$^{S}[(COOH\text{-}COOH)]_{dim}$ / B3LYP-D	2.866	5.801	0.3	1.4	4.4	87.7	83.6
$^{S}[(Me\text{-}Me)]_{dim}$ / B3LYP	3.023	8.479	0.0	0.2	0.0	-	-
$^{S}[(Me\text{-}Me)]_{dim}$ / B3LYP-CP	2.897	3.552	0.0	1.2	1.2	-	-
$^{S}[(Me\text{-}Me)]_{dim}$ / B3LYP + acetonitrile	3.014	4.022	0.4	0.8	0.5	-	-

On the other hand, the closed π-dimer structures proposed with different linkers and calculated with different methods show significant structural differences. For the substituents, which are expecting to exhibit interactions with a strong importance of the dispersion, corrections have been added to the B3LYP functional, mainly corrections developed by Grimme[1]. For the methyl terminal group, we also used the correction developed by DiLabio[17], which has been designed for such systems.

From a general point of view, when the methods with dispersion corrections included are used, we obtain a geometry wherein both viologen units are much more parallel (Table2.5, $\alpha \approx 0°$ and the distance D1 roughly equals to D2) which is not the case when we use B3LYP without corrections (Table2.5, $\alpha \approx 0°$ and the distance D1<D2). This is a result of the fact that dispersion energy contribution to the stabilization energy is quite important. B3LYP without correction is not a suitable level of theory to describe the geometry of the systems where dispersion is expected to be strong. On the other hand, when strong H-bonds are mainly responsible of the interaction for the acid carboxylic terminal group, the inclusion of the Grimme correction does not lead to a strong modification of the geometry. This result seems consistent with a strong H-bond between the carboxylic group since electrostatic is an important part of this type of interaction and thus can be well described by the B3LYP functional. We can also note that for the phenyl terminal group, for which the Grimme

corrections and the CAM correction were used, the Grimme correction seems to overestimate the interaction compared the CAM-B3LYP results. This trend of the Grimme correction has already been observed[37,38] and indeed the Grimme correction is expected to slightly over-estimate the long range corrections.

The inclusion effect for $[(Me\text{-}Me)]_{dim}$ leads also to a strong modification of the geometry in favor of the closed form.

Table2.6. Relative energy differences between the closed and open conformations calculated for $^{S}[(UnitA\text{-}UnitB)^{2+}]_{dim}$ with different methods; the closed conformation is favored when $\Delta E<0$. For all these methods 6-31+G** basis set was used**.**

Complex/Method	ΔE [kcal/mol]
$^{S}[\mathbf{2}^{+}]_{dim}$ / B3LYP	-0.90
$^{S}[\mathbf{2}^{+}]_{dim}$ / B3LYP-D	-21.31
$^{S}[\mathbf{2}^{+}]_{dim}$ / CAM-B3LYP	-16.30
$^{S}[(CF3\text{-}CH2OH)^{2+}]_{dim}$ / B3LYP	+6.53
$^{S}[(Ome\text{-}Ome)]_{dim}$ / B3LYP	+5.85
$^{S}[(CH2OH\text{-}CH2OH)]_{dim}$ / B3LYP	+1.47
$^{S}[(2CH2OH\text{-}2CH2OH)]_{dim}$ / B3LYP	+2.99
$^{S}[(CF3\text{-}CF3)]_{dim}$ / B3LYP	+6.67
$^{S}[(PhF\text{-}Ph)]_{dim}$ / B3LYP	+5.49
$^{S}[(PhF\text{-}Ph)]_{dim}$ / B3LYP-D	-13.16
$^{S}[(Ph\text{-}Ph)]_{dim}$/ B3LYP	+5.41
$^{S}[(Ph\text{-}Ph)]_{dim}$/ CAM-B3LYP	-4.64
$^{S}[(Ph\text{-}Ph)]_{dim}$/ B3LYP-D	-8.12
$^{S}[(Ph\text{-}Ph)]_{dim}$/B3LYP+acertonitrile// B3LYP-D	-1.81
$^{S}[(COOH\text{-}COOH)]_{dim}$ / B3LYP	0.00
$^{S}[(COOH\text{-}COOH)]_{dim}$ / B3LYP-D	-11.54
$^{S}[(Me\text{-}Me)]_{dim}$ / B3LYP	+5.80
$^{S}[(Me\text{-}Me)]_{dim}$ / B3LYP-D	-6.18
$^{S}[(Me\text{-}Me)]_{dim}$ / B3LYP-CP	-7.63
$^{S}[(Me\text{-}Me)]_{dim}$ / B3LYP + acetonitrile	-6.42

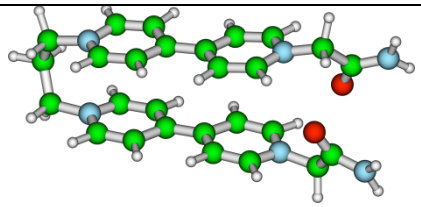
The energy difference between the closed and open forms of the reduced π-dimer, are mainly in accordance with the difference found for the geometry: the smaller the D2 distance is, the more stable is the closed form compared to the open conformation. Thus, the inclusion of

long range corrections can lead to a difference of energy as high as 18.5 kcal/mol in the case of $[(PhF\text{-}Ph)]_{dim}$. Similarly, the inclusion of the solvent effect (acetonitrile ($\varepsilon = 35.688$), which is the experimental solvent for $[2^{+}]_{dim}$) for $^{S}[(Me\text{-}Me)]_{dim}$favors strongly the closed form by 12 kcal/mol. On the other hand, for $[(COOH\text{-}COOH)]_{dim}$the energy difference with and without correction is around 11.5 kcal/mol which is quite a big value close to the value of 13.5 kcal/mol found for $[(Ph\text{-}Ph)]_{dim}$ whereas for this latter the difference in the geometry was clearly more impressive.

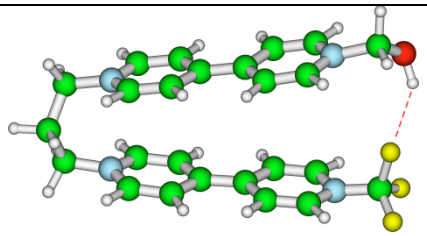
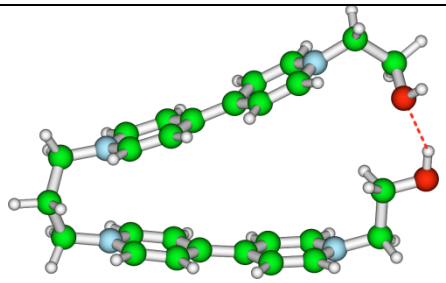
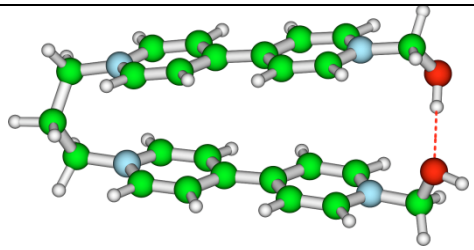
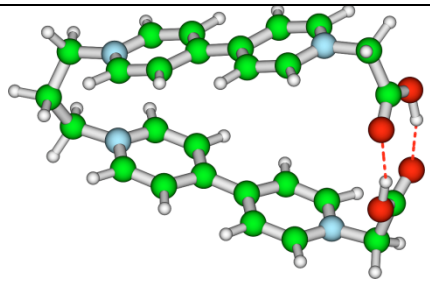
Among the factors responsiblefor the stabilization of the closed form of the system are the H-bonding interactions. Strangely enough, the amide terminal group despite exhibiting a strong energetic preference for the closed from does not show any H-bond. We wanted to gather the geometries of these systems with H-bond or potential H-bond in order to compare them, seeFigure 2.8.

It is clear from this figure that to have a closed form with small D2 distances, we need, of course, a force which will bring the terminal group together, and in this respect, terminal group with OH or CO2H group are quite promising but we have also to deal with the steric hindrance.

Figure 2.8Models used which presents H-bonding interactions for the closed conformation. The order of these geometrical models, presented here, are from the weakest to the strongest H-bonding model.

Model / Method	Geometrical model	H-bonding parameters[3]
$^{S}[\mathbf{2}^{2+}]_{dim}$ / B3LYP		Do not present H-bonding interactions. D2=4.153 Å

[3] $D_{Dn—An}$ represents the H-bonding distance between the donor and the acceptor and $A_{Dn\text{-}H—An}$ is the angle of the H-bonding between the donor and the acceptor. The *n* represents the number of H-bonding in the complex. The *D2* represent the distance between the two nitrogen that link the unit A and unit B to the bis-viologene units.

$^{S}[(CF3\text{-}CH2OH)^{2+}]_{dim}$ / B3LYP		$D_{D1—A1}$=3.136 Å $A_{D1\text{-}H—A1}$=144.9° D2=4.626 Å
$^{S}[(2CH2OH\text{-}2CH2OH)]_{dim}$ / B3LYP		$D_{D1—A1}$=2.929 Å $A_{D1\text{-}H—A1}$=153.8° D2=6.072 Å
$^{S}[(CH2OH\text{-}CH2OH)]_{dim}$ / B3LYP		$D_{D1—A1}$=2.840 Å $A_{D1\text{-}H—A1}$=174.4° D2=4.376 Å
$^{S}[(COOH\text{-}COOH)]_{dim}$ / B3LYP		$D_{D1\text{-}A1}$=2.670 Å $A_{D1\text{-}H1—A1}$ =163.0° $D_{D2—A2}$=2.669 Å $A_{D2\text{-}H2—A\ 2}$=163.0° D2=6.353 Å

Typical distance for the Donor-Acceptor distances for H-bonded system is already around 3 Å and according to Table 2.5, the D2 distance can be as small as 3.5 Å. So the units A and B, which havealready a certain size, prevent the formation of the dimers, in closed form. This leads to a decreasing of the relative energy between the closed and open form. For example $^{S}[(2CH2OH\text{-}2CH2OH)]_{dim}$ and $^{S}[(CH2OH\text{-}CH2OH)]_{dim}$ where we have two or one units of CH_2 respectively, we can see that the geometrical parameters and energetically point of view (ΔE) are in favor of the model with only one CH_2 group present in the terminal units A and B. We observe if D1 distance is quasi the same 3.010 and 3.015 Å respectively (Table2.5 entry 4 and 5), the D2 distance presents an important displacement of approximate 2 Å in the favor

of the $^{S}[(CH2OH\text{-}CH2OH)]_{dim}$ model (Table2.5). This is responsible for increasing stabilization energy in the favor of the closed form by 1.52 kcal/mol, which represents the $\Delta\Delta E$ of this two model. (Table2.6).

Finally, we wanted to gain insight into the interaction between the two amides and the two carboxylic groups. We computed the SAPT[39] interaction energy, which give not only the interaction energy but also give the physical contribution to the interaction energy: namely electrostatic energy, exchange energy, dispersion and induction energy. SAPT is a very accurate method to compute and decompose the interaction energy but it is also a very demanding one both from the point of view of the memory and the disk space needed. So, we computed the interaction energy using SAPT but limited to the terminal group, fulfilling the valence with an additional hydrogen to the CH2 group. It can be seen as a rough approximation but the CH2 group should stop the propagation of the mesomer effects between the viologen and the terminal group. Nevertheless, we are aware of the approximation. We used the geometry of the full system optimized at the B3LYP-D level. The SAPT results show that indeed the $[(COOH\text{-}COOH)]_{dim}$ exhibits all the properties of a H-bonding system with a high contribution from the electrostatic part of -35.0 kcal/mol quenched by an exchange contribution of 36.7 kcal/mol. The induction is also strong around -23 kcal/mol and the exchange-induction part is around 10 kcal/mol. Finally the dispersion energy is far from negligible around -10 kcal/mol and the total SAPT interaction energy is of -15.4 kcal/mol. On the contrary, for the amide dimer the electrostatic and induction contribution are small -4.6 kcal/mol and 2.0 kcal/mol respectively partially quenched by the exchange contribution 4.5 kcal/mol and the exchange-induction contribution 2.2 kcal/mol. The dispersion contribution is proportionally important -3.8 kcal/mol compared to a total SAPT interaction energy of -3.08 kcal/mol. Thus, contrary to what one might think, the interaction of the amide group is more similar to the interaction, which can be found in the system where dispersion is the driving force of the interaction.

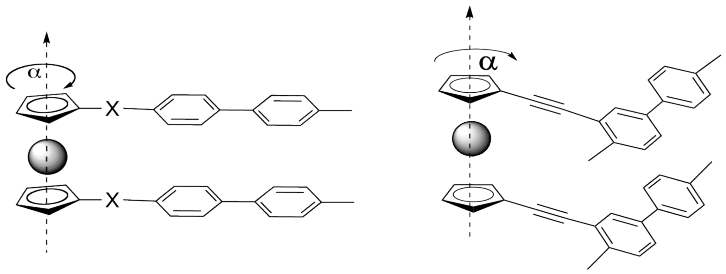
In conclusion, to have systems which promote closed forms relative to open forms, we must have terminal units which favor dispersion energy of these units: the $[(PhF\text{-}Ph)]_{dim}$ system is most favorable in this respect and also the $[2^{2+}]_{dim}$even if this system at first glance would not be placed in this category.

With terminal groups with H-bonding involved and minimal hindrance, it is also possible to favor strongly the closed form as for example with the $^{S}[(COOH\text{-}COOH)]_{dim}$ system but the geometry of the closed form is less "closed" than in the systems where the driving force is the dispersion. From our models proposed here the most suitable, form an energetical point of view, are in this order $^{S}[(PhF\text{-}Ph)]_{dim}$, $^{S}[(COOH\text{-}COOH)]_{dim}$ and $^{S}[(Ph\text{-}Ph)]_{dim}$.

2.4 Molecular rotors based on Ferrocene-bis-pyridinium units

The present study mostly aims to demonstrate, from a computational point of view, that the redox-activable properties of viologenes can be exploited to control the rotation of ferrocene between an open charge-repelled form to a closed state wherein two reduced π-radicals would be ideally positioned face to face to promote an efficient overlapping stabilizing the sandwich-like π-dimer form. For this study we use the new hybrid exchange-correlation functional CAM-B3LYP[40] which combines the hybrid qualities of B3LYP[16,14] with the long-range corrections proposed by Tawada et all.[41] to obtain a functional, which can be used successfully to compute properties, like charge transfer or polarizability of long molecular chains. For all the calculations we use the well known basis set 6-31G* for the C,N,H atoms and the SDD[42,43] pseudopotential for Fe atom, combined with SDD[44] basis sets for the same element. The structures are optimized in gas and/or in the n,n-DiMethylFormamide (ε = 37.219) solvent using PCM model.

Reduced form of Ferrocene-X-bis-viologen is abbreviated $[X^{+2}]_{dim}$,where X units are highlighted in Figure 2.9:



X=Methylene $[Methylene^{+2}]_{dim}$

X=Ethylene $[Ethylene^{+2}]_{dim}$ Ethynyl $[Ethynil^{+2}]_{dim}$

X=Phenyl $[Ph^{+2}]_{dim}$

Figure 2.9Models proposed for the Ferocene-bis-pyridinium units

Geometry optimization of non-reduced form $[X^{+4}]_{dim}$ of these systems leads invariantly to an open structure. The relative energy of the $[Ph^{+4}]_{dim}$ in DMF solvent was determined by a relaxed scan calculation imposing the dihedral angle between the ferrocene linkers. The 0° point was set as the energy reference point and the all other points was calculated relatively to this point.

Potential energy surface of the $[Ph^{+4}]_{dim}$ in the DMF solvent was determined by an relaxing scan procedure rotating the α angle (where α is defined in Figure 2.9) from 0 to 180 degrees. $\alpha = 0^{\circ}$ is the situation when the two bipyridines are superposed corresponding to the closed form. For $\alpha = 180^{\circ}$, the two bipyridines are completely separated corresponding to a maximum D2 distance. In Figure 2.10 the results are plotted as relative energies to the closed form ($\alpha = 0^{\circ}$). The minimum is found to be to $\alpha = 144^{\circ}$. At $\alpha = 144^{\circ}$ we perform full optimization geometry of the system. The energy differences between the closed form and the minimum and between the closed form and the $\alpha = 180^{\circ}$ conformation are found to be 3.34 kcal/mol and 2.88 kcal/mol respectively. The fitting of these points, using a dihedral force field type function $F(x) = A \cdot (1 + \cos(B \cdot x + C))$, presented in the Figure 2.10 (dashed line), allows us to find the parameters A, B and C of this function. The parameters are found to be: A=-1.67481, B=1.27674 and C=3.14155.

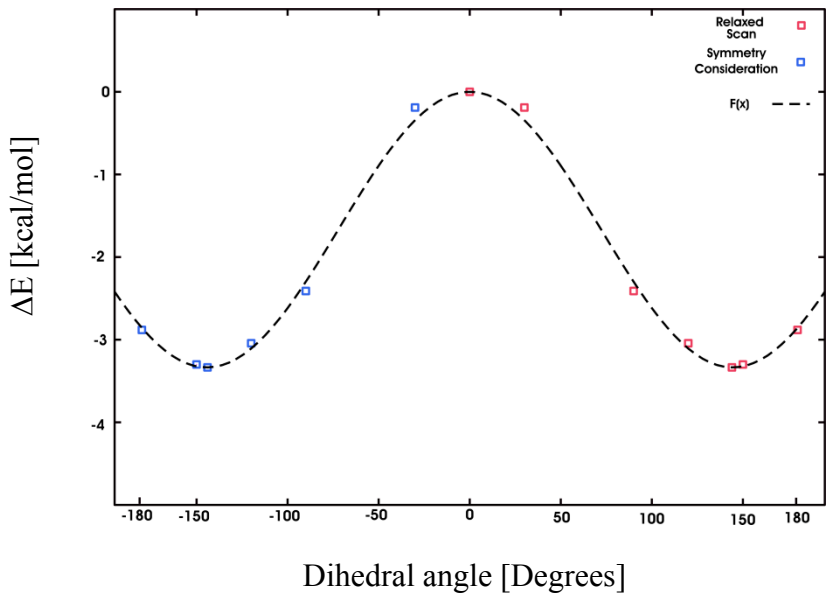
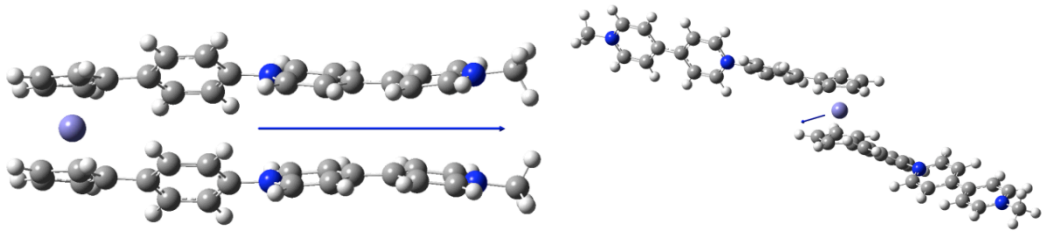


Figure 2.10Potential energy surface of the $[Ph^{+4}]_{dim}$ in the DMF solvent. The red points are the relaxed scan optimization form 0° to 180°. The blue points are considered from a C2 symmetry of the complex $[X^{+4}]_{dim}$ and the F(x) represents the functional form of the potential with $F(x) = A \cdot (1 + \cos(B \cdot x + C))$ used for fitting.

The theoretical description of the doubly reduced species, $[X^{+2}]_{dim}$, is investigated and both the open and closed π-dimer formswere characterized. These doubly reduced species were characterized also by ESR (Electron Paramagnetic Resonances) and the silent character found in the ESR spectra indicates that the coupling between bis-viologene π-radical dimers is made in a such way that the reduced form of these systems has singlet spin state nature. Based on these experimentally observations, the reduced form $[X^{+2}]_{dim}$ was characterized, from a computational point of view only for the singlet spin state.

Table2.7. Energetic difference between closed and open form of the complex $^{S}[X^{2+}]_{dim}$in reduced state with different X terminal units; the closed conformation is favored when ΔE<0

Dimer	$\Delta E_{Gas\text{-}phase}$ [kcal/mol]	$\Delta E_{DMF\text{-}solvent}$ [kcal/mol]	Total dipole moment in Gas [Debye]		Total dipole moment in DMF [Debye]	
			Closed	Open	Closed	Open
$[Ph^{+2}]_{dim}$	-2.26	-27.81	28.4	8.2	36.9	9.3
$[Ethynil^{+2}]_{dim}$	-8.04	-27.74	15.8	5.6	21.7	6.4
$[Ethylene^{+2}]_{dim}$	+1.55	-21.11	22.8	3.0	26.7	2.7
$[Methylene^{+2}]_{dim}$	-2.15	-18.47	16.1	0.2	19.3	1.1



$[Ph^{+2}]_{dim}$ – closed form $[Ph^{+2}]_{dim}$ – open form

Figure 2.11Dipolar moment of the $[Ph^{2+}]_{dim}$ closed and open conformations.

We can see that the inclusion of solvent effects strongly stabilizes the closed conformation (see Table2.7). This trend can be related to the different value of the dipole moments. Indeed, closed forms exhibit great values of dipole moment ranging from 15 Debye to 28 Debye whereas for the open forms, the highest value computed is only 8.2 Debye. Compared to the

value without the ferrocene, for example compared to the $^{S}[(Me\text{-}Me)]_{dim}$ system, we can see that the closed form is strongly favored when the ferrocene is present. The linkers which favor the mesomery ($[Ph^{+2}]_{dim}$ and $[Ethynil^{+2}]_{dim}$ are clearly favored in the closed form when solvent effect is included.

Like previously, the reduced conformation exhibits planar viologens whereas the non-reduced conformation is tilted.

2.4.1 Geometrical parameters

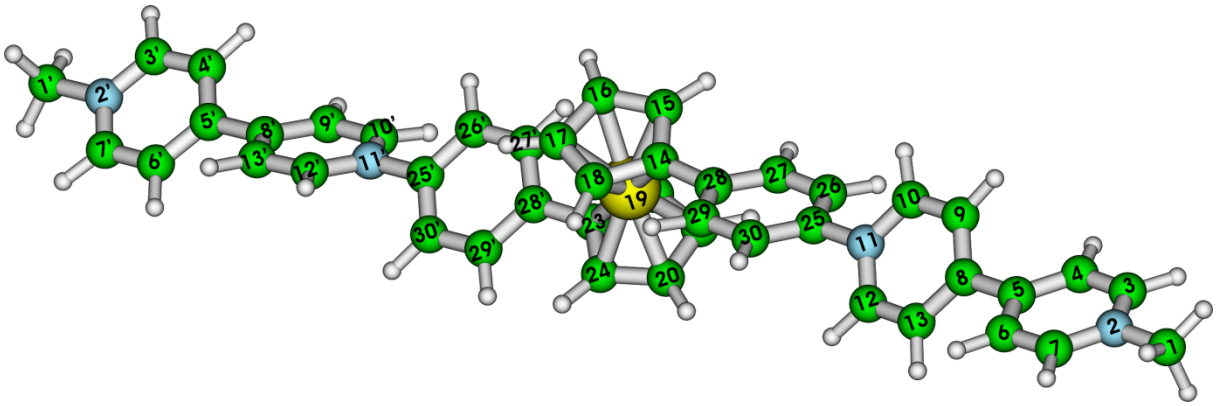


Figure 2.12Atoms label of the $[Ph^{+4}]_{dim}$

Table2.8. Geometrical parameters tables:

viologene geometry. In all these tables the distances are measured in Angstroms and the angles are in degrees.

	$[Ph^{+4}]_{dim}$	$[Ph^{+2}]_{dim}$	$[Ethynil^{+2}]_{dim}$	$[Ethylene^{+2}]_{dim}$	$[Methylene^{+2}]_{dim}$
C1-N2	1.486	1.463	1.464	1.465	1.466
N2-C3	1.347	1.364	1.365	1.364	1.361
C3-C4	1.380	1.361	1.361	1.361	1.363
C4-C5	1.397	1.430	1.429	1.428	1.427
C5-C6	1.397	1.428	1.429	1.429	1.426
C6-C7	1.380	1.361	1.381	1.361	1.363
C7-N2	1.347	1.365	1.365	1.364	1.362
C5-C8	1.487	1.428	1.429	1.428	1.431
C8-C9	1.398	1.426	1.429	1.427	1.426

C9-C10	1.379	1.363	1.359	1.380	1.362
C10-N11	1.352	1.366	1.363	1.365	1.364
N11-C12	1.351	1.368	1.378	1.361	1.356
C12-C13	1.380	1.360	1.379	1.364	1.364
C13-C8	1.398	1.428	1.418	1.426	1.425
N11-C25	1.453	1.437	1.463	1.482	1.479
γ_1	40.1	0.3	1.3	0.9	4.4
α [4]	144.8	9.4	0.0	0.2	30.1

Cp-ring

	$[Ph^{+4}]_{dim}$	$[Ph^{+2}]_{dim}$	$[Ethynil^{+2}]_{dim}$	$[Ethylene^{+2}]_{dim}$	$[Methylene^{+2}]_{dim}$
C14-C15	1.432	1.431	1.435	1.428	1.429
C15-C16	1.415	1.417	1.414	1.420	1.416
C16-C17	1.422	1.420	1.422	1.418	1.419
C17-C18	1.415	1.417	1.414	1.418	1.417
C18-C14	1.431	1.430	1.436	1.425	1.427
$Dist_{Cp\text{-}Cp}$ [5]	~3.3	~3.3	~3.4	~3.4	~3.5

Linkers : Geometry of Phenyl linker

Complex	C25-C26	C26-C27	C27-C28	C28-C29	C29-C30	C30-C25
$[Ph^{+4}]_{dim}$	1.392	1.385	1.403	1.402	1.386	1.390
$[Ph^{+2}]_{dim}$	1.393	1.384	1.400	1.400	1.387	1.391

Geometry of the Ethynyl linker

Complex	Cp-C1	C1-C2	C2-Viologene
$[Ethynyl^{+2}]_{dim}$	1.415	1.211	1.415

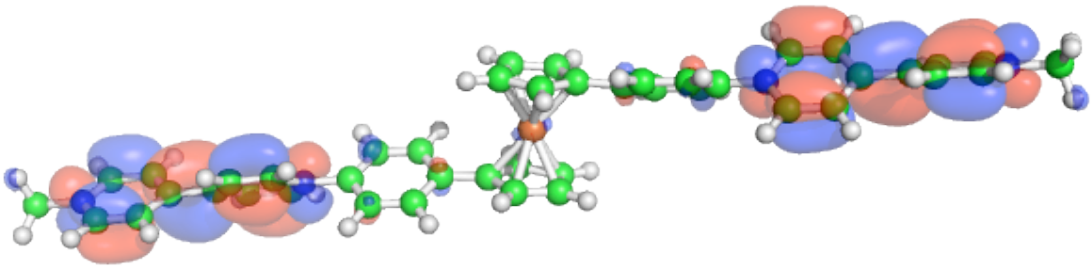
[4] γ_1 is the dihedral angle between the two pyridines and α is the rotation angle presented in figure 2.9

[5] $Dist_{Cp\text{-}Cp}$ represents the distance between the two Cyclopentadiene rings of the Ferrocene unit.

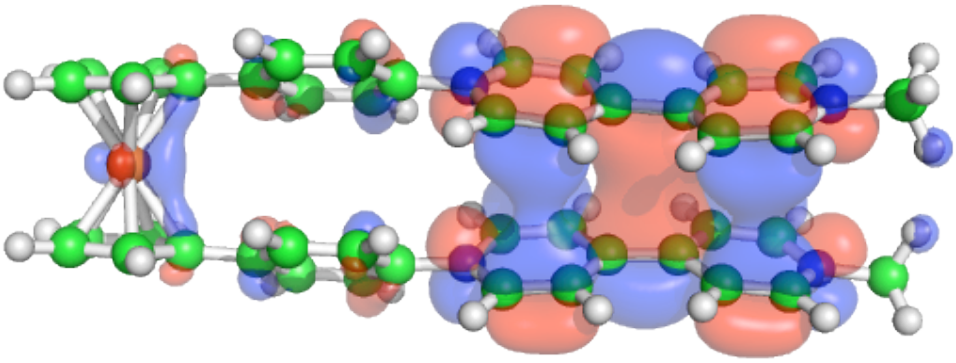
Geometry of the Ethylene and Methylene linker

Complex	Cp-C1	C1-C2	C2-Viologene
$[Ethylene^{+2}]_{dim}$	1.508	1.522	1.482
$[Methylene^{+2}]_{dim}$	1.497	-	1.479

The molecular orbital analyses of the $[Ph^{+2}]_{dim}$ in the closed form are shown in the Figure 2.13. The LUMO (Lowest Unoccupied Molecular Orbital) of the open starting material $[Ph^{+4}]_{dim}$ becomes the HOMO (Highest Occupied Molecular Orbital) of the doubly reduced $[Ph^{+2}]_{dim}$ species with a bonding character associated to the C8-C5 (C8'-C5') bonds. The same bonding character was observed for the C4-C3 (C6-C7) and C9-C10 (C12-C13) bonds, whereas an antibonding character is associated to the C4-C5 (C5-C6), N2-C3 (N2-C7) and N10-C11 (N11-C12) bonds. The electron doping process has thus significant consequences on the geometry of the viologen units as shown by the differences observed between the bond distances tabulated in Table 2.8. Reducing $[Ph^{4+}]$ into $[Ph^{2+}]_{dim}$ led especially to the significant bond lengths increases for N2-C3 (N2-C7), C4-C5 (C5-C6) and to the concomitant decreases of the C8-C5, C4-C3 (C6-C7) bond distances. As seen inFigure 2.13, the HOMO of $[Ph^{+2}]_{dim}$is also clearly responsible for the association between both viologen fragment forming the closed π-dimer $[Ph^{+2}]_{dim}$ species. From the HOMO of $[Ph^{+2}]_{dim}$we can also note that the π system is not limited to the viologen unit but has also contribution from the Phenyl linker and from the ferrocene.



LUMO of $[Ph^{+4}]_{dim}$



HOMO of $[Ph^{+2}]_{dim}$

Figure 2.13Frontier molecular orbitals calculated for $[Ph^{+4}]_{dim}$ and $[Ph^{+2}]_{dim}$. The electron density isosurface is 0.02 a.u.

2.5 Conclusions

In conclusion, this theoretical investigation shows the upmost importance of dispersion energy in the modelisation of π-dimers. This inclusion of dispersion energy and its effects has been made possible through the use of dispersion terms or corrected potentials in conjunction with DFT methods. These DFT-D methods proved to favor the closed structure $[\mathbf{2}^{2+}]_{dim}$ whatever the spin state considered or the functional used. Due to the structure of $[\mathbf{2}^{2+}]_{dim}$, there is no doubt about the importance of dispersion terms and about their necessity beyond the DFT level of theory[45]. Thus, a preference for the dimeric structure in the singlet state can be deduced from our theoretical investigations but the observation of the open structure as a minimum on the PES, led us to assume that interconversion between the two conformations has to be considered.

We investigate also the influence of the terminal group using different kind of terminal groups and we have shown that the most favorable to a nice closed form where the groups whose interactions is due to dispersion contribution. Strangely, we have shown that the amide terminal group falls in this category both from a geometrical and energetical point of view. The H-bonding systems will be better fitted for a system with the dimer has a distance between its two sub-units slightly bigger around 4 - 4.5Å.

Relative bibliography to the Chapter 2

(1) Stefan Grimme Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.***2006**, *27*, 1787-1799.

(2) Riley, K. E.; Pitoň ák, M.; Jureč ka, P.; Hobza, P. Stabilization and Structure Calculations for Noncovalent Interactions in Extended Molecular Systems Based on Wave Function and Density Functional Theories. *Chem. Rev.***2010**, *110*, 5023-5063.

(3) Mackie, I. D.; DiLabio, G. A. Accurate dispersion interactions from standard density-functional theory methods with small basis sets. *Phys. Chem. Chem. Phys.***2010**, *12*, 6092.

(4) J. L. Sessler; P. A. Gale; W.-S. Cho *Anion Receptor Chemistry*; RSC: Cambridge, 2006.

(5) Gale, P. A.; García-Garrido, S. E.; Garric, J. Anion receptors based on organic frameworks: highlights from 2005 and 2006. *Chem. Soc. Rev.***2008**, *37*, 151.

(6) Beer, P. D.; Bayly, S. R. Dans *Anion Sensing*; Topics in Current Chemistry; Springer Berlin / Heidelberg, 2005; Vol. 255, p. 125-162-162.

(7) Garcia-Espada, E.; Diaz, P.; Llinares, J. M.; Bianchi, A. Anion coordination chemistry in aqueous solution of polyammonium receptors. *Coor. Chem. Rev.***2006**, *250*, 2952-2986.

(8) Paul M. S. Monk *The Viologens: Physicochemical Properties, Synthesis and Applications of the Salts of 4,4'-Bipyridine*; 1998 éd.; Wiley, 1998.

(9) Johnson, E. R.; Mackie, I. D.; DiLabio, G. A. Dispersion interactions in density-functional theory. *J. Phys. Org. Chem.***2009**, *22*, 1127-1135.

(10) Mackie, I. D.; DiLabio, G. A. Interactions in Large, Polyaromatic Hydrocarbon Dimers: Application of Density Functional Theory with Dispersion Corrections. *J. Phys. Chem. A***2008**, *112*, 10968-10976.

(11) Ben Fredj, A.; Ben Lakhdar, Z.; Ruiz-Lopez, M. Six-coordination in Chlorophylls: The fundamental role of dispersion energy. *Chem. Phys. Lett.***2009**, *472*, 243-247.

(12) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.***2005**, *167*, 103-128.

(13) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A.

Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford CT, 2004.

(14) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B***1988**, *37*, 785.

(15) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A***1988**, *38*, 3098.

(16) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.***1993**, *98*, 5648.

(17) DiLabio, G. A. Accurate treatment of van der Waals interactions using standard density functional theory methods with effective core-type potentials: Application to carbon-containing dimers. *Chem. Phys. Lett.***2008**, *455*, 348-353.

(18) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.***1934**, *46*, 618.

(19) Cancès, E.; Mennucci, B.; Tomasi, J. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.***1997**, *107*, 3032.

(20) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. Ab initio study of ionic solutions by a polarizable continuum dielectric model. *Chemical Physics Letters***1998**, *286*, 253-260.

(21) Mennucci, B.; Tomasi, J. Continuum solvation models: A new approach to the problem of solute's charge distribution and cavity boundaries. *J. Chem. Phys.***1997**, *106*, 5151.

(22) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. New developments in the polarizable continuum model for quantum mechanical and classical calculations on molecules in solution. *J. Chem. Phys.***2002**, *117*, 43.

(23) VandeVondele, J.; Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.***2007**, *127*, 114105.

(24) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B***1996**, *54*, 1703.

(25) Hartwigsen, C.; Goedecker, S.; Hutter, J. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B***1998**, *58*, 3641.

(26) Krack, M. Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals. *Theor Chem Acc***2005**, *114*, 145-152.

(27) Kannappan, R.; Bucher, C.; Saint-Aman, E.; Moutet, J.; Milet, A.; Oltean, M.; Métay, E.; Pellet-Rostaing, S.; Lemaire, M.; Chaix, C. Viologen-based redox-switchable anion-binding receptors. *New J. Chem.***2010**, *34*, 1373.

(28) Matteo, A. D. Structural, electronic and magnetic properties of methylviologen in its reduced forms. *Chem. Phys. Lett.***2007**, *439*, 190-198.

(29) Saielli, G. Ion-Pairing of Octyl Viologen Diiodide in Low-Polar Solvents: An Experimental and

Computational Study. *J. Phys. Chem. A***2008**, *112*, 7987-7995.

(30) Femoni, C.; Iapalucci, M.; Longoni, G.; Tiozzo, C.; Wolowska, J.; Zacchini, S.; Zazzaroni, E. New Hybrid Semiconductor Materials Based on Viologen Salts of Bimetallic Fe–Pt and Fe–Au Carbonyl Clusters: First Structural Characterization of the Diradical π-Dimer of the Diethylviologen Monocation and EPR Evidence of its Triplet State. *Chem. Eur. J.***2007**, *13*, 6544-6554.

(31) Monk, P. M. S. *The Viologens: Physicochemical Properties, Synthesis and Applications of the Salts of 4,4'-Bipyridine*; Wiley, 1998.

(32) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proceedings of the National Academy of Sciences of the United States of America***2002**, *99*, 12562-12566.

(33) Laio, A.; Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.***2008**, *71*, 126601.

(34) Michel, C.; Laio, A.; Mohamed, F.; Krack, M.; Parrinello, M.; Milet, A. Free Energy ab Initio Metadynamics: A New Tool for the Theoretical Study of Organometallic Reactivity? Example of the C−C and C−H Reductive Eliminations from Platinum(IV) Complexes. *Organometallics***2007**, *26*, 1241-1249.

(35) Loehr, H. G.; Voegtle, F. Chromo- and fluoroionophores. A new class of dye reagents. *Accounts of Chemical Research***1985**, *18*, 65-72.

(36) Reynes, O.; Bucher, C.; Moutet, J.; Royal, G.; Saint-Aman, E.; Ungureanu, E. Electrochemical sensing of anions by redox-active receptors built on the ferrocenyl cyclam framework. *Journal of Electroanalytical Chemistry***2005**, *580*, 291-299.

(37) Pernal, K.; Podeszwa, R.; Patkowski, K.; Szalewicz, K. Dispersionless Density Functional Theory. *Phys. Rev. Lett.***2009**, *103*, 263201.

(38) Podeszwa, R.; Patkowski, K.; Szalewicz, K. Improved interaction energy benchmarks for dimers of biological relevance. *Phys. Chem. Chem. Phys.***2010**, *12*, 5974.

(39) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.***1994**, *94*, 1887-1930.

(40) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.***2004**, *393*, 51-57.

(41) Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. A long-range-corrected time-dependent density functional theory. *J. Chem. Phys.***2004**, *120*, 8425.

(42) Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. Energy-adjusted ab initio pseudopotentials for the first row transition elements. *J. Chem. Phys.***1987**, *86*, 866.

(43) Andrae, D.; Hausermann, U.; Dolg, M.; Stoll, H.; Preus, H. Energy-adjustedab initio pseudopotentials for the second and third row transition elements. *Theoret. Chim. Acta***1990**, *77*, 123-141.

(44) T. H. Dunning, Jr.; P. J. Hay H. F. Schaefer III: New York, 1976; Vol. 3, p. 1-28.

(45) Milet, A.; Korona, T.; Moszynski, R.; Kochanski, E. Anisotropic intermolecular interactions in van der Waals and hydrogen-bonded complexes: What can we get from density functional calculations? *J.*

*Chem. Phys.***1999**, *111*, 7727.

Chapter 3 WEAK INTERACTIONS STUDIED BY DFT AUGMENTED BY EMPIRICAL DISPERSIONS TERMS "DFT-D"

Résume

Dans ce chapitre, nous étudions l'énergie d'interaction d'un système non-covalent formé par l'interaction entre un système comme le cryptophane complexé avec différentes petites molécules. Le cryptophane peut servir d'hôte pour les petites molécules et l'énergie d'interaction des complexes est fortement dépendante des interactions à longue distance. Dans la deuxième partie de ce chapitre, nous présentons les développements réalisés dans le domaine de la DFT-D par notre groupe en particulier dans les paramétrages et la mise en œuvre des paramètres pour l'atome de platine à l'aide de méthode du Grimme[1]dans le program Gaussian09[2]. À la fin de ce chapitre, nous avons étudié la réaction d'un complex Pt(II)-utilisé comme composé catalytique dans une réaction organique de cycloaddition [2+1] en phase gazeuse.

.

3.1 Abstract

For computational chemistry the study of the weak interactions as non-covalent bonding or H-bonding is an open field under continuous development. It is well known that *ab-initio* post Hartree-Fock methods are needed to describe this kind of interactions at the chemical accuracy but these simulations are very expensive from a computational time point of view and computational resources (RAM and disk space). To overcome these problems, in the last couples of years the computational chemistry community has proposed promising solution to these problems. The alternative solution is to use DFT methods with corrections for the dispersion energy added empirically on the top of KS energy or in the new exchange correlation functional designed to correct the long-range correlation term.

In this chapter, we study the binding energy of a non-covalent bonding system between an host-guest system such as Cryptophane complexed with different small molecules. Cryptophane supra molecular complex can serve as a host for small molecules and the binding energy of complexes is strongly dependent on the long-range interactions. In the second part of this chapter we present the developments made in the field of the DFT-D by our group especially in the parameterizations and implementation part of the parameters for the Platinum atom using Grimme's method[1] into the Gaussian09 program[2]. At the end of this chapter we studied a reactivity pathway of a Pt(II)-complex used as catalytic compound in an organic reactions of [2+1] cycloaddition in gas phase.

3.2 Introduction

Since the discovery of the first Cryptophane, known under the name of Cryptophane-A by Jacqueline Gabard and André Collet[3], in 1981, numerous studies and research have been conducted to understand the chemical properties of such compounds with molecular recognition properties. This class of organic compounds has taken its interest because of its particular structure of this compound, which shows a central cavity which can specificity encapsulate and recognize molecules with a given specific structure. Because of this property, cryptophanes become very promising agents for the recognition and specific storage for many organic molecules especially small gaseous molecules such as methane and chloroform. Another interesting application of cryptophane is the use of this compound as potential storage cage for H_2 molecules, with an important impact in the environmental chemistry. However, the lack of experimental data about hydrogen molecules encapsulated in

Cryptophane and due to the difficulty of characterizing this interaction by ^{1}H-NMR technique, theoretical studies are highly required. Indeed, computational chemistry may provide information useful for a better understandingof the process of encapsulation and interaction between H_2 and Cryptophane.

In our study we will calculate the associations energies and interactions energies for different models Ligand-Cryptophane using different DFT functionals. The fact that these types of interactions are non-covalent (weak) rises considerable difficulties for the use of ordinary DFT methods. The strategy used to overcome this problem is to use DFT-D method, in which the dispersion term has been introduced empirically, or functionals corrected for long-range interactions. The functionals used in this study are: B3LYP[4,5], CAM-B3LYP[6], B97D[1], LCwPBE[7-10] and wB97XD[11] with the standard basis set 6-31G*.

3.3 Cryptophane

These organic molecules are cyclic; they have an aromatic-based kernel, posses a spherical structure with an internal cavity, lipophilic, responsible for their role as agents of encapsulation and molecular recognition[12]. Typically, the cryptophane is an association of two racemic (cyclotribenzylene) units stacked and connected by three bridges to limit an internal cavity having a specific size influenced by the linker units used for synthesis. The two racemic units can be linked in such way to give a *syn* or *anti* cryptophane isomers.

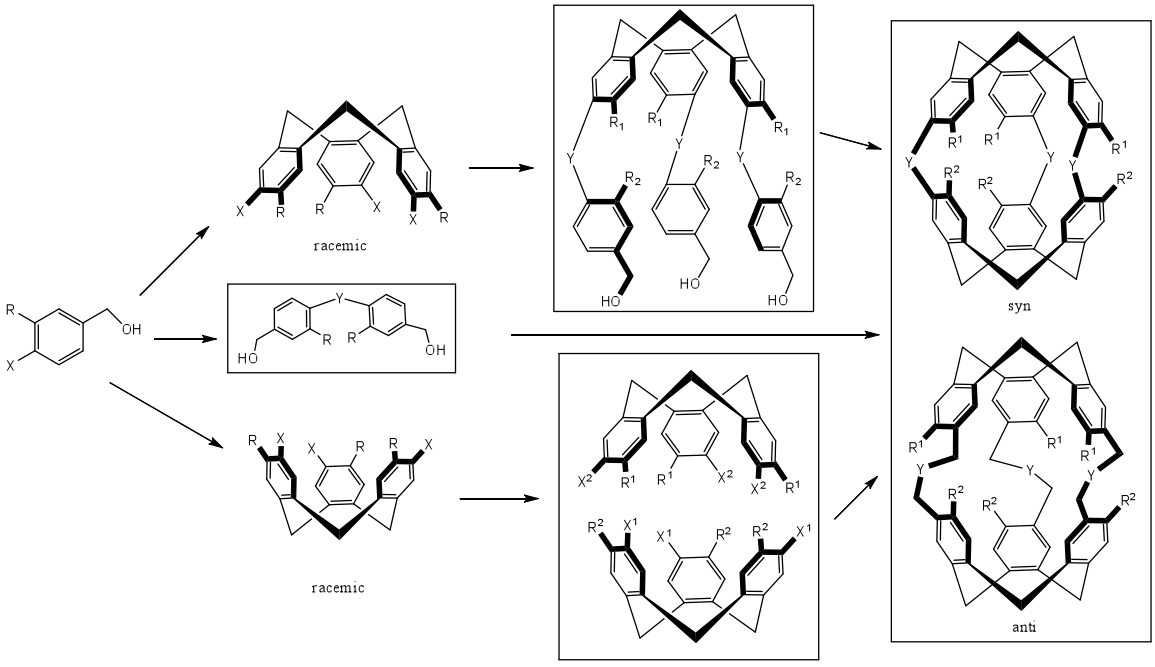
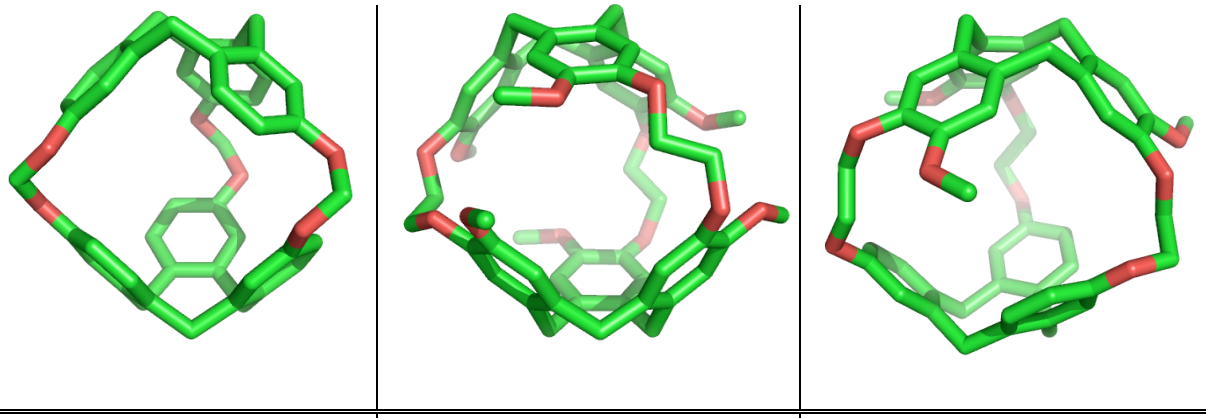


Figure 3.1The Cryptophane formation reaction

We use three models for Cryptophane named: Cryptophane-111 the smallest cyptophane, Cryptophane-A and Cryptophane-C. For the ligands molecules we used respectively: H_2, CH_4 and two enantiomers of HCFClBr, the **R**-form and the **S**-form. These molecules have also been chosen as ligands because, they have molecular volumes small enough to enter the cavity of the Crypthophanes. The last system Cryptophane C in interaction with the enantiomers of HCFClBr is quite a challenge since Cryptophane C is able to discriminate between the two enantiomers. Thanks to this interaction, the absolute configuration (link between the configuration and the deviation of light) of the R and S HCFClBr has been determined.



Cryptophane 111	Cryptophane A	Cryptophane C
Molecular formula $C_{45}H_{36}O_6$	Molecular formula $C_{54}H_{54}O_{12}$	Molecular formula $C_{51}H_{48}O_9$

Figure 3.2The structures of the three investigated Cryptophane complexes. Hydrogen atoms are removed for clarity.

3.3.1 Cryptophane-111 in interaction with H2

Cryptophane-111 is the smallest compound from his family possessing a rigid cavity of only ~ 81 $Å^3$. This compound presents remarkably binding behavior for xenon in organic solutions[13]. Currently, this cryptophane-111 complex presents interest because it can encapsulate small molecules like di-Hydrogen molecules. According to experimental results[14]the volume of the di-hydrogen molecule is around 8.6 $Å^3$ and the distance between the two interacting H_2 molecules is around 6.5 bohr[15]. From an experimental point of view, it was quite difficult to study these interactions by H-NMR technique and the number of encapsulated H_2 molecules remains unknown[14]. Thus, we sought, from a theoretical point of view, to obtain information on the number of hydrogen molecules, which can be included in

the cavity of the cryptophane-111 complex. Due to the volume of one single H_2 molecule, we did not study the inclusion of only one H_2 molecule. To solve this problem, we chose as a starting point, several systems of cryptophane-111 with two or three H_2 molecules inside the cryptophane cavity. With 4 and up di-hydrogen molecules, all our attempts to optimize the cryptophane with more than 3 di-hydrogen molecules inside the cryptophane's cavity failed. For each DFT methods, we performed optimizations for the dissociated form (Cryptphane-111 and H2 alone) and for the associated form (nH$_2$@cryptophanes-111) with "n" numbers of Hydrogen molecules inside the cavity. Then, we calculate the energy of association ΔE. The results in kJ/mol are gathered in Table 3.1

The association energy is defined as: $\Delta E = E_{AB} - (E_A + E_B)$ .Each of the compounds involved in this later equation is optimized alone and the whole idea is to deal with it as a "normal" reaction. On the other hand, the interaction energy is calculated as: $E_{interaction} = E_{AB}^{AB}(r_{AB}) - \left(E_A^{AB}(r_{AB}) + E_B^{AB}(r_{AB})\right)$ with the energy of the monomers computed using their geometry in the dimer and the basis set of the whole complex. Indeed, $E_A^{AB}(r_{AB})$ is the energy of A (subscript) computed in the basis set of AB (superscript) in the geometry of AB (specified in the subscript of r in the parenthesis). This definition of the interaction energy is corrected from basis sets superposition error (BSSE) using counterpoise (CP) method. The basis superposition error using CP method is defined as =

$$E_{BSSE} = E_A^A(r_{AB}) + E_B^B(r_{AB}) - \left(E_A^{AB}(r_{AB}) + E_B^{AB}(r_{AB})\right).$$

Table 3.1Association energy of $(H_2)_n$ @ Crypthophane-111 host; with n-number of H_2 molecules.

Method	ΔE kJ/mol $(H_2)_2$@Cry-111	ΔE kJ/mol $(H_2)_3$@Cry-111
B3LYP/6-31G*	+10.0	+9.1
B97D/6-31G*	-16.8	-15.9
CAM-B3LYP/6-31G*	+0.5	+18.8
LCwPBE/6-31G*	+1.5	+21.0
wB97XD/6-31G*	-24.3	-22.2

We can observe that B3LYP clearly give results that emphasize that two molecules are not bind in the cavity of cryptofane-111. To a lesser extent, it is also the case for some DFT with the long-range corrections CAM-B3LYP and LCwPBE. The results for three H_2 molecules are even worse for these functionals especially for LCwPBE and CAM-B3LYP. We clearly observe only for these two functionals that the association energy is strongly less favorable for 3 H_2 molecules than for 2 di-hydrogen molecules by around 19 kJ/mol. For the other functionals, the difference is surprisingly quite small, around 1-2 kJ/mol since only the DFT with value closed to zero for two H_2 molecules exhibit a difference. The DFT-D corrected empirically for dispersion, as B97D and wB97XD give favorable association accordingly with the binding character of the H_2 molecules inside of the Cripthofane-111 complex.

When we correct the results of the Basis Sets Superposition's Errors (BSSE) using the counterpoise method (CP)[16] the conclusions remain the same with even more pronounced tendencies.

Table 3.2Basis set superposition error of the association energy of Cryptofane-111 encapsulating two H_2 molecules.

Method	ΔE kJ/mol $(H_2)_2$@Cry-111	$\Delta E_{corrected}$ kJ/mol $(H_2)_2$@Cry-111	BSSE kJ/mol $(H_2)_2$@Cry-111
B3LYP/6-31G*	+10.0	+17.6	+7.6
B97D/6-31G*	-16.8	-9.1	+7.4
CAM-B3LYP/6-31G*	+0.5	+8.5	+8.0
LCwPBE/6-31G*	+1.5	+4.3	+2.8
wB97XD/6-31G*	-24.3	-21.1	+3.2

Having computed the BSSE corrections, we calculate the interactions energy with BSSE corrections. The Basis set superposition errors is defined as: . B3LYP and CAM-B3LYP show positive interaction energy whereas LCwPBE exhibit nearly no interaction energy. On the other hand, B97D and wB97XD show negative interaction energy of around 10 kJ/mol between the Cryptophane-111 cage and the two Hydrogen molecules inside of the cavity.

Table 3.3Interaction energy between $(H_2)_2$@Cry-111 complex.

Method	$E_{interaction}$ kJ/mol $(H_2)_2$@Cry-111
B3LYP/6-31G*	+16.6
B97D/6-31G*	-9.4
CAM-B3LYP/6-31G*	+7.7
LCwPBE/6-31G*	-0.2
wB97XD/6-31G*	-10.4
PM6	-0.2

The binding energy between the two di-hydrogen molecules and the cryptophane-111 host computed at the PM6 level of theory gives surprisingly good results. The value of 0.2 kJ/mol for the binding energy can be compared with the much more sophisticated DFT level corrected for long range interactions the LCwPBE functional. Another surprisingly good result for PM6 level is the orientations of the two di-hydrogen molecules in interior of the cryptophane-111 host cage.

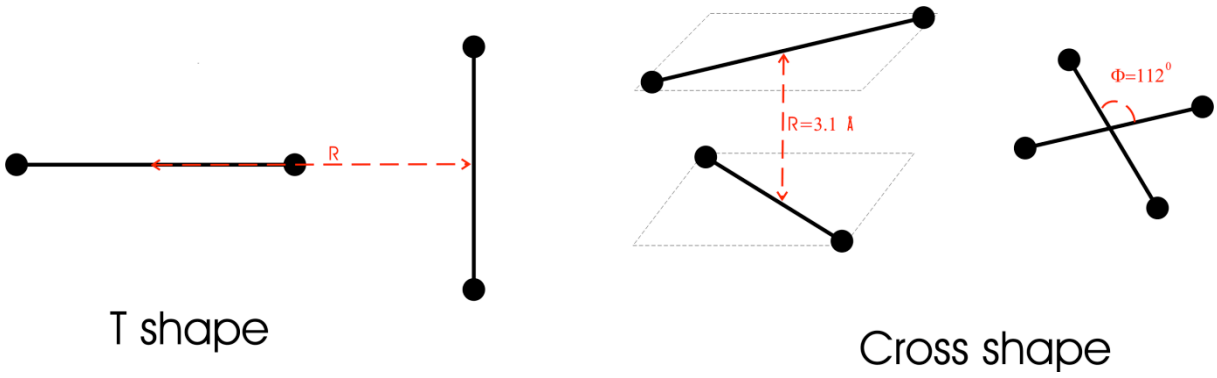


Figure 3.3Arrangement of the two di-hydrogen molecules inside of the crypthophane-111.

Indeed, the arrangement of the two dihydrogen molecules inside the cavity is in a "T" shape, which also corresponds to the optimal geometry of this dimer in the vacuum. The interaction energy of the H2 dimer has been computed at high level of theory (post HF and in

conjunction to complete basis set limits) and the values of this interaction are found to be around 0.3-0.4 kJ/mol for this "T" shape form in ref[15]. After a close inspection of the arrangement of the two di-hydrogen molecules in the interior of the cryptophane-111 cavity we can conclude that all the functionals used except LCwPBE give arrangements in "T" shape between the two di-hydrogen molecules with distance ranging from R=2.79 Å for B97D functional to R=3.42 Å for PM6. On the other hand, the geometry obtains for the arrangement of the hydrogen molecules using LCwPBE functional is in a cross from arrangement (see Figure 3.3).

It is interesting to compare these results with the results obtained for the H_2 dimer at its optimized T-shape geometry. The conclusions are quite similar to those observed previously. B3LYP quite logically does not reproduce this kind of interactions dominated by the dispersion, whereas the newly developed functionals like wB97XD and B97D are able to deal with this type of interactions with a tendency for B97D to overestimate the interaction. We can note that the dependency on the basis set's size is very dependent on the functional used. The B97D functional overestimates the interaction with the largest basis set whereas wB97XD does not show any influence of the basis set for the interaction energy corrected for BSSE.

Table 3.4Interaction energy E_{inter} of H_2–H_2, computed at different theoretical levels.

Method/basis set	E_{inter} [cm^{-1}] (BSSE-corrected)	BSSE
B3LYP/6-31G(d)	+32.8	18.2
B97D/6-31G(d)	-29.7	18.1
CAM-B3LYP/6-31G(d)	+0.6	19.8
LCwPBE/6-31G(d)	+8.2	15.8
wB97XD/6-31G(d)	-26.0	14.0
B3LYP/aug-cc-pVDZ	+20.8	21.5
B3LYP/aug-cc-pV5Z	+24.1	0.1
B97D/aug-cc-pVDZ	-49.7	21.3
B97D/auf-cc-pV5Z	-43.8	0.1
CAM-B3LYP/aug-cc-pVDZ	-5.9	21.9
LC-wPBE/aug-cc-pVDZ	7.5	15.5

wB97XD/aug-cc-pVDZ	-27.2	11.6
wB97XD/aug-cc-pV5Z	-29.8	0.2
MP2/aug-cc-pVDZ[6]	-19.5	
MP3/aug-cc- pVDZ[6]	-25.8	
MP4/aug-cc- pVDZ[6]	-26.4	
FCI/ aug-cc- pVDZ[6]	-26.4	
CCSD(T)/aug-cc- pVDZ[6]	-27.1	

3.3.2 Cryptophane-A host and CH4 guest

For this system, we deduce the binding energy between the host guest systems only by metadynamics simulations. The binding energy is deduced by the difference between the associated form and the dissociated form. The parameters used for metadynamics simulations are:

- Two collective variables: the first one is the distance between the center of the mass of the cryptophane and center of mass of CH_4 molecule and the second one is the potential energy of the system.
- Integration time step is set at 0.5 fs and the total simulation time was 7.5 ps.
- The $\tau_G = 5\,fs$
- The height of the Gaussians is set to $w = 0.1$ kcal/mol, which represents 0.418 kJ/mol. The width of the CV1 and CV2 are $\delta s_1 = 0.5$ Bohr and $\delta s_2 = 5$ kcal/mol ~ 21 kJ/mol respectively.
- The method used is PM6.

It must be mentioned that during the metadynamics simulations, we set at 25 Bohr from the center of the mass of the Cryptophane an artificial wall to prevent the ligand from slipping away from the box. This wall inverts the velocity vectors of the atoms when they reach this limit of 25 Bohr ~ 13 Å.

[6] Reference [15]

The simulation is stopped when CH_4 has crossed two times from the cage to outside of the cage. The free energy surface is represented in the next picture.

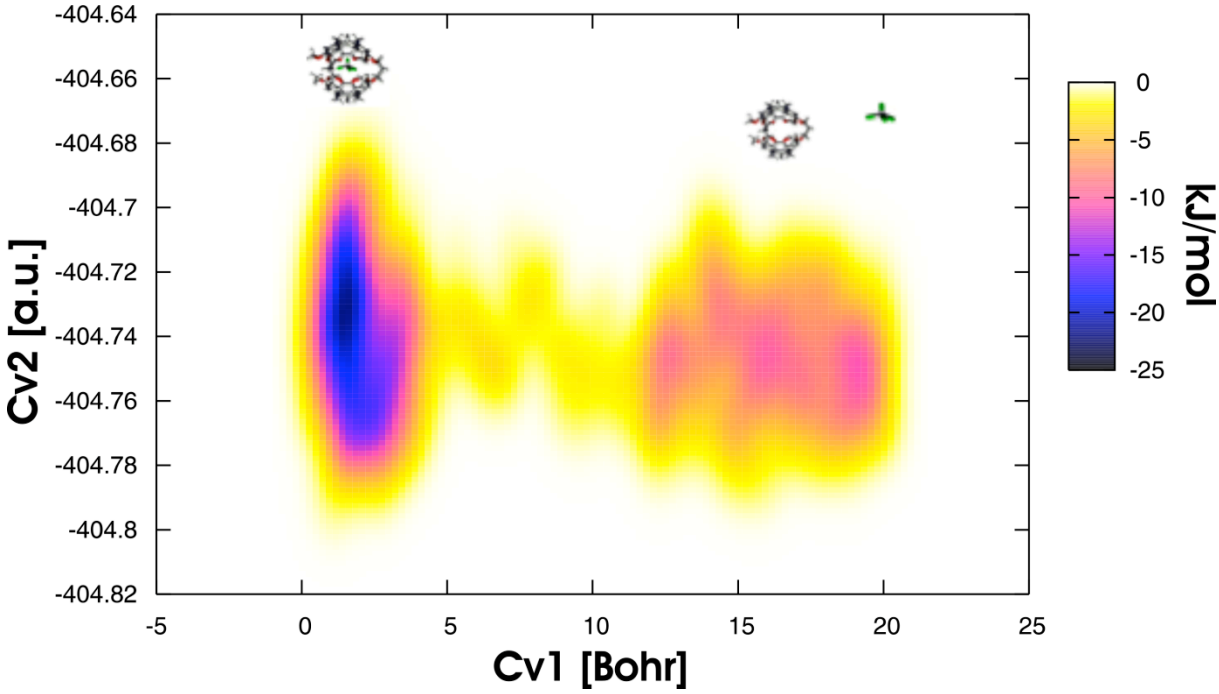


Figure 3.4Free energy surface representation of the CH4@Cryptophane-A as a function of the CV1(distance between the centers of masses of the Cryptophane and CH_4) and CV2 (Potential Energy)

After a close inspections of the results, we find an association energy of around -12 kJ/mol which is in excellent agreement with the experimentally results of -11.1 kJ/mol[14].

3.3.3 Interaction energy between Cryptophane-C and HCFClBr enantiomers.

One of the first objectives of the Crypthophane project was to design a molecule in order to discriminate between the enantiomers of small molecules and CHFClBr molecule is one of the smallest chiral molecule. Collet et al. were the first to demonstrate this ability of the cryptophane to complex small chiral molecules.[17]

Figure 3.5Enantiomers R **and** S **of the bromochlorofluoromethane molecule**

Increasing interest in this topic is due to new strategies to obtain optically pure compounds, and also to the recent investigations of their chiroptical properties and to the design of new materials catalytically active[17-21] Experimentally, from the 1H-NMR spectra, the chiral discriminations energy has been estimated at $\Delta\Delta G_{exp}$ = 1.1 (±5%) kJ/mol in favor of the **R** form complex with respect of the **S** form complex[12]. Using MD combined with MC simulations for the refinements of the thermodynamically parameters, McCammon and co-workers[22] find a $\Delta\Delta G_{calc}$ = 1.51 ± 1.41 kJ/mol and they concluded that the **R** form is more stable relative to**S** form with a value in agreement with the experimental one.

Using several functionals, we test if they are to describe experimental results obtained for the ΔΔG. The results are presented inTable 3.5. From these results we can see that all the functionals used herein correctly describe the right order of these two enantiomers **R** and **S** (**R** being more stable than **S** in the criptofane-C complex). Only PM6 reverses the order between the two enantiomers of the bromochlorofluoromethane molecule. After a closer inspection of the results from Table 3.5 we can see that all the functional augmented by dispersion corrections over-estimate the interactions in favor of the **R** form. Only B3LYP is closer to the experimental results, over-estimating with 2.5 kJ/mol the experiment. If we compare the results of B3LYP and the results of CAM-B3LYP we find a difference of 4.9 kJ/mol, which represents the long range corrections of the energy computed at B3LYP level of theory.

Table 3.5Relative energies of R **and** S **form @ Cryptofane-C**

Method	ΔE(R@ Cry-C) kJ/mol	ΔE(S@ Cry-C) kJ/mol	ΔΔE(R-S) kJ/mol
B3LYP/6-31G*	-25.0	-21.4	-3.6
B97D/6-31G*	-136.0	-127.0	-8.9
CAM-B3LYP/6-31G*	-50.0	-41.5	-8.5
LCwPBE/6-31G*	-47.3	-40.9	-6.5
wB97XD/6-31G*	-129.8	-122.5	-7.3
PM6	-42.8	-44.4	+1.5
Experimental data[7]			ΔΔG~ -1.1

Because of the relatively small 6-31G* basis set used to evaluate the energy with all these functionals, it is important to estimate the value of the BSSE for these calculations and the relative energy corrected for this error. The results after BSSE corrections using counterpoise methods are presented inTable 3.6. After BSSE corrections, the relative order between the **R** and **S** form has changed for the B3LYP functional and for LCwPBE, and they present an inversion of the relative order for these two conformers. Interestingly enough, these are the two functionals, which were the most unfitted to deal with dispersion energy in the former studies of this chapter. The over estimations of the interactions between the **R** and **S** form are reduced, comparatively with the non corrected ones, and the results are closer to the experimental values of 8 kJ/mol.

The functionals that are the most suitable to be used are the B97D and wB97XD. Indeed they give the better agreement with the experiences and have the physical background to deal with dispersion. These give a relative difference with the experiment of only 1.6 kJ/mol and 1.5 kJ/mol respectively for the ΔΔG value. Thus these are most suitable to reproduce the correlations between the cryptopfane and the ligand (HCFClBr enantiomers molecules) and must be used in the future studies.

All these calculations reveal the importance of the dispersion energythat is indispensable to obtain reasonable results for these kinds of systems characterized by weak interactions. Because the BSSE presents a non negligible part of these simulations it is mandatory to

[7] Ref [12]

reduce these errors for these calculations. The main conclusion is that we can predict the behavior of another's small molecules only if we use an appropriate functional which can correctly predict the dispersion interactions and with an appropriate basis set which reduce the BSSE errors.

Table 3.6Basis set superposition error corrections to the relative energies of S **and** R **form @ Cryptofane-C**

Method	ΔE_{corr} (R@Cry-C) kJ/mol	BSSE kJ/mol	ΔE_{corr} (S@ Cry-C) kJ/mol	BSSE kJ/mol	$\Delta\Delta E_{corr}$ (R_{form}-S_{form}) kJ/mol
B3LYP/6-31G*	46.6	71.6	44.3	65.7	+2.3
B97D/6-31G*	-52.8	83.2	-50.1	76.9	-2.7
CAM-B3LYP/6-31G*	27.4	77.4	27.7	69.2	-0.3
LCwPBE/6-31G*	24.9	72.2	24.1	64.9	+0.8
wB97XD/6-31G*	-52.5	77.3	-49.9	72.6	-2.6
Experimental data[1]	ΔG~ -8 kJ/mol				-1.1

Using metadynamics simulations we were also able to determinate the relative encapsulation free energy between the Cryptophane-C and the two enantiomers **R** and **S**. The metadynamics parameters are:

- Two collective variables: the first one is the distance between the center of the mass of the cryptophane and center of mass of the two enantiomers **R** or **S** and the second one is the potential energy of the system.
- Integration time step is set at 0.5 fs and the total simulation time was 10 ps.
- The $\tau_G = 5\,fs$
- The height of the Gaussians is set to $w = 0.1$ kcal/mol, which represents 0.418 kJ/mol. The width of the CV1 and CV2 are $\delta s_1 = 0.5$ bohr and $\delta s_2 = 5$ kcal/mol ~ 21 kJ/mol respectively.
- The method used was PM6.

The binding energy was calculated using the differences between the free energy surface of the system with **R** or **S** inside of the cavity and when the **R** or **S** is outside the cavity. The $\Delta F(R@CryC) = -19kJ/mol$ and $\Delta F(S@CryC) = -13kJ/mol$. The $\Delta\Delta F(R_{form}\text{-}S_{form})$ = -5.5 kJ/mol and in comparison with the experimental results which is find to be -1.1 kJ/mol is in good agreement.

It is interesting to note the difference observed between the results of these metadynamics simulations ΔF and the result obtained through optimizations. Indeed, the value obtained from metadynamics is clearly inferior to the one obtained by the difference of energy. It can be explained by the inclusion of the entropic contribution in the thermodynamic functions F or G. Indeed, the translation entropy is expected to increase the cost of the association's process[23]. This argument is sustained by static calculations of ΔG where we obtain -16.8 kJ/mol and -15.1 kJ/mol for the binding free energy between the **R** and **S** enantiomers respectively to Cryptophane-C complex calculated at same level of theory like metadynamics (PM6). These results comparatively with the results obtain for ΔF after metadynamics simulations are in excellent agreement with static results of ΔG.

3.4 Parameterization and implementation in Gaussian 09 framework for Platinum atom of the empirically long – range dispersion term following Grimme's work.

Grimme's method, described in the ref [1], was followed to determine the coefficients C_6 and R_0 for Pt atom. Using this approach we calculate the ionization potential Ip and static dipole polarizabilities α at DFT/PBE0 level of theory using QZVP basis set for the Platinum atom. The Grimme's equation $C_6^{Pt} = 0.05NI_p^{Pt}\alpha^{Pt}$(eq. 14 from ref [1] was used to find the C_6 coefficients, which uses the vertical ionization potential. It can be approximated by the absolute value of the HOMO orbital. The static electric polarizabilty α expresses the capacity of the charge density of a system to be distorted by an external electric field. Usually, only the average value of its diagonal is given: $\langle\alpha\rangle = \frac{1}{3}(\alpha_{xx} + \alpha_{yy} + \alpha_{zz})$ which gives the value for α^{Pt} used in the equation of the C_6 coefficient. For the Pt atom we find the C_6^{Pt} coefficient equals to 42.44 J nm^6 mol^{-1} and the R_0^{Pt} coefficient equals to 1.75 Å.

The implementation of these new coefficients in the G09 source code is not so difficult because these coefficients are kipped in the utility routines as separate data in two arrays,

which are part of the ***R6DC6*** and ***R6DR0*** functions for C_6 and R_0 respectively. These two functions are double precision functions and can be found in the "**utilnz.F**" file. In the same utility file, it can be found the ***R6DGrm*** double precision procedure, which is responsible for the computation of the empirically dispersion corrections to energy. This procedure calculates the Energy, the first derivatives and the second derivatives of the energy with these corrections counted. The procedure structure is highlight in the next logical structureFigure 3.6.

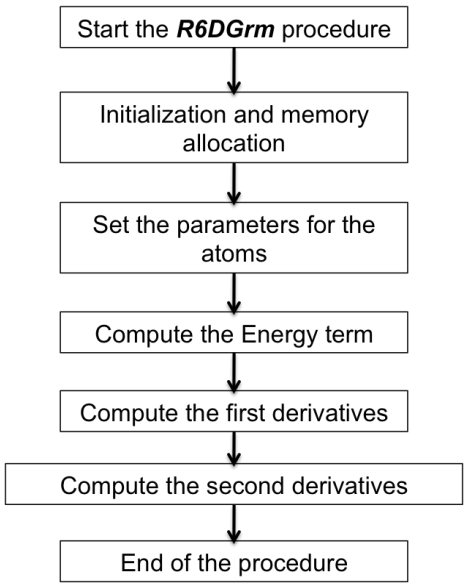


Figure 3.6The logical diagram of the *R6DGrm* procedure

In the first revision (A.02) of the G09 software only a few functional are implemented which can be used with the Grimme's dispersion corrections parameterizations. But the implementations of new one are not difficult because the routine that is responsible for choosing the functional and allocate the parameters for the damping function proposed by Grimme is well structured. This routine is ***R6DPar*** and it can be found in the same utility file "utilnz.F" and the parameters are listed in the Table 3.7

Table 3.7List of functionals implement in Gaussian 09 revision A.02 software, which can use the empirically atom pair corrections for the dispersion forces.

Functional	S6 coefficient
B97	1.25
B2PLYP	0.55

mPW2PLYP	0.40
PBE	0.75
BLYP	1.20
B3LYP	1.00
BP86	1.05
TPSS	1.00
wB97X[8]	1.00 and d=6.0

From the list presented in Table 3.7only the wB97X functional uses a different form of the damping function, all others being implemented with Grimme's damping function.

To use explicitly in a G09 job these new options of empirically long range corrections for the dispersion term, the IOP() keywords in root section of the job must be used. Thus, **IOP(3/124=3)** enables the option responsible for the empirically dispersion term and for the debugging purpose only. The keywords **IOP(3/33=3)** and **IOP(3/136=10)** enable the printing flag to print, in output file, all parameters used in an human readable format. The use of the IOP options is raised by the fact that Gaussian framework is a collection of independent programs, named links, which use internal files (the so called read write files) to communicate between them. The link responsible for the calculation of empirically dispersion term is link 301, which is also responsible to generate the basis sets information and to set the functional parameters for a DFT job.

This parameterization of the C6 and R0 for Platinum atom is important in the field of organometallic chemistry where platinum complexes are used as catalyst and the proper description of the interaction between organic ligands and Pt complex may be crucial for the understanding of the mechanism involved.

3.5 Platinum complex [2+1] cycloadition reactions

Recently, it has been shown that Palladium (II) and Platinum(II) complexes stabilized by secondary phosphine oxide ligands catalyze a very unique [2+1] cycloaddition of norbornene derivatives with various terminal alkynes to give functionalized alkylidenecyclopropanes[24-26].

[8] This functional use a different damping function as it is described in the work of Chai and Head-Gordon for more details see ref [11]

The question about the mechanisms is still an open question and some experimental Mass Spectrometry data as well as theoretical work suggest different mechanisms depending on the metal used, Pd or Pt[27]. The specific role of the ligand compared with the Herrmann-Beller complex is still at stakes[28]. So in this part of the chapter, we will present the theoretical work we have performed on the mechanism with the Platinum complexes. Our work has focused mainly on the mechanisms involving not the norbornadiene but the O-tBu norbornadiene.The reactivity observed with this substituted norbornadiene is indeed different from the one observed with the norbornadiene. Due to the possible role of weak interactions and to gain a better description of the ligand/metal interaction, we have implemented a version of the B97D in the Gaussian09 program for the platinum atom as previously discussed.

So the first paragraph of this part will be devoted to the presentations of the different mechanisms postulated, then we will present the results when the OtBu substituted norbornadiene is used then we will discuss the data we have on the mechanism with the norbornadiene.

3.5.1 General Mechanisms

In a first article[24], the authors have shown that the use of Palladium (II) and Platinum(II) complexes stabilized by secondary phosphine oxide ligands catalyze a [2+1] cycloaddition of norbornene derivatives to lead to alkylidenecyclopropanes, see Scheme below.

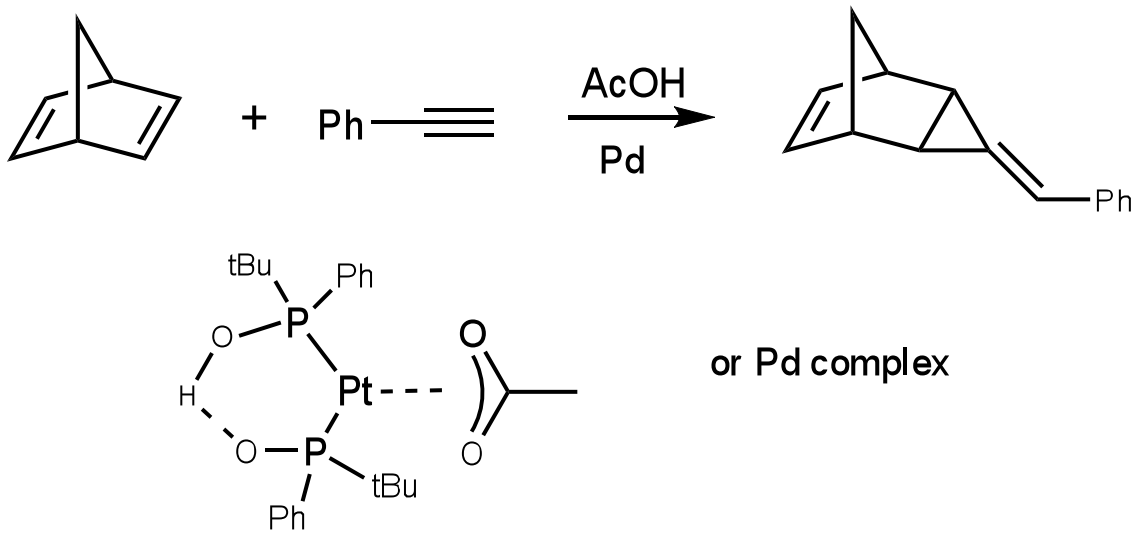


Figure 3.7 Pd(II) [2+1] cycloaddition reaction for norbornene derivatives

On the contrary when OtBu substituents are used, the reactivity observed with Pt complexes

is quite different, as shown in Figure 3.8.

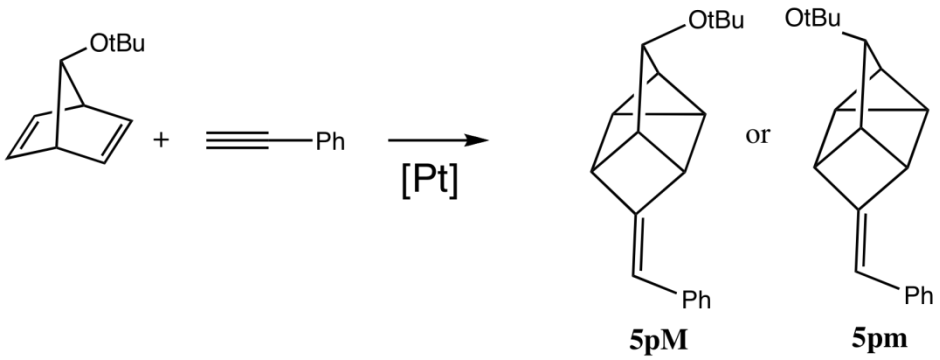


Figure 3.8 Mechanism of norboradiene substituted with OtBu

The reogiochemistry is particularly well defined and the product **5p** is obtained with a high yield 64% and high selectivity of 70% for 5pM that we will denote *trans* product.

First, we have estimated the relative stabilities of the different possible products with and without the OtBu substituent. We have taken into account the different possibilities and the results are gathered in the figure below.

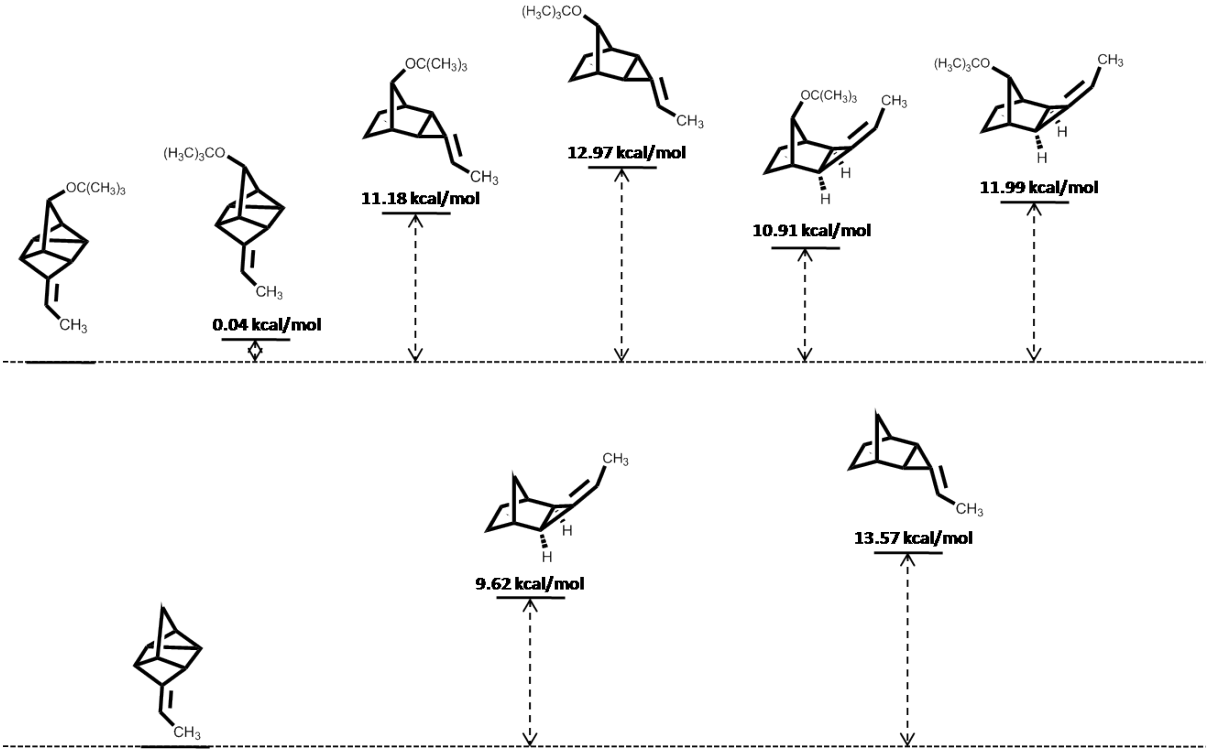


Figure 3.9 Relative energy chart for different possible products of OtBu-norbornadiene or norbornadiene

With or without the OCH3 substituent, the **5p** type product is more stable than the benzylidenecyclopropane by 10kcal/mol. Moreover, the endo product is found more stable than the exo product by 2-4 kcal/mol.

This reaction is particularly intriguing and the mechanism has raised many questions. Even if we will focus on the Pt complex catalyst, it is worth noting a similar reaction known as the Hermann-Beller (HB) reaction with a phosphapalladacycle[28]products a 5-alkynylbicyclo[2.2.1]hept-2-ene. Both these reactions need acetic acid to be effective.

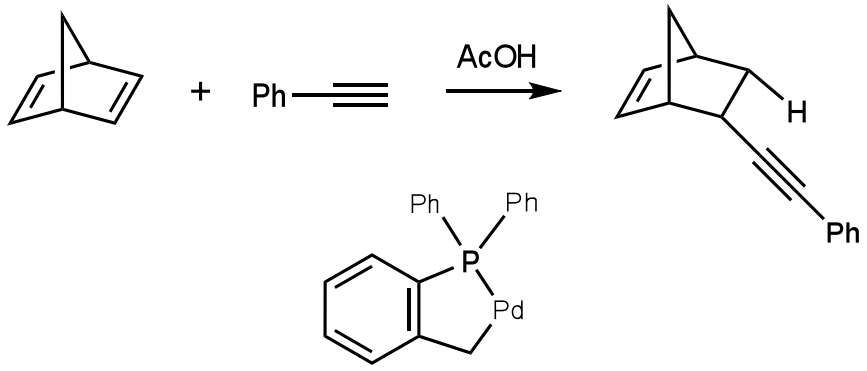


Figure 3.10 Model of a Hermann-Beller reaction

Among the proposed mechanisms shown below, the blue one is first similar to that proposed for the HB reaction with the formation of an acetylenure ligand, whereas the red one proposes the formation of a carbene ligand through a H-shift. Numerous investigations have been performed to elucidate the mechanism and particularly mass spectrometry analysis. Even if many questions are still open, from the last experimental data[27]available, it seems reasonable to assume a different pathway for the Pd and Pt complexes. For the Pt complexes, the formation of the carbene ligand seems the viable route. It is this assumption we have made to investigate the reaction with the OCH3 substituent.

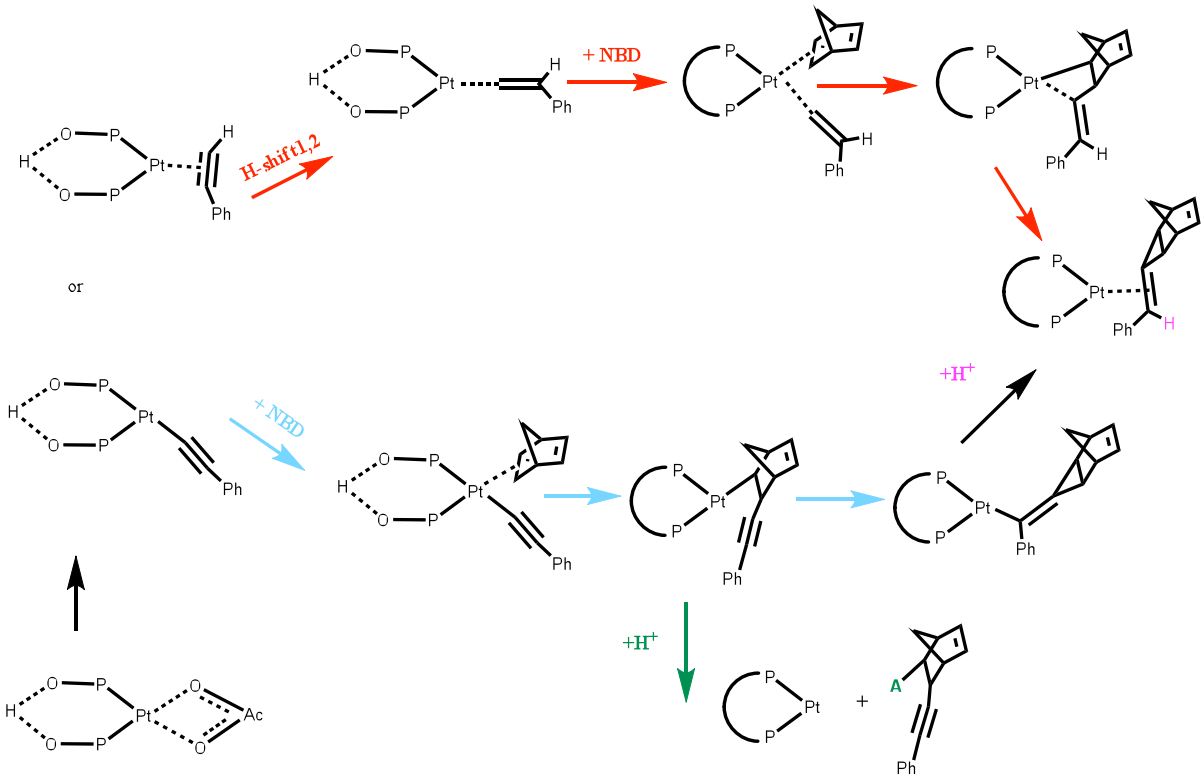


Figure 3.11 Proposed Mechanisms for the reaction of the Pt(II) complex with norbornadiene

Thus, we propose the following mechanism starting from the complex with the carbene and the norbornadiene ligand.

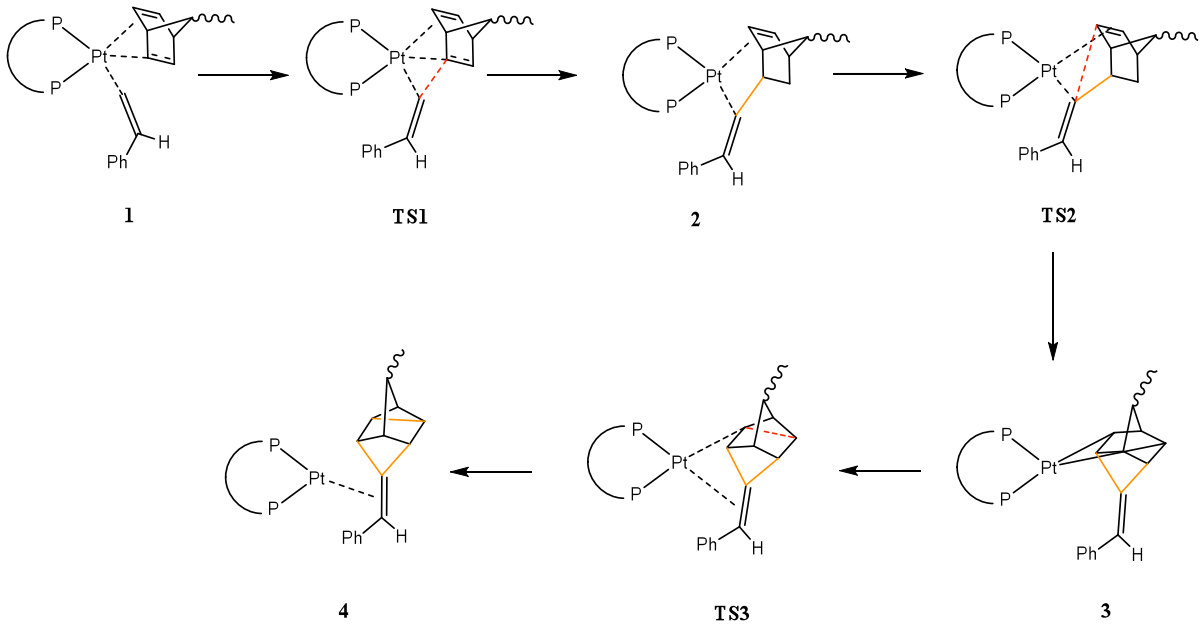


Figure 3.12 The mechanism proposed for the Pt(II)-complex with the carbene and norbornadiene ligands

The complex **1** is the complex with both the norbornadiene and the carbene already coordinated. From 1, the first carbon-carbon bond can be formed to give complex **2**, then the second carbon-carbon bond is formed leading to the clycloropane pattern. Finally, the TS corresponding to the last carbon-carbon formation connects complex **3** to the complex **4**.

Next, we will describe this pathway and will try to answer the following questions:

- Is this pathway feasible?
- Which step controls the selectivity and how?

3.5.2 Mechanism:

To investigate the main features of the mechanism, we used a complex with the secondary phosphine oxides (SPOs) ligand{[(*t*-Bu)(Ph)PO]$_2$H} modelized by a {[(Me)PO]$_2$H} ligand. More controversially, Ph substituent of the alkyne was modeled by a methyl group. Additional calculations were performed on the real complex and will be discussed below. Except when indicated otherwise, the calculations have been performed with the Gaussian03[29] code using the B3LYP[4,5] functional with the lanl2dz[30,31] basis set with added polarization functions on the P, C, O atoms[32]. First, we will describe the reactivity using the model complex

First CC bond formation

As shown on the reactive pathway, we start from a Pt(II) complex with the norbornadiene and the carbene as ligands and of course the SPOs ligand. But on the contrary to the general mechanism proposed with the norbornadiene, we do not start our mechanism from a η^2 coordinated norbornadiene but from a η^4 coordinated norbornadiene. It seems a reasonable hypothesis the OtBu is definitively a bulky ligand, which may prevent the association of the norbornadiene in a η^2 mode with an agostic interaction.

Due to the presence of the OtBu substituent and of the Ph susbituent of the alkyne, 4 pathways (see Figure 3.13) have indeed to be studied. If the dissymmetry due to the phosphine ligand is taken into account, which is not the case with our model ligands, 8 pathways should be computed.

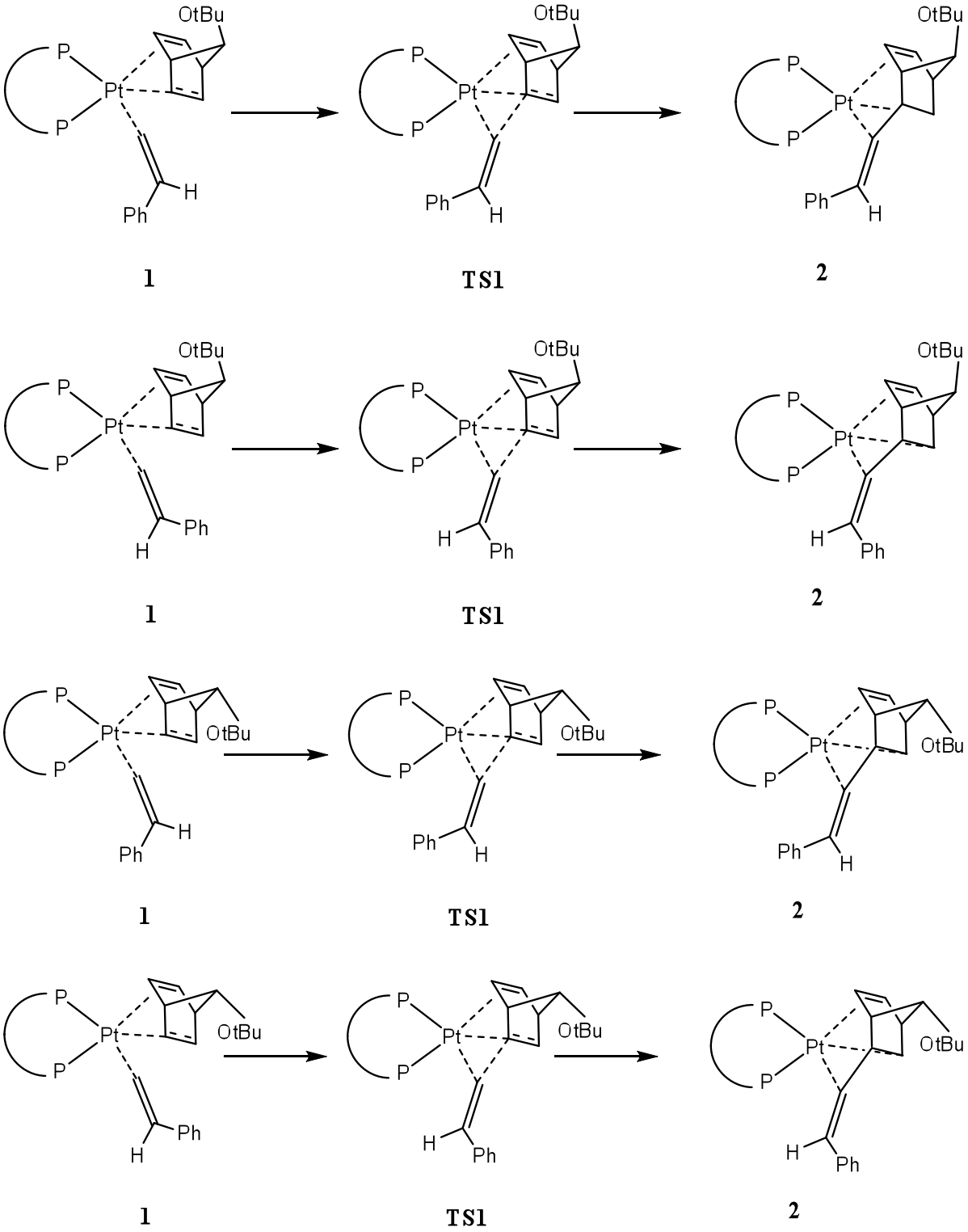
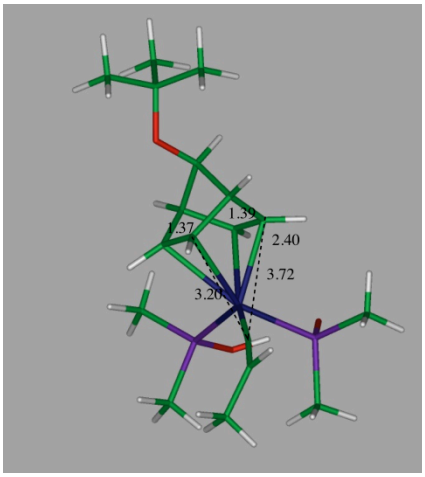
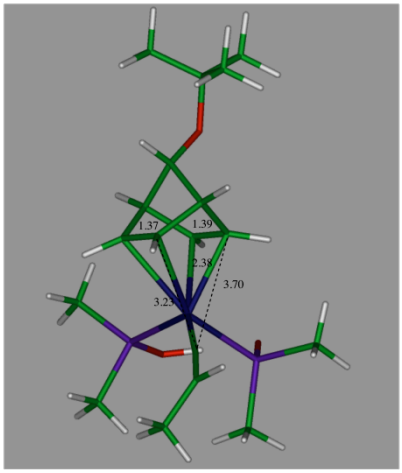


Figure 3.13 Four possible models for TS1 complex

All the 4 **TS1** exhibit small barriers around 5-8 kcal and the complex **2** is strongly stabilized by around 20 kcal/mol.



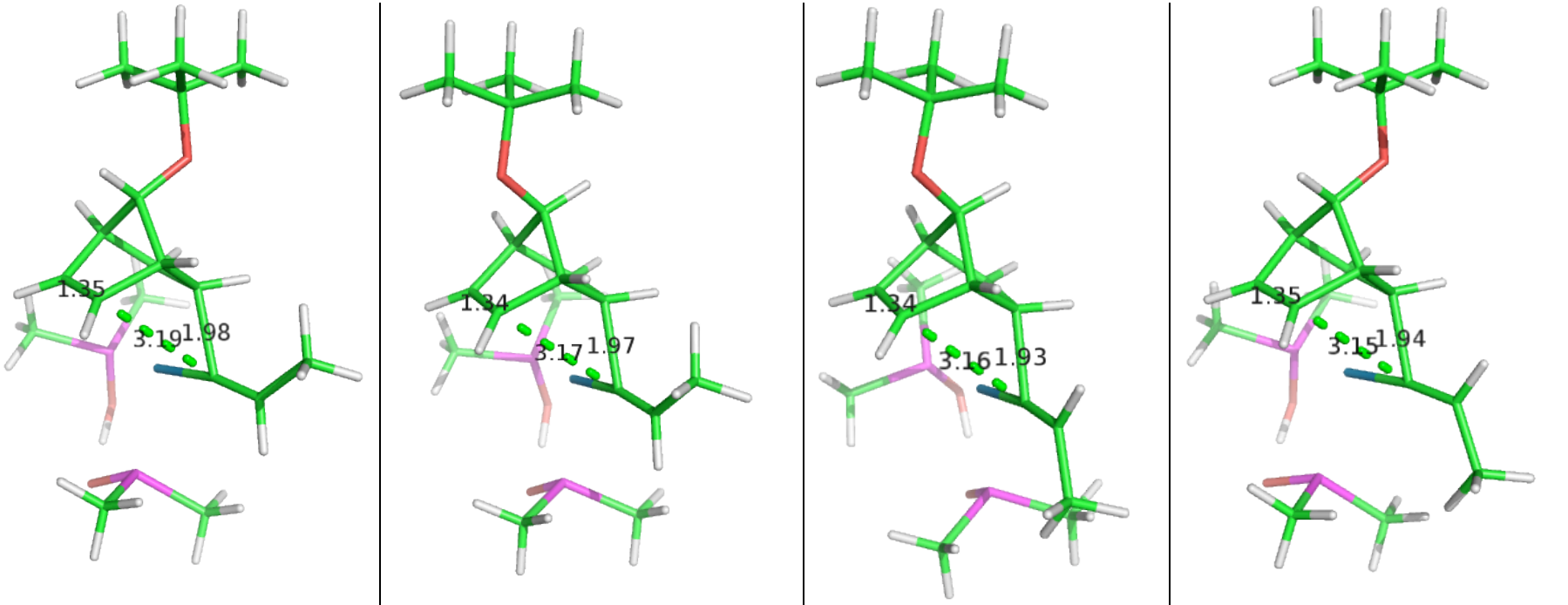


cis	*trans*

Figure 3.14 Starting complex, η^4 coordinated

Two of the starting complexes are shown in Figure 3.14. The distance between the Pt and the carbons of the norbornadiene are around 2.30-2.40 Å and the distance between the two carbons of the first bond to be formed is around 3.2 Å.

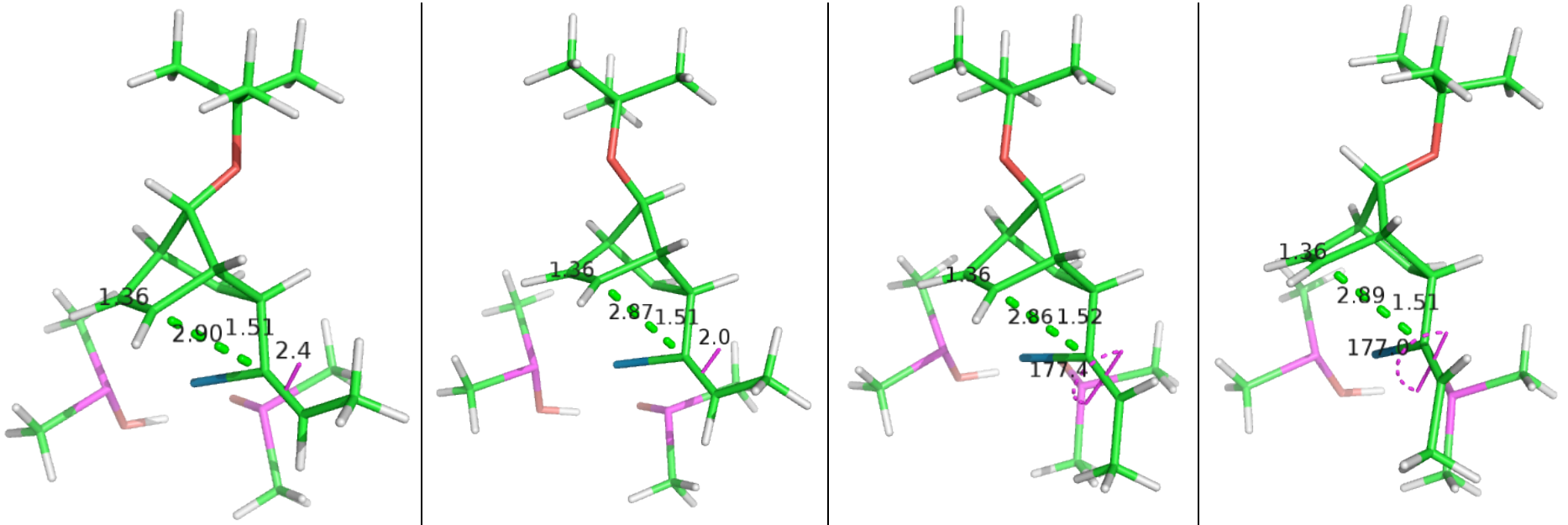
The corresponding TS are shown below and they exhibit similar C-C bond around 1.9/2.0 Å. The longer distance is obtained when the first C-C bond is formed with the carbon of the carbene substituted by the methyl.



b	d	c	a

Figure 3.15 Geometries of the transition state 1 (TS1) for all four models proposed. Distances are measured in Angstroms and angles are in degrees.

When the first CC bond is formed, complex **2** are obtained. In this geometry, the substituents of the carbene adopt an up/down position.

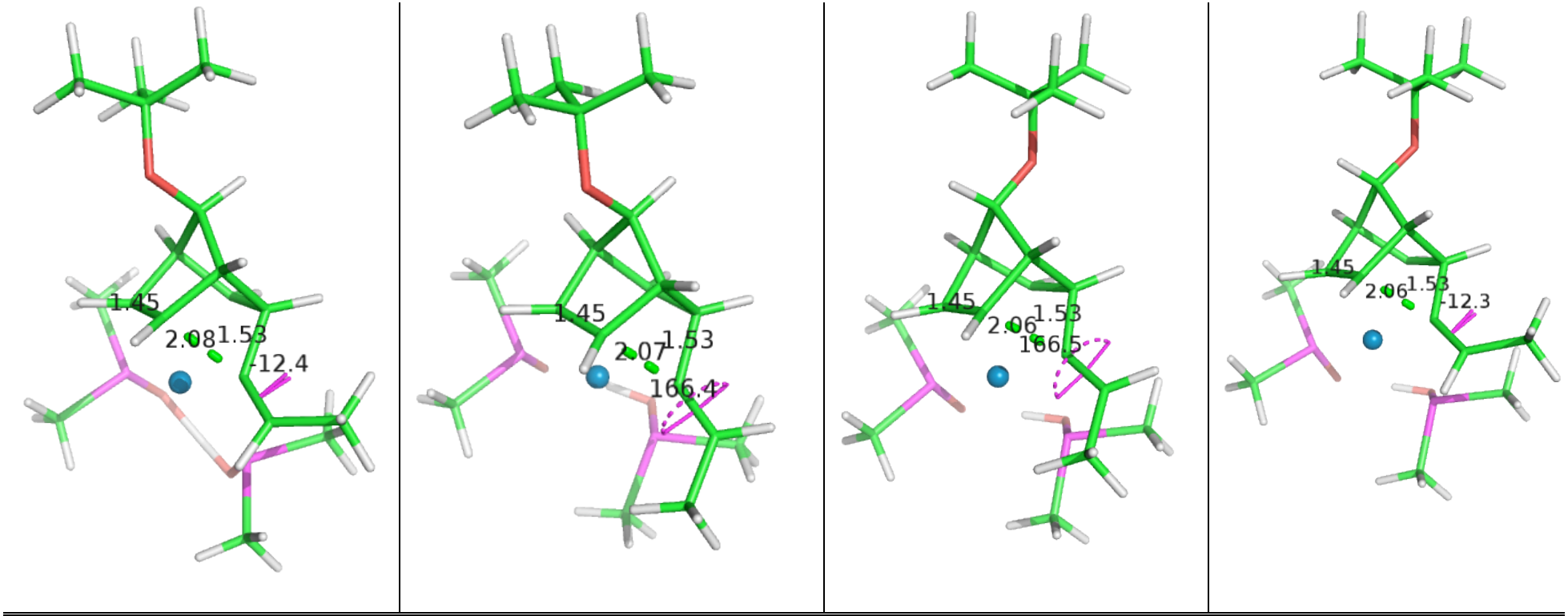


b	d	c	a

Figure 3.16 The geometry's of the firs CC bond formation of these models. Distances are measured in Angstroms and angles are in degrees

Second CC bond formation

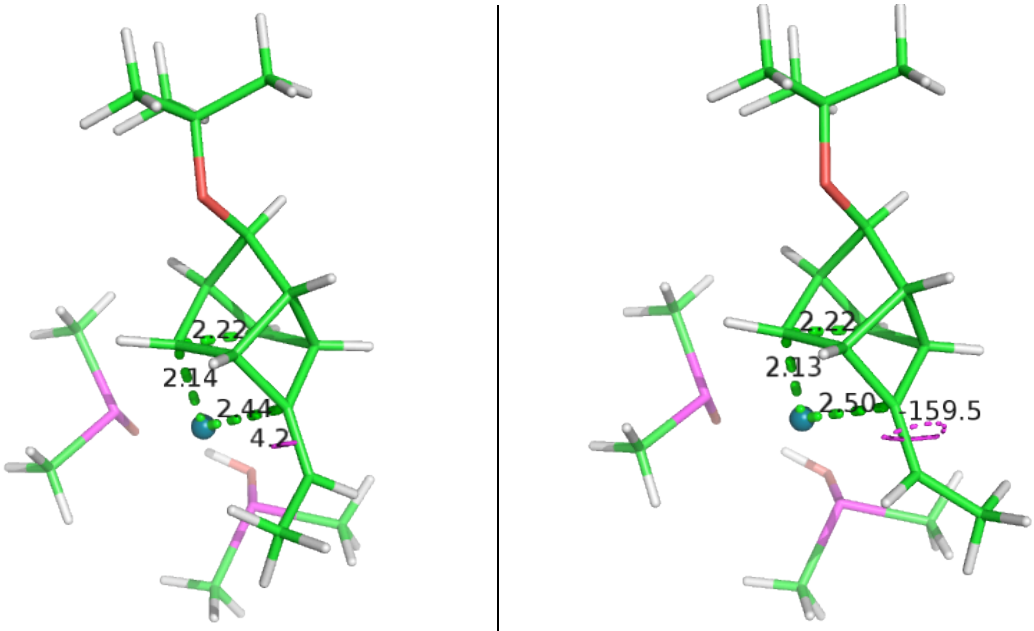
The barrier for the second bond formation is higher around 20 kcal/mol and the general process is slightly favorable since complex **3** is around 1kcal/mol more stable than complex **2**.



b	d	c	a

Figure 3.17 Geometry's of the second CC bond formations for these models Distances are measured in Angstroms and angles are in degrees

All the TS exhibit a tilt in the same direction, which will be explained later. The complex obtained have the cis/trans isomery definitively set. The resulting complexes **3** have very short distances between the two carbons, which will give the last CC bond. For the CC to be formed the bond distance is around 2.2 Å and the Pt-C bond around 2.15 Å.

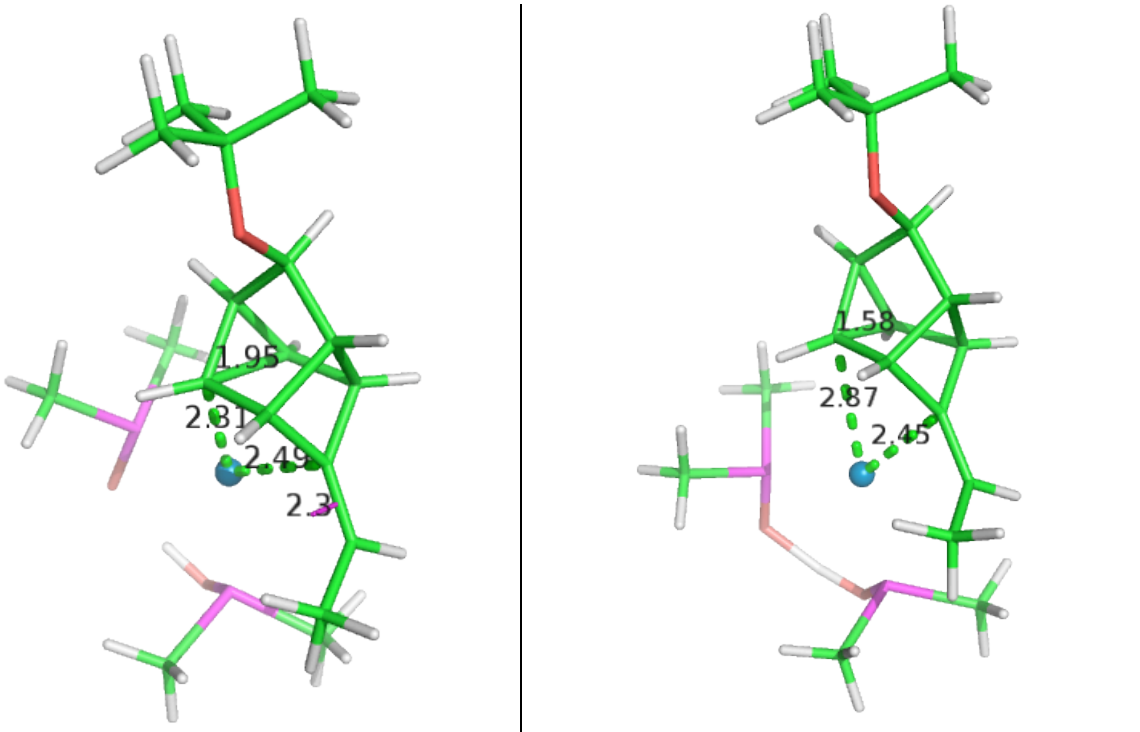


Cis	Trans

Figure 3.18 Cis and Trans conformations for complex 3. Distances are measured in Angstroms and angles are in degrees

Third CC bond formation

Finally, in this paragraph, we will study the last CC bond formation. The barrier is very small around 1-2 kcal/mol and leads to the product 10 kcal/mol below the **TS3**. The low barrier is consistent with a TS structure with a critical CC bond distance of 2 Å quite close to the value in the starting complex 2.2 Å. The Pt-C bond distances increases from 2.2 Å in the complex **3** to 2.3-2.35 Å in the **TS3**. From a general point of view, the product **4** is 25 kcal/mol lower in energy than the starting complex **1** and the product is linked to the Pt through the double bond since the Pt-C bond involving the carbons implied in the last bond formation is now around 2.9 Å.



TS3	4

Figure 3.19 Geometry conformations for TS3 and 4 complexes.Distances are measured in Angstroms and angles are in degrees

Conclusions

The presentation of the results starting from complex **1c (***vide infra* for definition) are represented below. We can notice that the pathway seems reasonable with the barriers in agreement with experiments at 60°C with a yield of around 65% We can also note that the highest transition state in energy is indeed the first one.

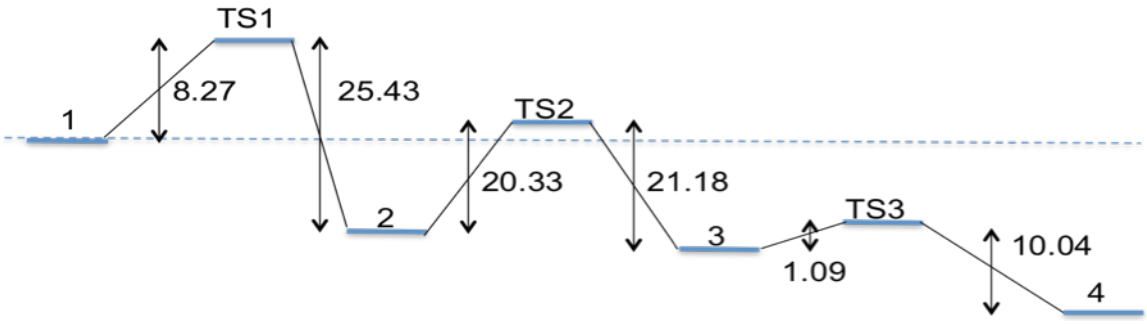
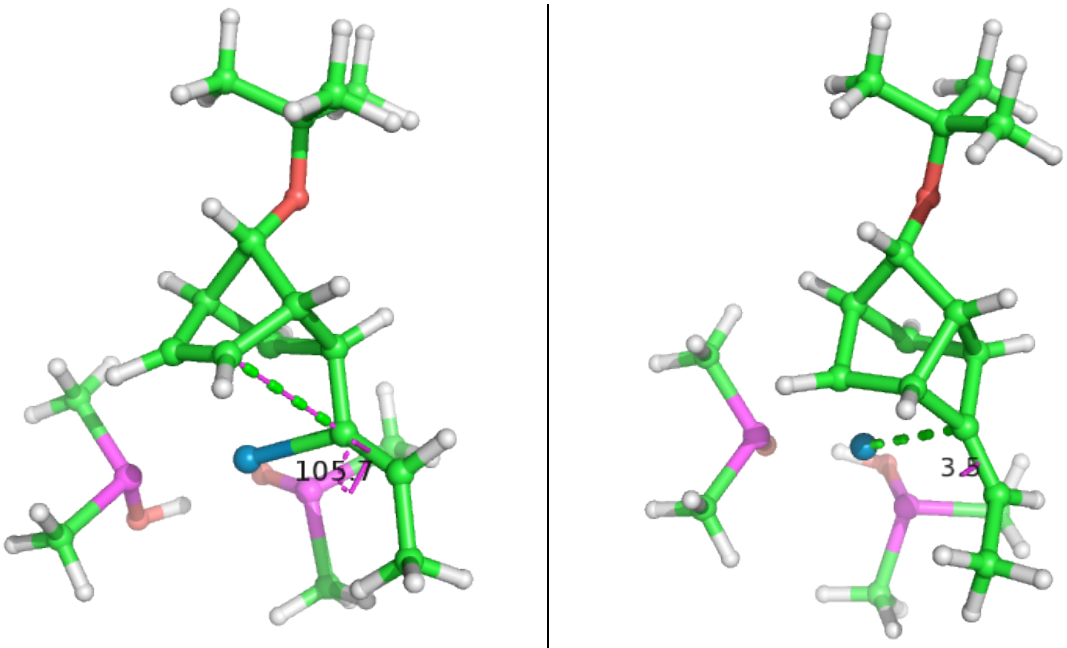


Figure 3.20 Relative energy path for model 1c. The energy is express in kcal/mol

3.5.3 Selectivity

From an experimental point of view, **5pM** is the major product and the first question, which jumps out, is where is the regiochemistry determined? In fact, a closer look at the structures **2** and **3** show that in complex **3** the regiochemistry is determined whereas in complex **2** the conformation of this complex seems to indicate that the conformation is still not determined. In complex **2**, the substituent H and Methyl of the alkene are indeed in a up/down position versus the OtBu substituent and not in the cis/trans position obtained in the final product.



complex 2 with H/Me in up/down position	complex 3 with H/Me in cis/trans position

Figure 3.21Molecular conformations of the complex 2 (left) and 3(right). Dihedral angle in degrees

Indeed, during the formation of the second carbon-carbon bond, a tilt of the double bond is observed and will lead to complex **3** with Me substituent either in a *cis* or *trans* position. Thus, two TS for the CC bond formation should be observed one with a direct tilt and the other one with the indirect tilt. But the energy between these two TS is quite high around 25 kcal/mol and so only the direct tilt is observed.

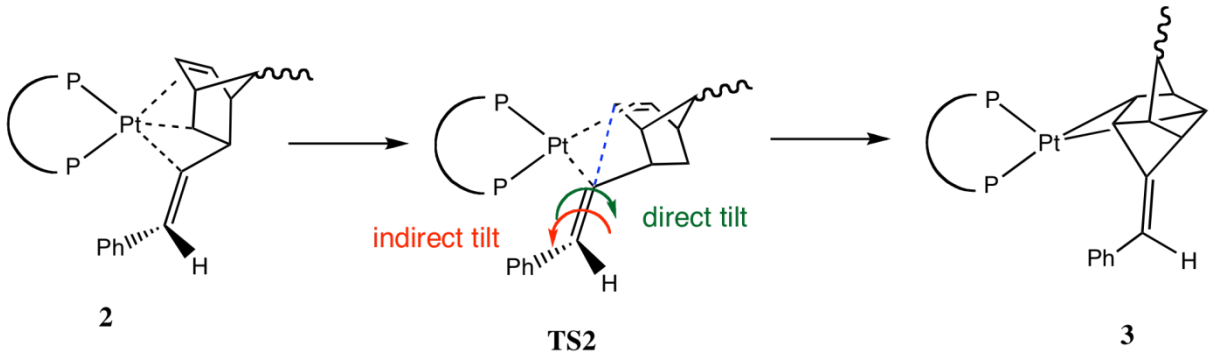
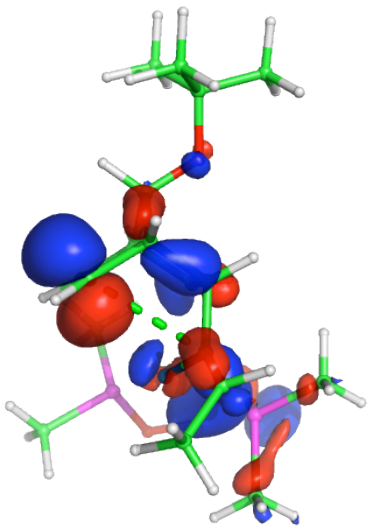


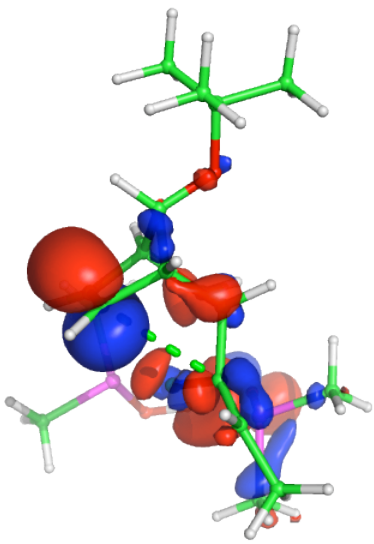
Figure 3.22 The second transitions state

Thus, the cis/trans conformation is not determined by the second step but indeed after the first C-C bond formation, the final conformation is already set. So when complex **1** is obtained the resulting regiochemistry is already determined.

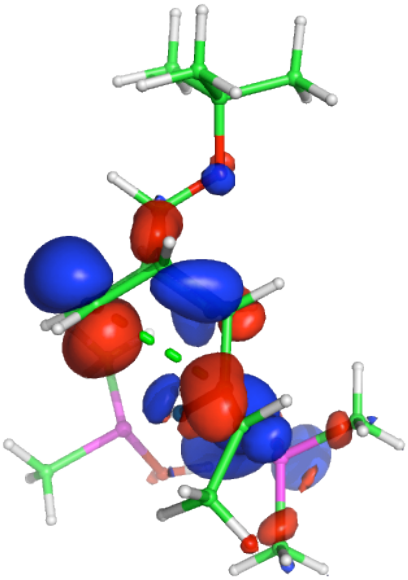
To explain why only the direct tilt is observed, we need to look at the Molecular Orbitals of the complex. Indeed, the easiest way to see this trend is to consider a direct and indirect tilt in our case of 45° and 30° from complex **2**. Thus the representation of the HOMO-1 orbital on the left clearly shows a bonding interaction between the two π systems, which are going to form a bond. Whereas on the representation on the right, we can clearly see that with this indirect tilt the overlap between the two π systems tends to vanish. These MO clearly explain the selectivity observed during the second carbon-carbon bond formation.



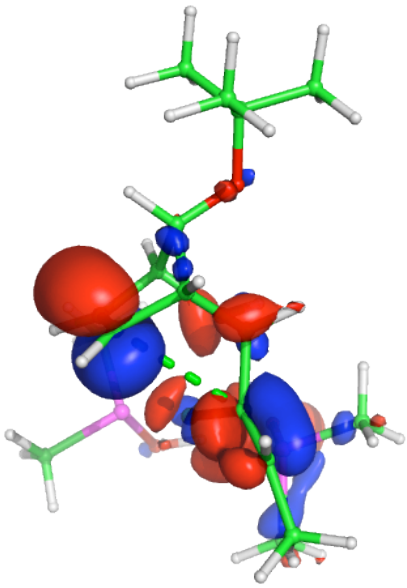
Direct tilt 30°



Indirect tilt 30°



Direct tilt 45°



Indirect tilt 45°

Figure 3.23 HOMO-1 molecular orbital for direct (left) and indirect (right) tilt model.The electron density isosurface is 0.02 a.u.

This means that the first bond formation may be the one leading to the selectivity. Nevertheless the low barrier around 5-8 kcal/mol and the experimental temperature 60°C is not really in favor of this hypothesis. But on the other hand, once complex **2** is formed, going back to complex **1** seems difficult since the barrier to come back to complex **1** fromcomplex **2** is around 25 kcal/mol.

To discuss, precisely the barriers, we will use the notation described below in the Figure 3.24:

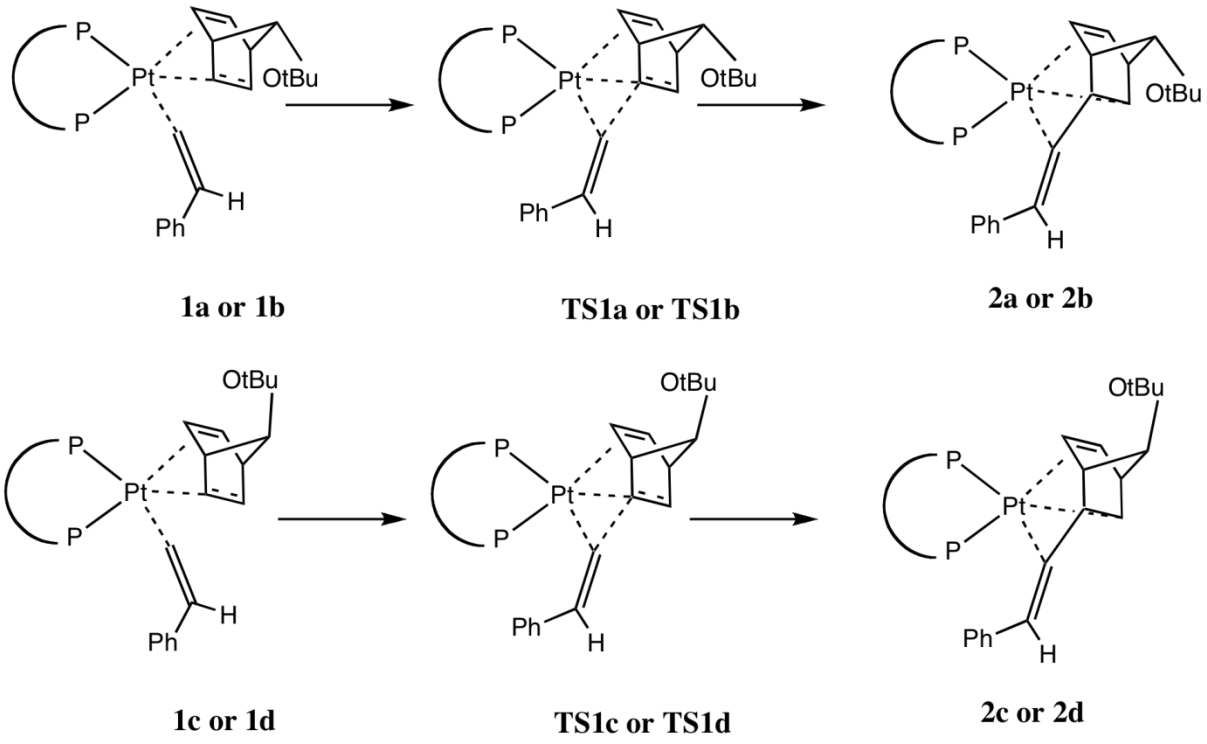


Figure 3.24 Transitions states complex models

We will denote **TS1a** or **TS1b** the transition state structure with the double involved in interaction with the OtBu substituent of the bridge of the norbornadiene motif. We will call **TS1a** the TS structure with the H in the up position. Similarly, **TS1c** and **TS1d** are the TS's structure with the OtBu in interaction with the double bond of the norbornadiene not involved in the first CC bond formation. **TS1c** leads to complex **2c** with the H in up position. With this definition, **1b** and **1c** will lead to the major product. The biggest energy difference between **1a/1b/1c/1d** is 0.09 kcal/mol so we will discuss the energy of **TS1** and the barrier given in the table below (Table 3.8).

Table 3.8Energy barriers for TS1

Pathway from	First barrier in kcal/mol	Second barrier in kcal/mol	Final complex
1a	6.16	22.99	Minor
1b	5.42	22.32	Major
1c	8.27	20.89	Major
1d	7.10	24.63	Minor

We can note that due to the interaction between the OtBu substituent and the double bond, the first bond formation is easier from complex **1a** and **1b**. The lowest barrier is obtained from the complex, which will lead to the major product but the energy difference is not so significant even if a rough calculation using Boltzman distribution leads to a 0.6 kcal/mol energy difference for a 70/30 ratio at a temperature of 60°C.

3.5.4 Comparison of the models

As said previously, the main drawback of this study is the use of modelisation of the bulky substituents by methyl for the substituents of the phosphine. That's why we switch to the real models. All the main characteristics of the pathway remain exactly the same, with energy in the same range as previously described with a difference of max ±2kcal/mol. The discussion about the selectivity is also quite similar and the analysis of the **TS1** gives the following results. **TS1b** is the lowest TS with a barrier of 4.68 kcal/mol, followed by **TS1d** 1.84 kcal/mol higher in energy, then **TS1a** 4.07 higher and finally **TS1c** 5.24 kcal/mol higher in energy. The order is not exactly the same but the trends are similar. The lowest **TS1** is still the one leading to the major product and the energy difference with the following **TS1** has increased. When we compare the first bond formation with the Ph or H substituent in an up position, we can see that the lowest barrier is observed when the double bond of the norbornadiene involved is in interaction with the OtBu substituent.

In order to see if the interactions between all the substituents play a role in the description of this reactivity, we computed the pathway using the B3LYP-D functional. The results using the ΔG values for the **1c** model pathway does not show any influence of both the change from E to G or for the change from B3LYP to B3LYP-D.

Similarly, single points at the B97D level on the optimized geometry at the B3LYP level for the real system does not lead to any change. The barrier for **TS1b** is still the lowest with a value of 4.17 kcal/mol. **TS1d, TS1a** and **TS1c** are respectively 0.95, 1.78 and 3.78 kcal/mol higher in energy.

Thus, the selectivity is already determined by the formation of the first C-C bond and indeed some difference of energy between TS1 have been found even if the energy difference reaches the accuracy of the method.

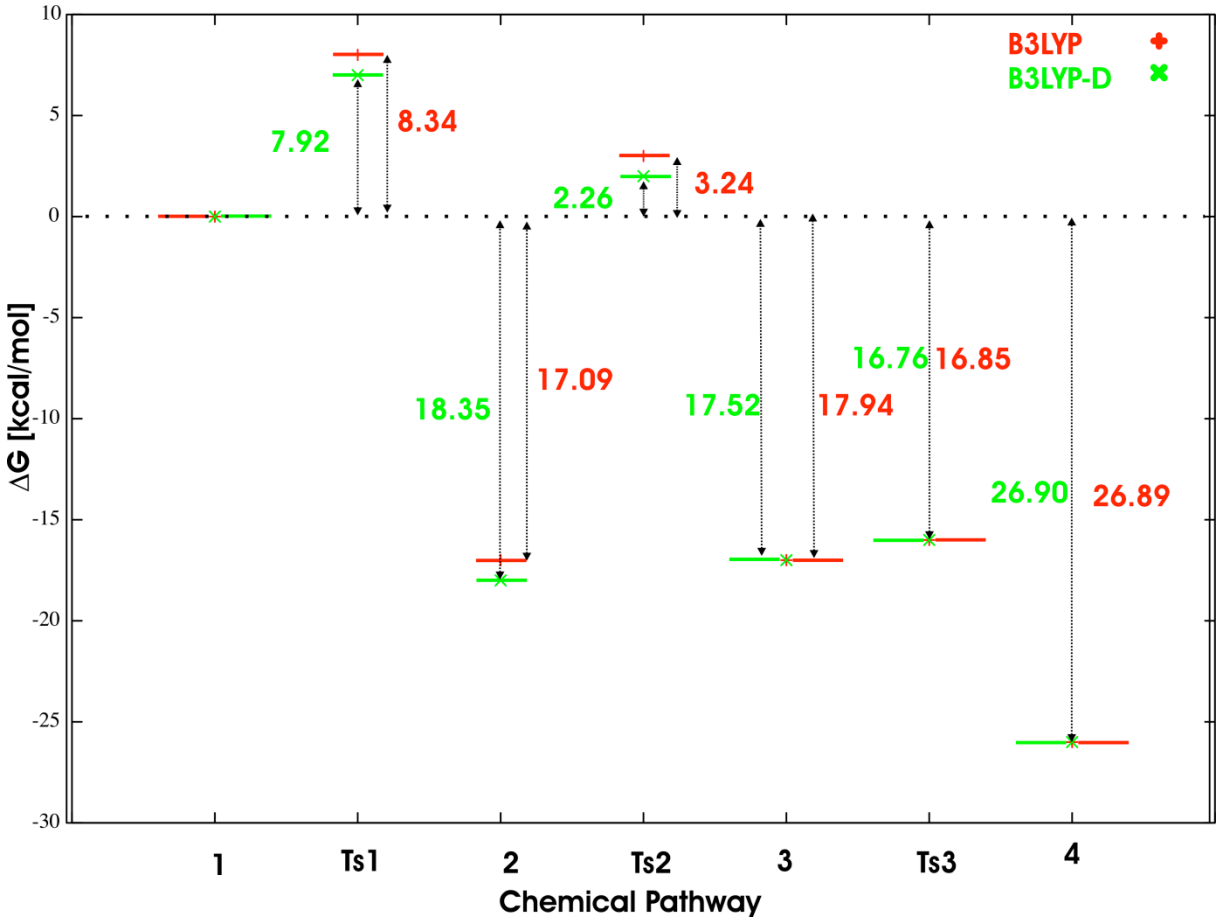


Figure 3.25 Chemical pathway for the model 1c computed at B3LYP and B3LYP-D level of theory

3.5.5 Conclusions

With this study, we have proposed a reasonable general mechanism for the formation of product **5p** using Platinum(II) complexes stabilized by secondary phosphine oxide ligands. One of the aspects of this reaction is its selectivity. We studied it and found that this selectivity is indeed already controlled when the first C-C bond is formed. The analysis of the barrier for this first step is indeed in good agreement with the experimental ratio but the energy diffence found is quite small and quite close to the accuracy of the method used even if the same trend has been found using different models and theoretical levels of calculations.

3.6 Conclusion remarks and perspectives

This work's aim was to test the DFT-D functionals to describe the encapsulation of small gas molecules with the cryptophane, which is driven by weak interactions. Moreover, dynamic studies, using metadynamics were conducted to determine the free energies of association during the encapsulation process. In comparison with other DFT methods, the DFT-D seems

more appropriate for calculating the energies of weak interactions for this kind of systems Host-Guest type. The best's results were obtained using wD97XD and B97D functionals.

On the other hand metadynamics simulations showed their success as a relatively efficient tool for constructing the surface free energy 'F' associated with the encapsulation process Ligand-Cryptophane molecules for the CH_4 as Ligand.

Relative bibliography to Chapter 3

(1) Stefan Grimme Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.***2006**, *27*, 1787-1799.

(2) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian 09, Revision A.02, Gaussian, Inc., Wallingford CT, 2009.

(3) Gabard, J.; Collet, A. Synthesis of a (D3)-bis(cyclotriveratrylenyl) macrocage by stereospecific replication of a (C3)-subunit. *J. Chem. Soc., Chem. Commun.***1981**, 1137-1139.

(4) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.***1993**, *98*, 5648.

(5) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B***1988**, *37*, 785.

(6) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.***2004**, *393*, 51-57.

(7) Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. A long-range-corrected time-dependent density functional theory. *J. Chem. Phys.***2004**, *120*, 8425.

(8) Vydrov, O. A.; Scuseria, G. E. Assessment of a long-range corrected hybrid functional. *J. Chem. Phys.***2006**, *125*, 234109.

(9) Vydrov, O. A.; Heyd, J.; Krukau, A. V.; Scuseria, G. E. Importance of short-range versus long-range Hartree-Fock exchange for the performance of hybrid density functionals. *J. Chem. Phys.***2006**, *125*, 074106.

(10) Vydrov, O. A.; Scuseria, G. E.; Perdew, J. P. Tests of functionals for systems with fractional electron number. *J. Chem. Phys.***2007**, *126*, 154109.

(11) Chai, J.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.***2008**, *10*, 6615.

(12) Brotin, T.; Dutasta, J. Cryptophanes and Their Complexes—Present and Future. *Chemical Reviews***2009**, *109*, 88-130.

(13) Fogarty, H. A.; Berthault, P.; Brotin, T.; Huber, G.; Desvaux, H.; Dutasta, J. A Cryptophane Core

Optimized for Xenon Encapsulation. *Journal of the American Chemical Society***2007**, *129*, 10332-10333.

(14) Chaffee, K. E.; Fogarty, H. A.; Brotin, T.; Goodson, B. M.; Dutasta, J. Encapsulation of Small Gas Molecules by Cryptophane-111 in Organic Solution. 1. Size- and Shape-Selective Complexation of Simple Hydrocarbons. *The Journal of Physical Chemistry A***2009**, *113*, 13675-13684.

(15) Diep, P.; Johnson, J. K. An accurate H2–H2 interaction potential from first principles. *J. Chem. Phys.***2000**, *112*, 4465.

(16) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.***1970**, *19*, 553.

(17) Canceill, J.; Lacombe, L.; Collet, A. Analytical optical resolution of bromochlorofluoromethane by enantioselective inclusion into a tailor-made cryptophane and determination of its maximum rotation. *Journal of the American Chemical Society***1985**, *107*, 6993-6996.

(18) Tambuté, A.; Canceill, J.; Collet, A. Optical resolution of C3 cyclotriveratrylenes and D3 crytophanes by liquid chromatography on chiral stationary phase Chiralpak-OT(+). *Bull. Chem. Soc. Jpn.***1989**, *62*, 1390-1392.

(19) Soulard, P.; Asselin, P.; Cuisset, A.; Aviles Moreno, J. R.; Huet, T. R.; Petitprez, D.; Demaison, J.; Freedman, T. B.; Cao, X.; Nafie, L. A.; Crassous, J. Chlorofluoroiodomethane as a potential candidate for parity violation measurements. *Phys. Chem. Chem. Phys.***2006**, *8*, 79.

(20) Canceill, J.; Collet, A.; Gabard, J.; Gottarelli, G.; Spada, G. P. Exciton approach to the optical activity of C3-cyclotriveratrylene derivatives. *Journal of the American Chemical Society***1985**, *107*, 1299-1308.

(21) Brotin, T.; Barbe, R.; Darzac, M.; Dutasta, J. Novel Synthetic Approach for Optical Resolution of Cryptophanol-A: A Direct Access to Chiral Cryptophanes and Their Chiroptical Properties. *Chem. Eur. J.***2003**, *9*, 5784-5792.

(22) Costante-Crassous, J.; Marrone, T. J.; Briggs, J. M.; McCammon, J. A.; Collet, A. Absolute Configuration of Bromochlorofluoromethane from Molecular Dynamics Simulation of Its Enantioselective Complexation by Cryptophane-C. *J. Am. Chem. Soc.***1997**, *119*, 3818-3823.

(23) Zhou, H.; Gilson, M. K. Theory of Free Energy and Entropy in Noncovalent Binding. *Chemical Reviews***2009**, *109*, 4092-4107.

(24) Bigeault, J.; Giordano, L.; Buono, G. [2+1] Cycloadditions of Terminal Alkynes to Norbornene Derivatives Catalyzed by Palladium Complexes with Phosphinous Acid Ligands. *Angewandte Chemie International Edition***2005**, *44*, 4753-4757.

(25) Bigeault, J.; Giordano, L.; de Riggi, I.; Gimbert, Y.; Buono, G. Platinum(II)-Coordinated Phosphinous Acid-Catalyzed Alkylidenecyclopropanation of Bicyclic Alkenes with Terminal Alkynes. *Org. Lett.***2007**, *9*, 3567-3570.

(26) Bigeault, J.; de Riggi, I.; Gimbert, Y.; Giordano, L.; Buono, G. Tandem [2+1] Cycloaddition-Ring Expansion of Bicyclic Alkenes with Tertiary Propargylic Acetates Catalyzed by Palladium(II)-

Coordinated Phosphinous Acid. *Synlett***2008**, *2008*, 1071,1075.

(27) Thota, R.; Lesage, D.; Gimbert, Y.; Giordano, L.; Humbel, S.; Milet, A.; Buono, G.; Tabet, J. Gas-Phase Study of Phenylacetylene and Norbornadiene on a Palladium(II) Phosphinous Acid Complex: Importance of the Order of Introduction of the Organic Partners. *Organometallics***2009**, *28*, 2735-2743.

(28) Tenaglia, A.; Giordano, L.; Buono, G. Herrmann−Beller Phosphapalladacycle-Catalyzed Addition of Alkynes to Norbornadienes. *Org. Lett.***2006**, *8*, 4315-4318.

(29) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford CT, 2004.

(30) Hay, P. J.; Wadt, W. R. Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. *J. Chem. Phys.***1985**, *82*, 270.

(31) Ortiz, J. V.; Hay, P. J.; Martin, R. L. Role of d and f orbitals in the geometries of low-valent actinide compounds. Ab initio studies of U(CH3)3, Np(CH3)3, and Pu(CH3)3. *J Am Chem Soc***1992**, *114*, 2736-2737.

(32) T. H. Dunning, Jr.; P. J. Hay H. F. Schaefer III: New York, 1976; Vol. 3, pp. 1-28.

Chapter 4 FROM FREE ENERGY TO ENTROPY *VIA* METADYNAMICS

Résume

Dans le quatrième chapitre, nous avons déterminé les contributions thermodynamiques pour la réaction de Claisen et les réactions de Diels-Alder intermoléculaires en utilisant les développements réalisés dans notre équipe dans le cadre de la métadynamique. Ainsi, nous avons développe une application python «entropy» qui calcul cette contributions thermodynamique de l'énergie libre en fonction de l'énergie potentielle et une coordonne de la réaction. Dans les paragraphes suivants, nous présenterons la stratégie de mise en œuvre de cette nouvelle variable collective dans la nouvelle version de CP2K et le "bias_exe.py", l'application que nous avons développé pour effectuer « Bias-Exchange »

4.1 Abstract

In the fourth chapter we compute the thermodynamic contributions for Claisen reaction and Diels-Alders intermolecular reactions using the developments made in our laboratory in the framework of metadynamics. Thus, we develop a python application "EntroPy" to compute these thermodynamic contributions from the free energy surface as a function of the potential energy and one ordinary geometrical reaction coordinate. In the following paragraphs, we will present the strategy of implementation of this new collective variable in the new version of the CP2K and the "bias_exe.py", application that we develop to perform bias exchange metadynamics.

4.2 Introduction

In the last couple of years, progress was made to achieve more efficient metadynamics. This can be exemplified by the implementation of new collective variables and extensions of the metadynamics algorithms[1-5]like well-tempered metadynamicsandbias-exchange . In this chapter, I will focus on the developments that had been made in our group in the last several years in the field of metdaynamics and, especially on my work during my PhD.

In 2009 Carine Michel, Alessandro Laio and Anne Milet published a paper in the Journal of Chemical Theory and Computation, where they have used potential energy as a collective variable in the metadynamics framework at a finite temperature[5]. They were able to reconstruct the free energy surface as a function of the potential energy and a geometrical variable **s,** $F_{T_0}(E,s)$**.** They were also able toestimate the free energy surface at different temperatures. This successful attempt allows tracing the entropy and characterization ofthe thermodynamic quantities of a molecular system at all the temperatures by a single simulation.

4.3 Energy as collective variable; technical aspects

The CP2K code is written using object orientated programming approach using the programming language FORTRAN 95. In the CP2K code, the metadynamics is handled by the ***metadynamics*** modulus.

In order to trace the entropy of a system along the reactivity pathway, we implemented a new collective variable, Energy (the potential energy of the system), in the new code Quickstep that is part of the CP2K package.

The subroutines needed by algorithm of the metadynamics which has been described in the chapter 1 of this manuscript, are located in the ***metadynamics*** modulus and these are the following:

- ***metadyn_integrator***, which is a general driver for applying metadynamics.
- ***metadyn_forces*** that add forces to the subsystem due to the metadynamics run.
- ***hills*** subroutine that is the major driver for adding hills, computing forces due to the history dependent term and write down the HILLS file. Described by the equation 1.25 from chapter 1
- ***metadyn_write_colvar*** subroutine which is responsible for writing the information's from a metadynamics run in COLVAR file at each step of dynamics.

This metadynamics modulus must be called by the MD main loop once the forces have been computed and before the positions/velocities are updated for the next step. One important aspect of ***metadynamics*** modulus is to compute the derivative of the history dependent potential $V_G(S(r),t)$ with respect to **r** using the chain rule: $\left(\frac{\partial V_G(S(r),t)}{\partial r}\right)=\left(\frac{\partial V_G(s,t)}{\partial s}\right)\left(\frac{\partial S(r)}{\partial r}\right)$, where **s** represents the value of the collective variables at that moment. It should be noted that the only restriction for a collective variable is to be a function of r and this condition is directly linked to this chain rule. If we imagine that we have the collective variable Energy then the derivatives of the collective variable, energy, with respect of the atomic position is the gradient of the energy $\frac{\partial E}{\partial r_i}=-f_i$. As we can see, this gradient is the opposite of the forces.

This calculation is helpful in our case because all these quantities are previouslycalculated in the code. So, for the programmer, it is an easy task to gather all these quantities together. Thus, in the subroutine ***energy_colvar*** all these tasks are performed and are kept in the ***colvar_methods*** modulus.

4.4 From Metadynamics to entropy

After a metadynamics simulation using at least 2 collective variables where one is Energy and the others are geometrical CV's, we should be able at this point to reconstruct the free energy surface (FES) as a function of the collective variables namely F_T(E,s) at temperature

T. From here, we can access the free energy, internal energy and entropy of the system as a function of these geometrical CV's.

All these features are coded in an application ***entropy,*** inwhich starting from the free energy surface F_T(E,s), we are able to obtain F_T(s), U_T(s) and $T \times S_T(s)$ where T is temperature, S the entropy, U the internal energy and F the free energy profiles. The aim of this code is to create a flexible, robust and easy to use (user-friendly) application, which has the scope to study Entropy along a reactivity pathway starting from a Metadynamics simulation. One of the goals is also to be able to extract one set of thermodynamic quantities from several similar runs.

The internal energy profile U_T(s) and the free energy profile F_T(s) are derived directly from F_T(E,s):

$$U_T(s) = \frac{\int dE\, E \exp\left(-F_T(E,s)/k_B T\right)}{\int dE \exp\left(-F_T(E,s)/k_B T\right)} \quad (5.1)$$

$$F_T(s) = -k_\beta T \cdot \log\left(\int dE \exp\left(\frac{-F_T(E,s)}{k_\beta T}\right)\right) \quad (5.2)$$

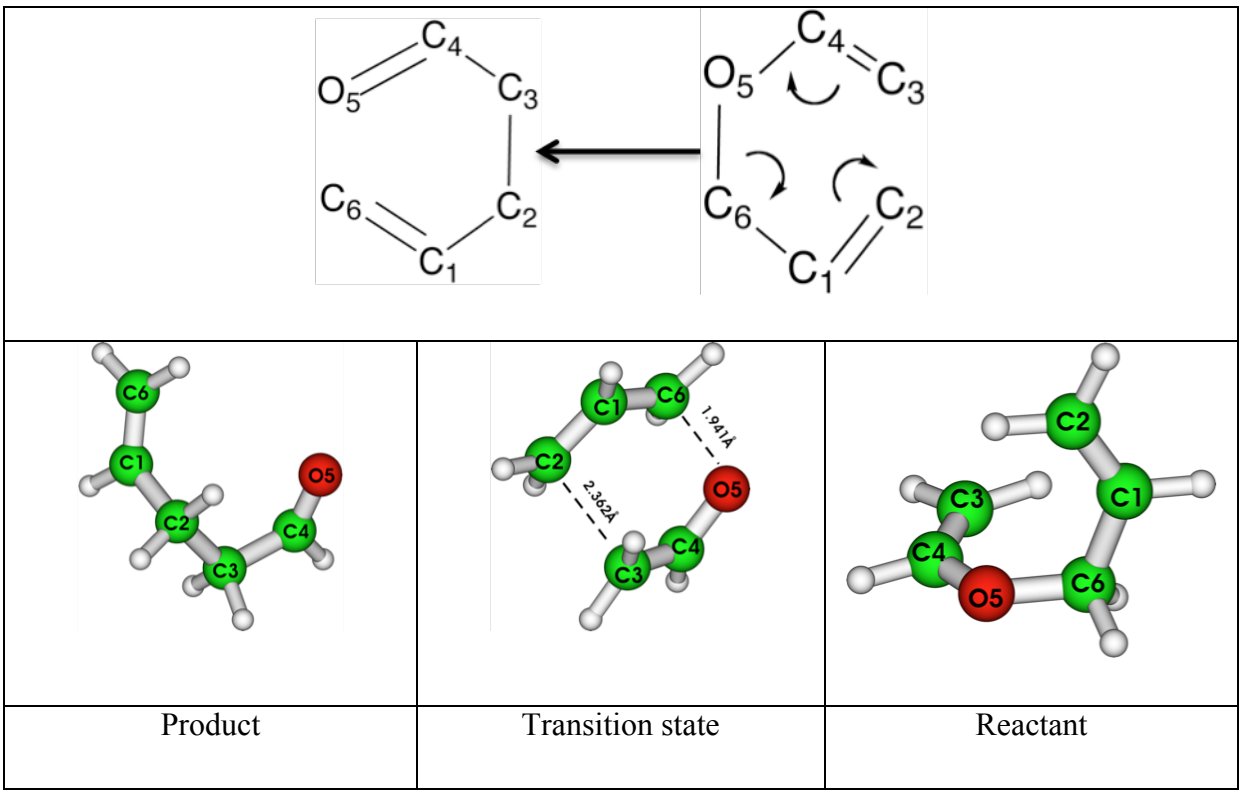
and the entropy S_T(s) as a function of the reaction coordinate "s" is given by the equation:

$$S_T(s) = \frac{U_T(s) - F_T(s)}{T} \quad (5.3)$$

The Entropy application is coded to be able to reconstruct the thermodynamic parameters from a single metadynamics run or from a set of independent but equivalent metadynamics simulations for a reaction. These simulations are independent but equivalent, because for each simulation, the initial velocities have been randomly and differently chosen according to Boltzmann distribution at the simulation temperature T. The interest to use more than one simulation is due to the statistical way used to compute U_T(s). U_T(s) is indeed directly computed as a mean (see equation above) and so the more data about F_T(E,s) we can have, the better. If we explore a finite region of the CV's space, the fact it has been more or less explored is directly linked to the number of Gaussians added. So to ensure that a region has been sufficiently explored, the history dependent potential must be imposed to be greater that

a multiple of K_bT. This condition assures that theregion is sufficiently explored by the Metadynamics simulation and that the reconstruction of $U_I(s)$ in this region is correct.

This procedure has been validated on a simple reaction, the Claisen rearrangement, treated at the PM6 level of theory. This reaction was chosen because both reactants and products are both linear chains with two double bonds and in between sp3 carbons, which allow flexibility. On the other hand, the TS structure is a strained cyclic like structure without flexibility. So because of this lack of conformational entropy in the TS region, a significant drop of the entropy around the TS structure is expected.



Product	Transition state	Reactant

Figure 4.1 Claisen rearrangement reaction with atoms number and geometrically CV's definition for the metadynamics simulation.

The construction of the $F_I(E,s)$ is done at T=250 K. The CV used are the potential energy **E** and a second collective variable **s** (Figure 4.2). This variable represents the difference between the two distances, and these distances were chosen to discriminate between the formation or breaking of the bonds in the Claisen rearrangement. So CV2 is defined as *s=[dist(C2-C3)-dist(O5-C6)]* according to the atom numbering that is presented in the Figure 4.1.

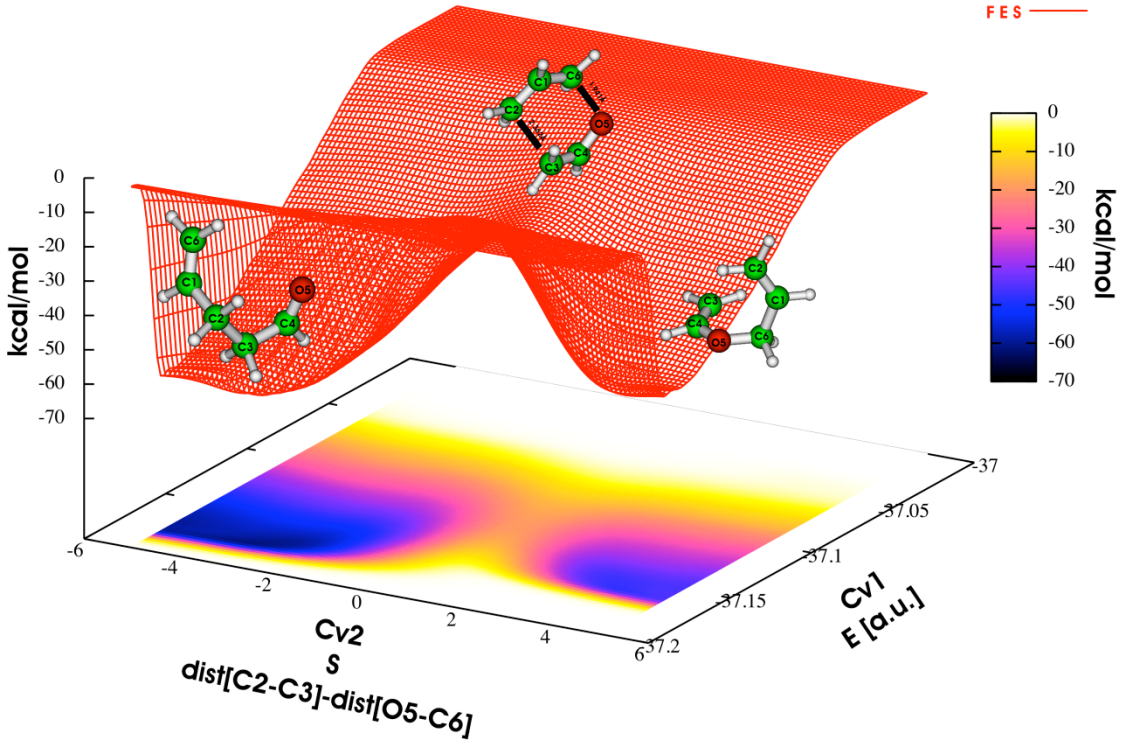


Figure 4.2Free energy reconstruction functions to the CV1 and CV2.

When the two distances are equal then the difference is equal to zero, which is the case around the transition state structure, whereas when CV2 is positive, the system is in the reactant region and when CV2 is negative, the system is in the product region. The shape of the added Gaussians is defined by the height (0.5 kcal/mol), the width along **s** (1 bohr), and the width along **E** (2.5 kcal/mol). The Gaussians are added every 5 fs. The total time of the simulation is 30 ps and after this simulation time, a diffusive behavior is observed in the CV's space. The free energy surface $F_T(E,s)$ is then reconstructed on a 100 X 100 grid in the (-5 <**s** < 5) X (-37.200 < E <-37.000) region. This region corresponds to the transition state and thus is selected to present the comportment of these three thermodynamically quantities (**F, U** and **S**) in this region. Starting from this reconstructed surface $F_T(E,s)$ generated with the ***Graph*** application of the CP2K source code and using our application, in which are implemented all the equations from 1 to 3, we will find the values F(s), U(s) and S(s). The procedure is emphasized in the Figure 4.3.

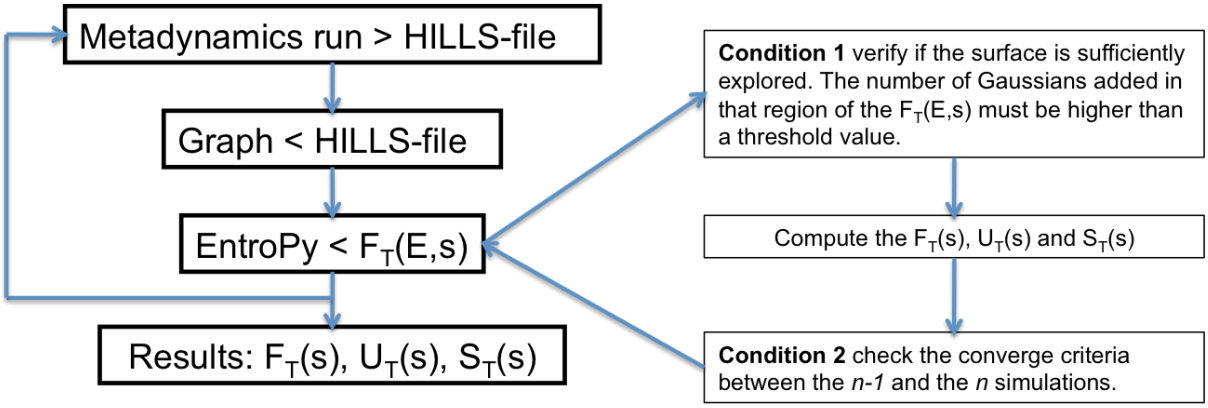


Figure 4.3Flow diagram of the procedure to have access to the thermodynamically parameters F**,** U **and** S

When $F_T(s)$ is reconstructed, a first test is performed: the test of the sampling of the surface through the density of deposited Gaussians. The link between these two issues has been already addressed. When we use several metadynamics runs, the average $F_T(E,s)$ is defined as follows; for example for five independent but equivalent simulations, the $F(E,s) = \frac{\sum_i F_T^i(E,s)}{5}$, this average value assuresa more correct sampling of the region explored according to statistical laws. So if the number of deposited Gaussians is below a given range of values of s, the surface is said not to have been sufficiently explored and is not reconstructed for the values of s under consideration.

In the code, another test is added: a convergence criterion is set for $F_T(s)$ or $U_T(s)$. The convergence criterion is defined as the difference between the average values of *n-1* metadynamics simulations and the average values of *n* metadynamic simulations. If this difference is smaller than a threshold value then the convergence criteria are accepted and the cycle described in Figure 4.3can stop. In Figure 4.4 from panel A to D is presented thevariation ofthe reconstruction of the $F_T(s)$ with the number of the simulation.

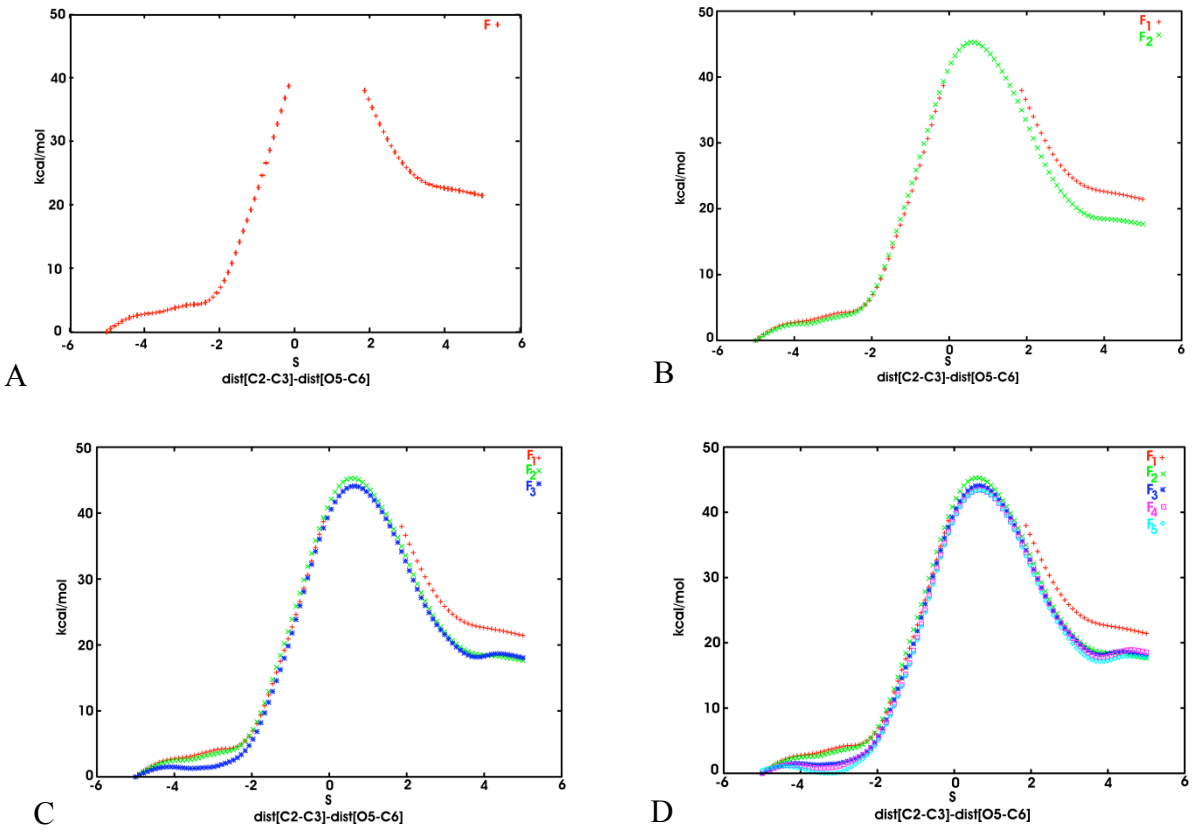


Figure 4.4Panel "A" represents the first metadynamic simulation of the Clasien reaction discussed earlier in the text. Panel "B" the first and the mean of the two equivalent simulations. InPanel "C" the blue points represent the mean of three metadynamic simulations. Panel "D" is the mean of fourth and fifth simulation reconstructions of the F(s), which are points colored in magenta and in cyan respectively.

The missing points are clearly in the TS region [-1,1.5] and indicate that the reconstruction of the F(s) is not done because this region is insufficiently explored with one metadynamics simulation. It is not surprising. Intuitively, the TS region is less explored than the reactant and product region. The program controls those points through the control of the number of deposited Gaussians from the metadynamics run used for the reconstruction of the F(s). If the number of deposited Gaussians is below a given threshold, the surface is not reconstructed in this region. The threshold to test the quality of the exploration of the region is an option to be defined by the user. In our case it is set to 10 KbT ~5 kcal/mol for a temperature of 250K, accordingly with the ref [5]. After five independent but equivalent simulations the differences between the mean of four simulations are smaller that 4 KbT meaning that the F(s) reconstruction has converged with the convergence criteria set to 4 KbT. The simulation temperature is 250K meaning 4 KbT ~ 2 kcal/mol.

To test our approach we have performed two sets of 12 independent simulations each, applied to the same Claisen reaction. The aim of this test is to see how the convergence criteria for $F_T(E,s)$ influence $U_T(s)$ therefore $S_T(s)$, which presents more oscillations thanthe $F_T(s)$ for only five simulations. The second reason was to check if the two independent sets of simulations will indeed converge to the same results using different sets of data.

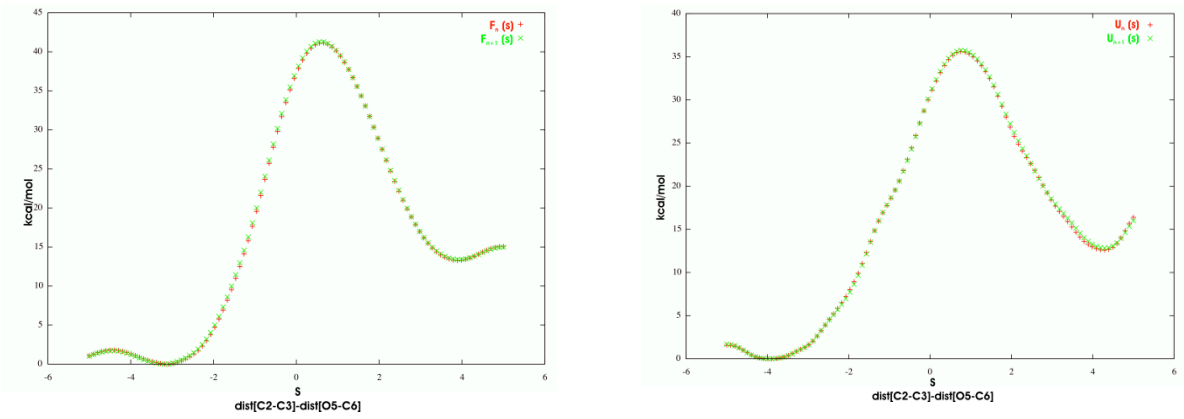


Figure 4.5Averages of the $F_T(s)$ (left) and $U_T(s)$ (right) curve taken from *n* and *n+1* numbers (read and green curve respectively) of independent metadynamics simulations.

Byincreasing the number of equivalents simulations, we obtain a difference of 0.5 KbT between the $F_T(s)$ average curve plotted at *n* and*n*+1 with *n*=11 representing the number of independent and equivalents simulations. At this point the $U_T(s)$ curve difference between *n* and *n*+1 steps of simulations is approximately 1 KbT. The $U_T(s)$ and $F_T(s)$ are in the same time converged, and this fact assures us that imposingthe convergence criteria to the F or to the U term is a good option. This task remains at the user choice, becauseitdependson the system under study and of the accuracy expectedfor this specific study.

Concerning the comparison of the results obtained from the two sets of twelve independent runs, we can see that the shape of the two curves F and U is clearly similar and they all give the same thermodynamic results. The relative energy's is calculated setting the "zero" to the minima of the product region. Minor differences appear especially in the region of the reactant $s > 1.5$ and product $s < -1.5$ region and may be due to the use of the wall to prevent the system from investigating region with high value or low value of s. This way, we can obtain more easily the diffuse regime of the simulation needed to extrapolate all the

thermodynamic data. This problem may also be amplified by the existence of many minima for the border values of s.

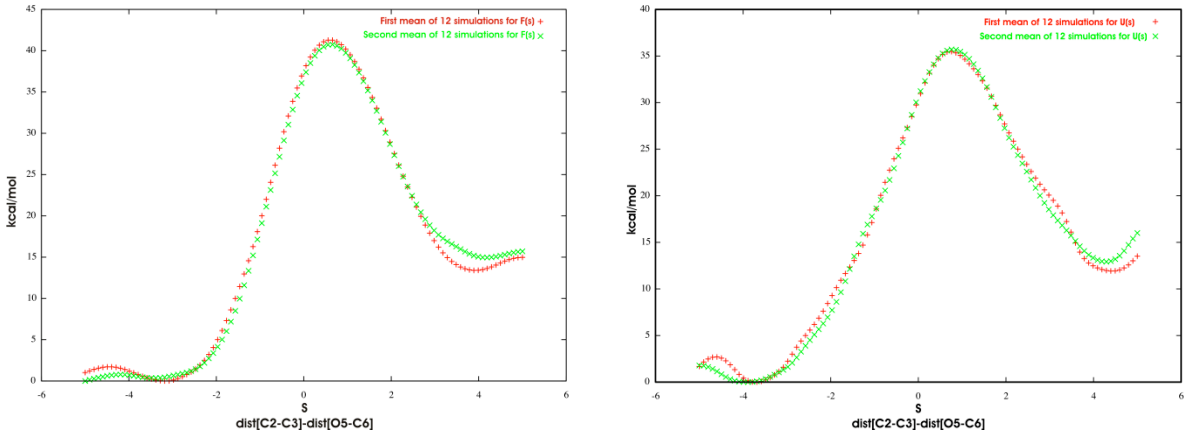


Figure 4.6Figures of the two independent sets of 12 simulations each, for the F_T(s) (left) and U_T(s) (right) curve.

The most important shifts between these two sets of simulations (12 for each set) can be seen in Figure 4.6. These differences between these two sets of independents simulations are predominately in the product and reactant regions where we have a maximum relative deviations of ~1.5kcal/mol for the U_T(s) curve.

Now, we will discuss the profiles of one of these sets of 12 independent but equivalent metadynamics runs.

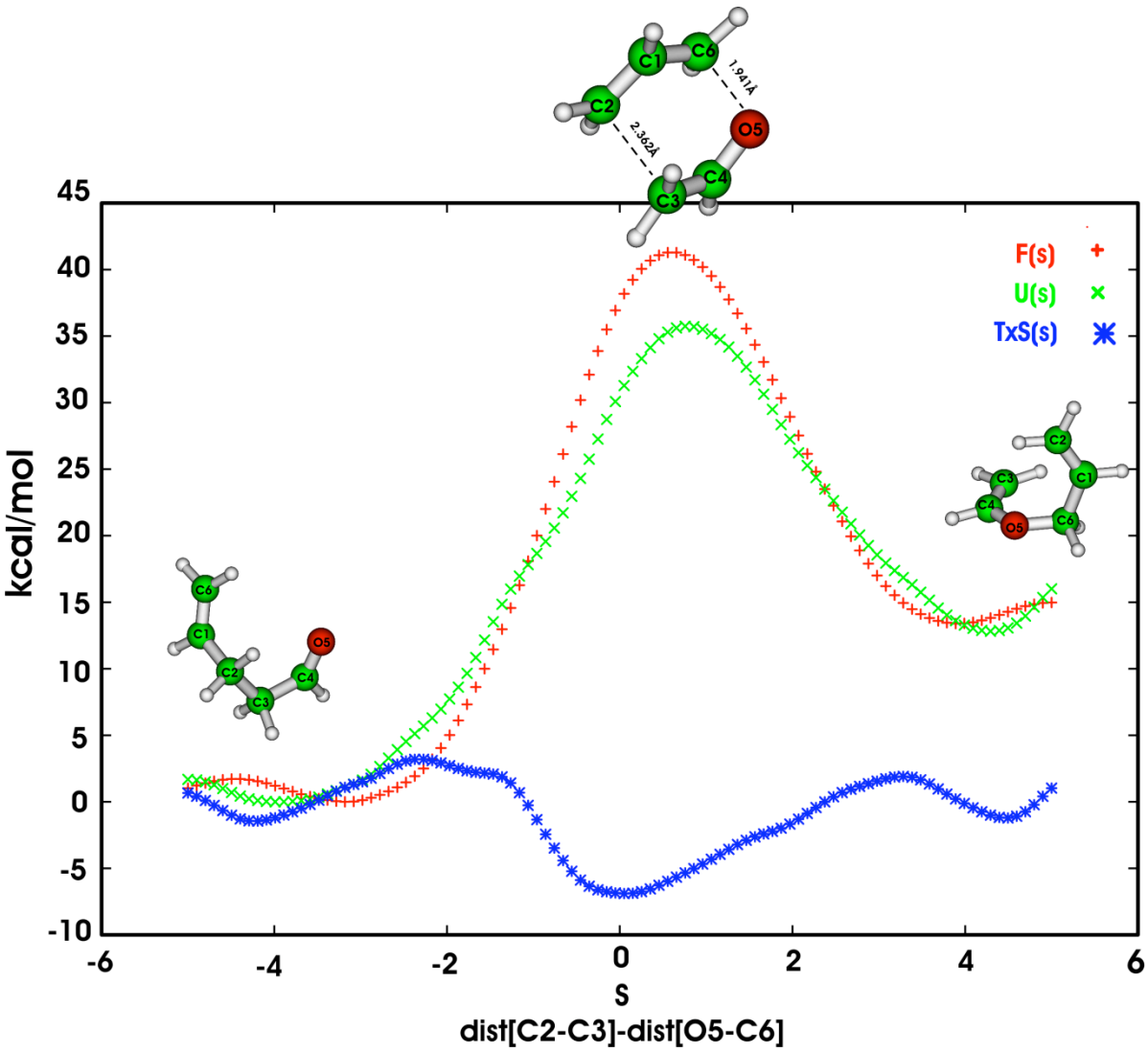


Figure 4.7Thermodynamically parameters $F_T(s)$, $U_T(s)$ and $TxS_T(s)$ for the Claisen rearrangements at 250K. The curves plotted are extracted from a mean of 12 independent metadynamics simulations.

The free energy profile $F_T(s)$ and the entropy profile $S_T(s)$ obtained at 250K are plotted together with the internal energy profile $U_T(s)$ in Figure 4.7. Both F and U present two minima at s=4 and s=-3 corresponding to the reactant and to the product conformations. From the reactant well to the transition state the free energy increases by 28.5 kcal/mol. The internal energy exhibits the same shape but with a lower barrier of 23.0 kcal/mol which corresponds toa change in the entropy profile of TxS= 5.5 kcal/mol. The free energy and the internal energy profile presents a maximum value around s=0.61 and s=0.76 respectively.

The entropic profile $S_T(s)$, is what we expected: from the region of the reactant, $s > 0$, $S_T(s)$ is monotonically decreasing and from the TS region, it is increasing to regain approximately the same value than in the reactant region. We can note that the slope of the curve is quite

different for s < 0 and for s >0 and also that the maximum of $S_T(s)$ does not correspond to the maximum of $F_T(s)$ hence the difference between the maximum of F and U.

Table 4.1Thermodynamic parameters from metadynamics simulation and from a static reaction simulation at the PM6 level of theory, taken with CP2K and Gaussian09 respectively. U **is internal energy eq(5.1).** F **is free energy eq(5.2) and** S **entropy eq(5.3).** H **and** G **are enthalpies and Gibss free energy respectively.**

Relative calculations of Thermodynamic parameters	Metadynamics simulations [kcal/mol]	Static simulations [kcal/mol]
Internal energy of reaction	$\Delta_r U(250K)$=-12.5~-13.0	$\Delta_r H(250K)$=-11.7
Free energy of reaction	$\Delta_r F(250K)$=-13.0	$\Delta_r G(250K)$=-12.7
Activation free energy	$\Delta F^{\ddagger}_{Fmax}(250K)$=-28.5 $\Delta F^{\ddagger}_{Umax}(250K)$=-28.0	$\Delta^{\ddagger} G(250K)$=-31.9
Activation internal energy	$\Delta U^{\ddagger}_{Umax}(250K)$=-23.0 $\Delta U^{\ddagger}_{Fmax}(250K)$=-22.5	$\Delta^{\ddagger} H(250K)$=-29.9
Temperature multiply with activation Entropy	$T x \Delta S^{\ddagger}_{max}(250K)$=6.5 $T x \Delta S^{\ddagger}_{Fmax}(250K)$=5.5 $T x \Delta S^{\ddagger}_{Umax}(250K)$=5.0	$T x \Delta^{\ddagger} S(250K)$=2.0

The thermodynamic quantities obtained from the cp2k simulation are in very good agreement with the static calculation at the same level of theory, which is expected for an intramolecular reaction. As the maxima of F and U and so of S do not correspond, we have computed the thermodynamic quantities for the different maximum. For the comparison with the static calculations, we will have to compare the $\Delta G^{\ddagger}$ static value with the $\Delta F^{\ddagger}$computed at the maximum of U, the same conclusion is also valid for $T x \Delta S^{\ddagger}$. We can note that in this case we are not able to recover all the change in entropy using static calculations since the maximum entropy deviation does not correspond with the maximum of U.

4.5 Intra-molecular Diels-Alder reaction

The intra-molecular Diels-Alder (IDA) reaction is not only a well-known[6-8] reaction for organic synthesis, but ithas become a testing ground for computational methods and prediction of stereo-selectivity due to the large number of experimental data available. And in this way we study the influence of the size of the cycloalkane unit on the thermodynamic parameter of the IDA reaction.

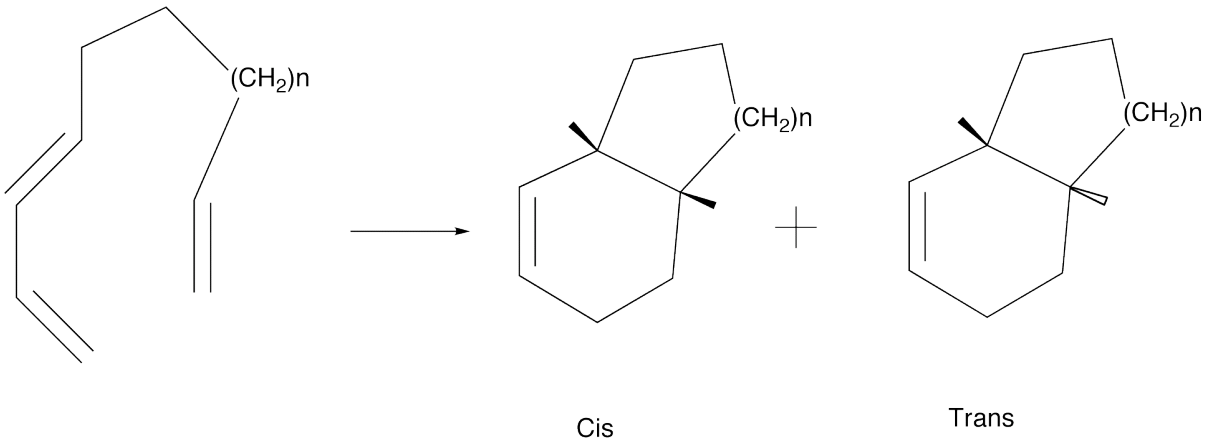
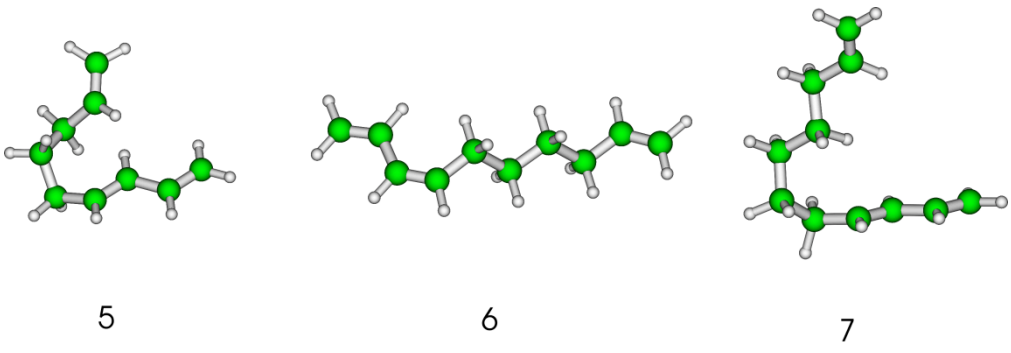


Figure 4.8Intra-molecular Diels-Alder reaction of 1,3,n-Triene

The diagram presented inFigure 4.8, represents the intra-molecular Diels Alder reaction and how the two products *cis* and *trans* can be formed. The only variable is ***n*** the number of carbon atoms that will determine the size of the cycloalkane.Referring to the size of ring formed, we use the notation (***5***) for 1,3,8 nonatriene, (***6***) for 1,3,9 decatrine and(*7*)for 1,3,10 undecatriene. For products and for transition states structures we can observe a *cis* or *trans* regio- selectivity.



A Reactants of Diels-Alder intramolecular reaction

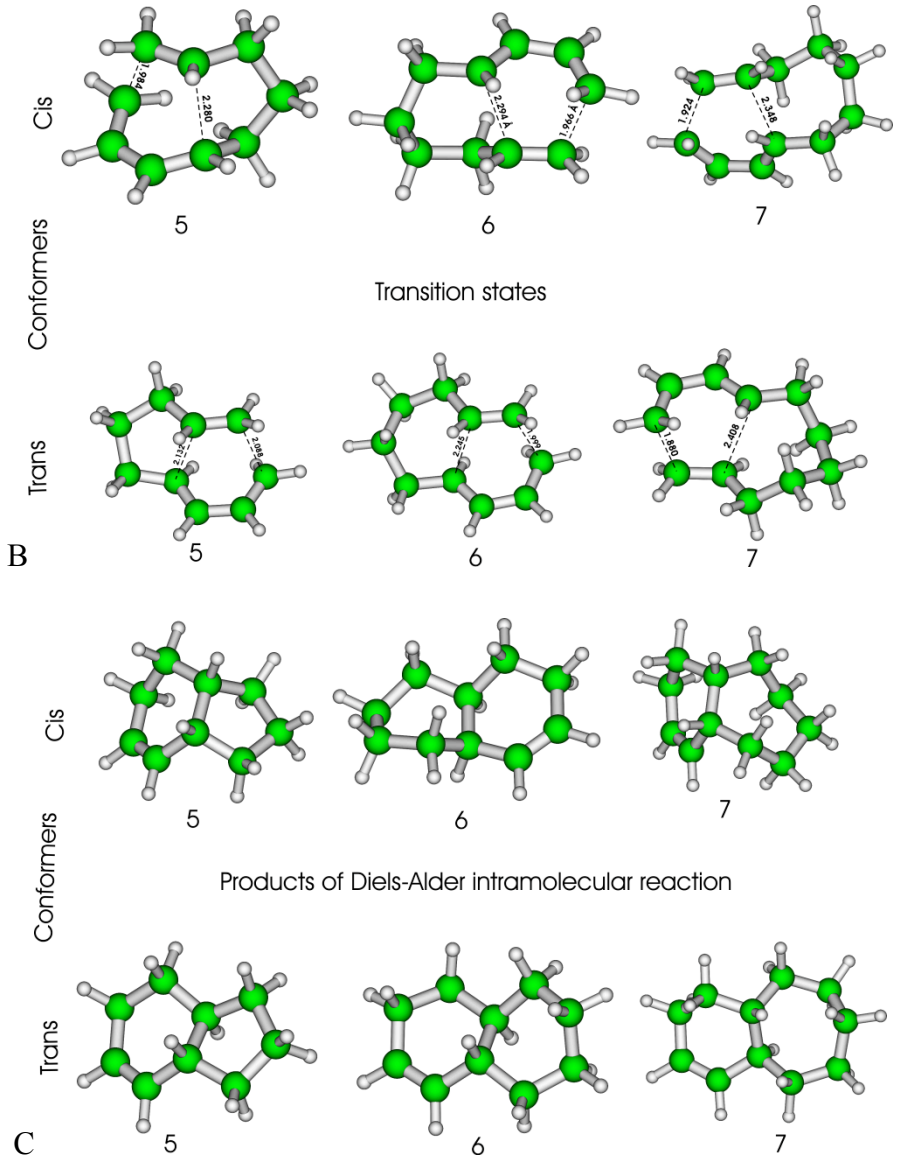


Figure 4.9Optimized geometries of reactants (panel A), transition states (panel B) and products (panel C) of the intra-molecular Diels Alder reaction.

The reaction was studied by static methods, optimizations of the reactants, transition states and products, and by metadynamics. The level of the theory used for the calculations is PM6 for both the static methods and for metadynamics. The static calculations are performed with the Gaussian09[9] suit of programs and dynamics simulations are undertaken with CP2K[10] suit of programs. The first objective of this study is to find the thermodynamic quantities: the free energy of reactions F_T(s), internal energy U_T(s) and entropy TxS_T(s) for these reactions. The second purpose is to reproduce the stereo-selectivity of this kind of reactions with the help of molecular dynamics simulations. We use the same procedure as in the case of the Claisen

rearrangements discussed earlier in this chapter for the metadynamics simulations for the computations of these thermodynamical quantities.

In the metadynamics simulations we start from the product well in order to find the barriers for the *cis* and *trans* conformations and to compare them. We start from the product well to be able to discriminate at the beginning of the simulation between the *cis* and *trans* to study the stereo-selectivity of these reactions. As we start from the product well, we must be able to define a geometrical collective variable, which will represent the break of the bonds between the carbons involved in this IDA reactions. Experimentally this reaction is stereo-selective in favor of *cis* conformer for complex **5** (1,3,8 nonatriene) only; cis/trans ratio is find to be 73:27 in ref [[11]] and 69:31 in ref[[12]]. For the **6** (1,3,9 decatriene) and **7** (1,3,10 undecatriene) complexes these reactions are found to be nonselective.

4.5.1 Metadynamics parameters

The metadynamic simulations were done in the two dimensional subspaces of the real space. The geometrical collective variable "***s***" was chosen as a sum function of two distances D1 and D2 s= D1+D2: where the D1 and D2 represents the distances between the carbon atoms as can be seen in the Figure 4.9panel B and the second collective variable was set to be the potential energy of the system "***E***". The product well is around the region situated from 5 to 10 Bohr of the CV1 and the transition state and the reactant is situated in the region of CV1 from 10 Bohr to 30 Bohr respectively.

The height of the Gaussians added during the metadynamics was set to 0.5 kcal/mol the width of the Gaussians was chosen $\delta_1 = 1$ Bohr and $\delta_2 = 2.5$ kcal/mol for the CV1 and CV2 respectively. The integration time step was chosen to be 0.5 fs and the period of the added Gaussians was chosen to be $\tau_G = 5$ fs, which correspondsto a ratio of 10 time step of dynamic. The simulations are done in the canonical ensemble at 250 K. To assure the conservation of the temperature during the simulations, the predictor corrector was the velocity rescaling. After 50 ps of simulations when the CV1 became diffusive we stopped the metadynamics and reconstructed the free energy surface $F_T(E,s)$ on a 100x100 grid points in the $(0.0 < s < 45.0)$ x $(E_{min} < E < E_{max})$ region. For each of these three simulations of metadynamics these two limits (Emin, Emax) are found by visual inspection of the potential energy collective variable CV2.

4.5.2 1,3,8 – Nonatriene (5)

After five independent metadynamics simulations the convergence criteria set to 5 KbT~ 2.5 kcal/mol is obtained for the $F_T(s)$ curve and so we can determinethe different thermodynamic parameters for the *cis* and *trans* conformers. Using metadynamics simulations, we find out that the *cis* productismore stable than the *trans*product with a difference of 9.5 kcal/mol. The results obtained with static calculationsat the same level PM6 and at the same temperature 250K for the thermochemistry analysis give a relative difference between the *cis* and *trans* productsofonly 2.6 kcal/mol in favor of the *cis* conformer (*cis* conformer most stable than the *trans* conformer). At the transition state we find,by metadynamics simulations, a difference between the *cis* and *trans*transition states of only 2.0 kcal/mol in favor of the *trans* (*cis* being higher that *trans* by 2.0 kcal/mol) conformer.The static determination of the transition states structures give a difference of 2.80 kcal/mol in favor of the same *trans* conformer.

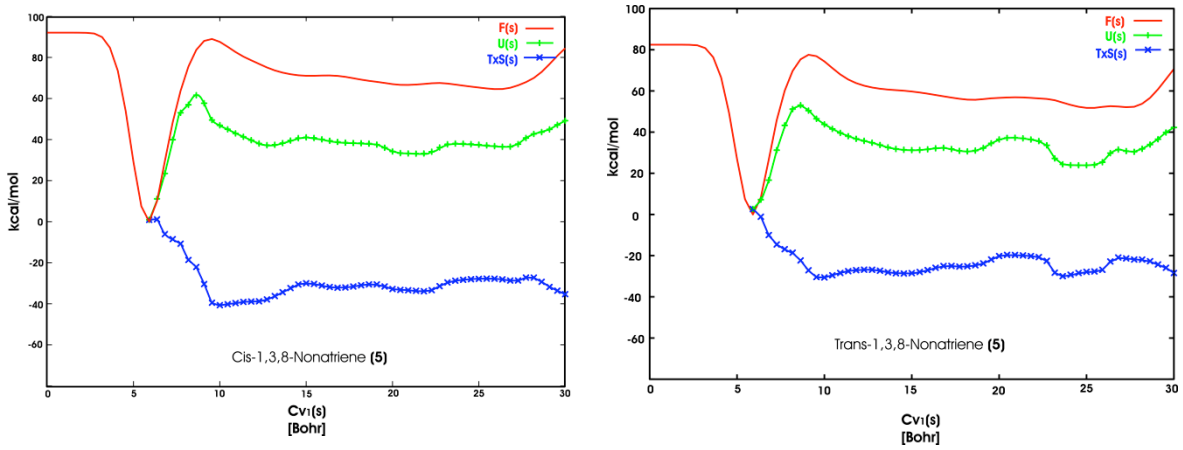


Figure 4.10Thermodynamical profiles reconstructions of the 1,3,8-Nonatriene for *cis* (left) and *trans* (right) conformers

The entropy curve is extracted from the five independent but equivalent metadynamics simulations as well as the free energy curve and internal energy curve. These curves are plotted in the Figure 4.10. These curves reveal the importance of the entropic effects in our simulated reaction. From the reactant well region (10 bohr < s < 30 bohr) the entropy decrease from -29.5 kcal/mol (s ~ 30 bohr) to -40 kcal/mol (s~10 bohr TS region) for the *cis* conformation and it decreases from -20 kcal/mol to -30 kcal/mol for the *trans* conformation. Then in both cases, the entropy increases to obtain its maximum around the product region. It is quite logical that the entropy increases from the TS region to the reactant since the TS is a

more strained structure compared to the linear flexible structure of the reactant. Nevertheless the increase in entropy from the strained bicycle product to the possibly less strained TS structure is less expected. But it has been computed this way from both static and metadynamics calculations and so indeed the TS structure seems to have less conformational entropy than the product.

The $T \times \Delta S^{\ddagger}$ value determined by metadynamics is -10.5 kcal/mol and -10 kcal/mol for the *cis* and *trans* respectively. By static calculations at the same level of theory as metadynamics simulations (PM6), these values are -5.4 kcal/mol and -5.3 kcal/mol for *cis* and *trans* respectively. At the B3LYP/6-31+G** level of theory we obtain by static calculations for $T \times \Delta S^{\ddagger}$ –4.0 kcal/mol and -3.8 kcal/mol for the *cis* and *trans* respectively. We also used B3LYP level of theory for comparisons purpose only in this case. It is well known that at this level of theory (B3LYP) the thermodynamically parameters are in good agreement with the experiments.[13-15]

Table 4.2Relative activation parameters for the Intramolecular Diels-Alder Reactions of 1,3,8-nonatriene (5) calculated by metadynamics and static simulations.

Conformers	Metadynamics kcal/mol	PM6 kcal/mol	B3LYP/6-31+G** kcal/mol
cis	$\Delta F^{\ddagger}$=30.0 $\Delta U^{\ddagger}$=19.5 Tx$\Delta S^{\ddagger}$=-10.5	$\Delta G^{\ddagger}$=40.5 $\Delta H^{\ddagger}$=35.1 Tx$\Delta S^{\ddagger}$=-5.4	$\Delta G^{\ddagger}$=36.9 $\Delta H^{\ddagger}$=33.0 Tx$\Delta S^{\ddagger}$=-3.9
trans	$\Delta F^{\ddagger}$=25.5 $\Delta U^{\ddagger}$=15.5 Tx$\Delta S^{\ddagger}$=-10.0	$\Delta G^{\ddagger}$=37.7 $\Delta H^{\ddagger}$=32.5 Tx$\Delta S^{\ddagger}$=-5.3	$\Delta G^{\ddagger}$=35.5 $\Delta H^{\ddagger}$=31.7 Tx$\Delta S^{\ddagger}$=-3.8
cis –trans at 250K	$\Delta\Delta F^{\ddagger}$=4.5 $\Delta\Delta U^{\ddagger}$=4 Tx$\Delta\Delta S^{\ddagger}$=-0.5	$\Delta\Delta G^{\ddagger}$=2.8 $\Delta\Delta H^{\ddagger}$=2.6 Tx$\Delta\Delta S^{\ddagger}$=-0.2	$\Delta\Delta G^{\ddagger}$=1.4 $\Delta\Delta H^{\ddagger}$=1.3 Tx$\Delta\Delta S^{\ddagger}$=-0.1

Another interesting aspect revealed are the results obtained in the reactant well with the reconstruction of the FES after metadynamics simulations. In figure 4.11 we present the $F_T(s)$ of the *cis* and *trans* curves.

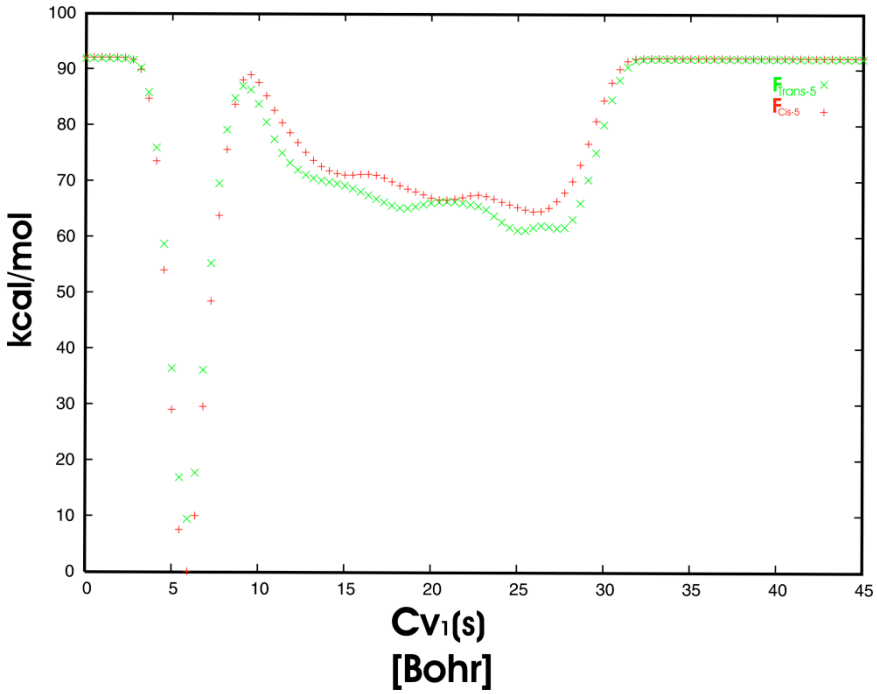


Figure 4.11The free energy surfaces of the *cis* and *trans* (red and green points respectively) conformers of the 1,3,8-nonatriene.

Wecan see in the 10 bohr <**s**<30 bohr region that represent the reactant well region, that the curve presents some local minima with low energetic barriers of ~ 5 kcal/mol between them. These barriers between these local minima are nothing else than the equivalent reactants conformers. This supposition is sustained by the optimizations of the reactant in different conformations and by the fact that the barrier obtain between these conformers are around 1.5 kcal/mol at the PM6 level of theory. The minima are demonstrated by the frequencies calculated for these conformers and all of them present positive frequencies. The most stable conformation of the reactant for the *cis* and *trans*conformers is in both cases the linear conformation that corresponds to the maximum value of the CV1. In the metadynamics simulation, wehave obtained for the most stable reactant conformer a value of CV1 of around 30 bohr and in the case of the optimizations we find a 32 bohr value. These close values show the correlations between the two independent methods.

4.5.3 1,3,9 – Decatriene (6) and 1,3,10 – Undecatriene (7)

The *cis* and *trans* parameters of F(s) of 1,3,9 –decatriene are similar to the 1,3,10–undecatriene results. The only difference is noticed for the width of the reactant well region, which increases since the number of the carbon atoms in the chain has increased. The product well remains well defined with a depth around 80 kcal/mol relative to the transition state. It can be seen that the reactant region is different for **6** and **7** conformations. The two curve of $F_T(s)$ in that region monotonically decreased to a global minimum situated around s=32.5 bohr and s=35 bohr for the **6** and **7** respectively. This difference corresponds more less to a carbon-carbon bond distance, which corresponds to the difference between the structures of **6** and **7.**

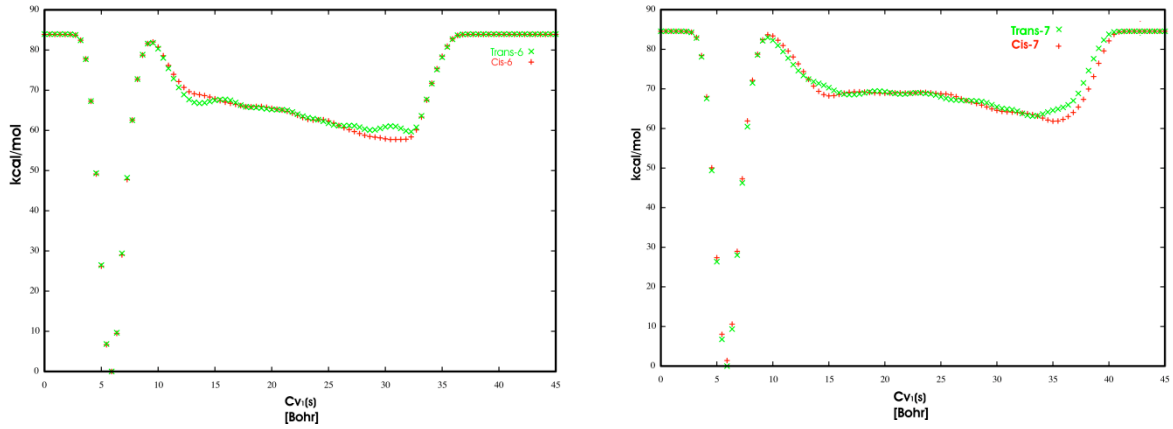


Figure 4.12. Free energy surfaces reconstructions for the Cis and Trans (red and green points) conformations of the 1,3,9 – decatriene (6) **(left side) and 1,3,10 – undecatriene** (7) **(right side).**

This behavior of these two curves in this 15 bohr <**s**<35 bohr region is controlled by the tail of the Gaussians added in our metadaynamics simulations. Because of the $\delta_l = 1$ bohr value set to the geometrical collective variable it is impossible to discriminate with this value the presence of all the local minima. So we can see from these two simulations that the local minima are not so well represented since the reconstruction of this region is affected by the size of the added Gaussian. The solution to this inconvenient is to choose the parameters for the metadynamics to be proportional to the local minima in this region. But this task is beyond the scope of this study. Of course if we choose smaller values forthe width (δ) and the height (w) of the addedGaussians to overcome this problem the time of the simulation increases. Nevertheless, our goal is to find the barrier and the energy difference between the reactant and products, which can be achieved with the parameters used.

The termodynamical parameters are computed by metadynamics and static determinations as previously discution in the case of the 1,3,8 – nonatriene (**5**) complex. These are presented in the next table:

Table 4.3Relative activation parameters for the Intramolecular Diels-Alder Reactions of 1,3,9-decatriene and 1,3,10-undecatriene (7) calculated by metadynamics and static simulations.

Conformers	Metadynamics kcal/mol	PM6 kcal/mol
(**6**) *cis –trans* at 250K	$\Delta\Delta F^{\ddagger}$~0.0 $\Delta\Delta U^{\ddagger}$~0.0 Tx$\Delta\Delta S^{\ddagger}$=1.5	$\Delta\Delta G^{\ddagger}$=0.2 $\Delta\Delta H^{\ddagger}$=-0.1 Tx$\Delta\Delta S^{\ddagger}$=0.3
(**7**) *cis –trans* at 250K	$\Delta\Delta F^{\ddagger}$~0.0 $\Delta\Delta U^{\ddagger}$~0.0 Tx$\Delta\Delta S^{\ddagger}$=3.0	$\Delta\Delta G^{\ddagger}$=-1.0 $\Delta\Delta H^{\ddagger}$=-1.0 Tx$\Delta\Delta S^{\ddagger}$~0.0

From theTable 4.3, we see that we can not discriminate between the reactant and the transitions state for the *cis* and *trans* conformers and so no selectivity is expected for these reactions. This fact is confirmedby the experimental product ratio of formation of*cis* and *trans*of 50-50 at the end of this intra-molecular Diels Alder reaction. These values are in excellent agreement with the experimental[12] data of these reactions, which for **6** and **7** reaction are essentially no stereoselective, with only a few percent excess *cis* product formed in both cases (51:49; 53:47 for **6** and **7** respectively).

In conclusion concerning the Diels Alder intermolecular reactions we can reproduce the experimental data using metadynamics simulations in combination with the developments made around this subject in our research group. Thecalculations were doneusing the Energy as a collective variables and also using Entropy application to reconstruct the $F_T(s)$, $U_T(s)$ and $S_T(s)$ starting from the free energy surface reconstructed on a grid in the space of the collective variables.

4.6 Bias Exchange molecular dynamics

One main drawback of metadynamics is the limited number of CV. Usually 2 CVs are used and using more CV is very time consuming since the time of the simulation scales exponentially with the number of CV. But one advantage of studying chemical reactivity using dynamics is to be able to deal with reactivity in solution. Indeed the number of minima and the flexibility of such systems are hardly compatible with static calculations. Nevertheless, “simple” reactivity like following a proton inside H-bonded network is not a chemical process easily described by one or two collective variables.

Solutions to this issue have been proposed in conjunction with the metadynamics framework, they are the use of Multiple walkers[3], Bias-exchange metadynamics[2] (BEM), and parallel tempering[4] (PT) metadynamics.

The last two methods use multiple walkers to improve the sampling but PT requires many simulations at different close temperatures, which is very time consuming. That’s why we decided to implement the extension of the ordinary Metadynamics called Bias-Exchange Metadynamics (BEM) in an external python program (a python library). The capability of the python program, combined with CP2K suite programs willovercome the limitations of the classical Metadynamics simulations.

My task within this project was to write and create this modulus and to test it. The test revealsif this modulus is correctly implemented, using some test cases suitable to the procedures described in BEM reference [2].

The result is thatafter a rigorous testing procedure, our group should be able to use this application for serial production of different biological applications or reactivity study in explicit solvents.

Because Bias-Exchange Metadynamics is an extension of the Metadynamics simulations, we can easily combine a computational code like CP2K, which performs efficiently metadynamics, with an external driver code to perform BEM. For the external driver code our option was for the Python language because it is an object oriented language and because it is very flexible when it comes toprocessing files.

4.6.1 Bias Exchange technical details

Two files compose the structure of this external code-named “bias_exe.py”. Thefirst one is the main code which contains the main procedure of the Bias Exchange Metadynamics and the second one is the library file where are gather all the procedures and functions which are necessary to performs BEMD. This is useful because if someone wants to switch from the CP2K suite of programs to another application, which performs Metadynamics, the only part that has to be written, is another library for this specific program. The main program of BEMD remains unchanged. The strategy is represented in the next flow diagram that emphasizes the concept of BEMD.

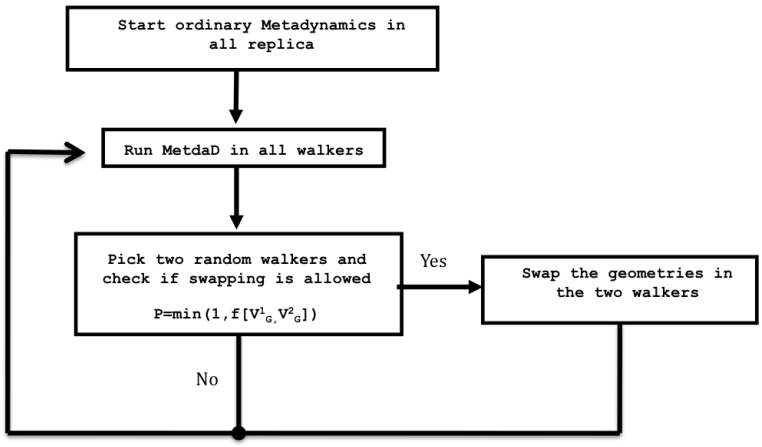


Figure 4.13Bias Exchange Metadynamics logical scheme

The program starts from a general input for CP2K where the all walkers are defined together and where the parser function of the program creates all the input files for the walkers. In the next step of the application the parameters are loaded for this simulation from an external file “PARAM_BIAS_EXCHANGE”. These parameters are:

- $\tau_{exchange}$ defines the time of swapping between two random walkers.
- Number of cycles of BEM – represent the number of loops of simulations required by the user.
- The restart option is responsible for the restart simulation of BEM. This option has two sub options that refer to the walker number and the last cycle number where the simulation must be restarted.

- Number of processors that are used for each walker to do Metadynamics.
- Option to do all the simulations in each walker simultaneously or sequentially. This function is useful when the application can be run to big computational centers.
- Definition of the collective variables for each walker.

The most consuming part is the metadynamics simulations where the bias potential is constructed for each walker. After the bias potential reaches a certain value (numbers of meta-steps in a regular metadynamics run) the swapping condition is verified and the program begins a new cycle in BEM simulation. A simple example will clarify how the swap moves are defined. Consider a system with two walkers and in each walker one relevant collective variable e.g. the distances D1 and D2. Then two replicas are run in parallel, one with a bias on D1 and the other on D2. At a given time $\tau_{exchange}$ a swap is attempted. If the move is accepted, the simulation proceeds as follows:

Coordinates		Before the swap		After the swap	
		CV	Bias potential	CV	Bias potential
Replica 1	x_1	D1	$V_G\left(D1(x_1),t\right)$	D2	$V_G\left(D2(x_1),t\right)$
Replica 2	x_2	D2	$V_G\left(D2(x_2),t\right)$	D1	$V_G\left(D1(x_2),t\right)$

The swap is accepted after we check the condition to the probability of swapping which is $P = \min\left(1, f\left[V_G^1, V_G^2\right]\right)$ with:

$$f\left[V_G^1, V_G^2\right] = \exp\left(\frac{\left(V_G\left(D1(x_1),t\right) + V_G\left(D2(x_2),t\right)\right) - \left(V_G\left(D2(x_1),t\right) + V_G\left(D1(x_2),t\right)\right)}{RT}\right)$$

representing the metadynamics potential for walker 1 and 2. The first two terms in the exponential function are the amount of Gaussians that the system feels before swapping. The second two terms are the Gaussians which system would feel if the swap is done. In other words, the criterion for swapping the walkers is the amount of Gaussians that are feeling before and after a possible swap. If the swap is accepted then the program will change the coordinates of walker 1 with the coordinates of walker 2 and will begin a new cycle of BME.

In the case when the swap is not accepted then in the new cycle of BEM the coordinates of the walkers remain unchanged.

The functionality of the ***bias_exe.py*** has been very well tested and all the exceptions like: numerical stabilities of the python cod, the swapping procedure and the correctness of the values obtain for the collective variables calculations implemented in our code were resolved during this testing period of time.

4.7 Following proton in an aqueous solution

Because our group is interested in studying the reactivity of complex systems in complex media like explicit solvent at different level of theory, we have created an application to follow the proton in an aqueous solution. To be able to discriminate between a normal water molecule H_2O, OH^- anion or H_3O^+ cation our strategy is to use a sum of Gaussian functions centered to the Oxygen atoms and to find the coordination number of O atoms. For normal water molecules the coordination number for the Oxygen atom is two because is coordinated to two Hydrogen atoms. But when we have an OH^- anion the coordination number will became one, or three for the H_3O^+ cation respectively. Because we are able to discriminate between the oxygen atom from a normal water molecule and for oxygen that is in OH- or H_3O^+ conformations at a certain time, as a consequence, we will be able to follow the proton through a Molecular Dynamics simulation also. From a technical point of view, this kind of challenge is divided in two independent parts. The first part is to discriminate between the oxygen's atoms from our system and to find the label of the oxygen that contains the hydroxyl anion or the H_3O^+ cation. The label of the Oxygen means the identification number of the atoms in a molecule, which remains unchanged between the steps of MD simulation. In the first part, what remains to calculate is the coordination number of each O atom after a formula like this:

$$O_{coordination\_nr} = \sum_{i=1}^{N_{Oxygen}} \sum_{j=1}^{N_{Hydrogens}} \exp\left(\frac{(r_{ij} - a)^2}{2\sigma^2}\right) \text{ with } r_{ij} < 1.5 \text{ Å ; a=0.96 and } \sigma = 0.96$$

For parameters "a" and "σ" we choose a value of the optimal water bond which is 0.96 Å, an Oxygen Hydrogen bond in a water molecule. The coordination number is a sum of Gaussians centered to each O and H for a distance between the O and H, less that 1.5 Å. If around O

atoms are one, two or three H atoms then the coordination number became one two or three respectively.

During an MD simulation at each step of dynamics we evaluate the position of the (OH^-) anion or cation (H_3O^+) and we are recording the label of the O atom where this anion or cation is located at this step of MD. Following this procedure during a MD simulation we are able to discriminate and to record the position of this deprotonated or protonated water molecules and used it further on to interpret the results.

Other great advantage to use Gaussian functions is that we can incorporate this method like a new collective variable during a metadynamics simulation. It is know that for a new collective variable is important to have the derivative of this function (CV's) $\frac{\partial s(r)}{\partial r}$ evaluated at each meta-step of dynamics. How s(r) is a Gaussian function with $s(r) = \exp\left(-\frac{(r_{ij}-a)^2}{2\sigma^2}\right)$ the derivative is $\frac{\partial s(r)}{\partial r} = -\frac{(r_{ij}-a)}{\sigma^2}\cdot\exp\left(-\frac{(r_{ij}-a)^2}{2\sigma^2}\right)$ and the metadynamic forces can be evaluated in a precise manner.

Using our approach to a test case, which involves the movement of a proton from a water molecule to another water molecule, we demonstrate the applicability of our method. We use a fictitious, protonated water system, as test case containing two water molecules separated at 4Å (O---O distance). When the proton is equally distant from these two water molecules we find the situation of a Zundel ion. The Zundel ion has a short symmetrical H bond. The system is characterized by a barrier of transition of low energy, between the double well potential regions. Region one is well delimited at the moment when proton is found around the first water molecule and region two delimits itself when the proton reaches the second water molecule. The sketch of Zundel ion is represented in theFigure 4.14

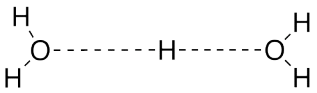


Figure 4.14Zundel ion

To represent the movement of the proton in an aqueous solution we decided to expressthis as afunction of a distance depending on the center of mass of the system and the position of the

proton. There are 2 major situations, one in which the proton acts as free agent in the aqueous solution and the other one corresponding to the case of the proton connected to the water molecule.For our test case the center of mass (CM) is situated between the Oxygen atoms. This is one of the possible situations that we take into consideration. In the other different studied situation, when the proton binds to one water molecule, we choose to compute the distance between the water Oxygen and the mass center of the system. The reason for this option is justified by the fact that we cannot predict which proton will flay apart from this water molecule. Using thisapproach, we were able to represent the function of the distance between the proton and the CM within a molecular trajectory. The results are plotted in Figure 4.15

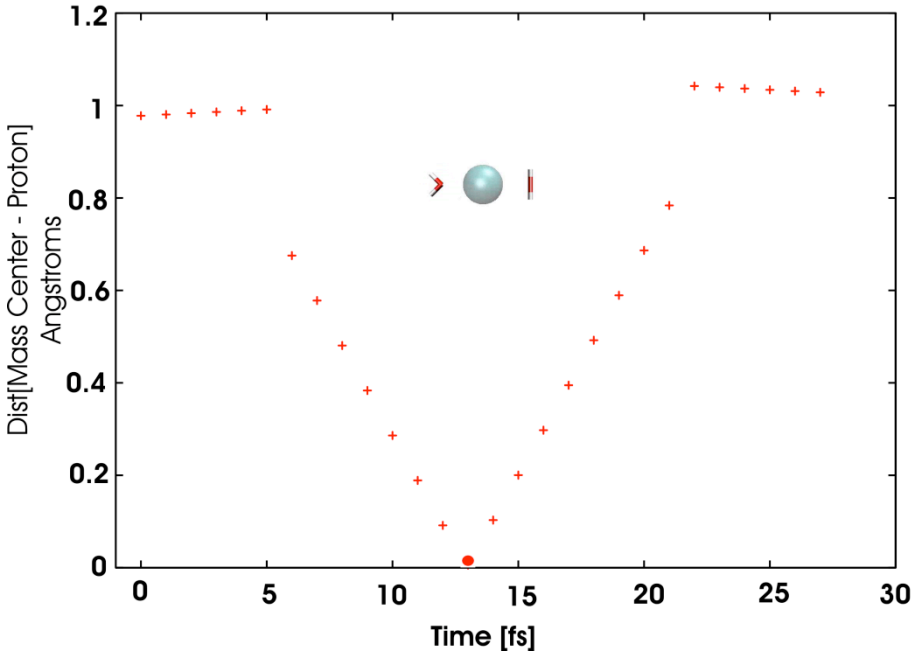


Figure 4.15Zundel ion trajectory.

It is very well defined the situation of the Zundel ion representedin Figure 4.15, showing that the distance CM-proton is about zero. Using this representation of the proton's trajectory we can clearly determine the position of the Zundel ion. This distance is zero because the CM of the system and the proton are onthe same axis (O---O) in this particular test case.

To see how we can treat a much more complex system we choose to use a molecular dynamics trajectory of a system of 48 waters molecules $[H(H_2O)_{48}]^+$. For the $[H(H_2O)_{48}]^+$ system the plot is given in Figure 4.16.

This simulation starts with a distance between the mass center and the proton around 3.2 Å. During this simulation the proton clearly shows a movement from the initial position to the surface of the cluster in the region (0-800 fs) and coming back to the initial start region of the cluster after (800-1500 fs) simulations time.

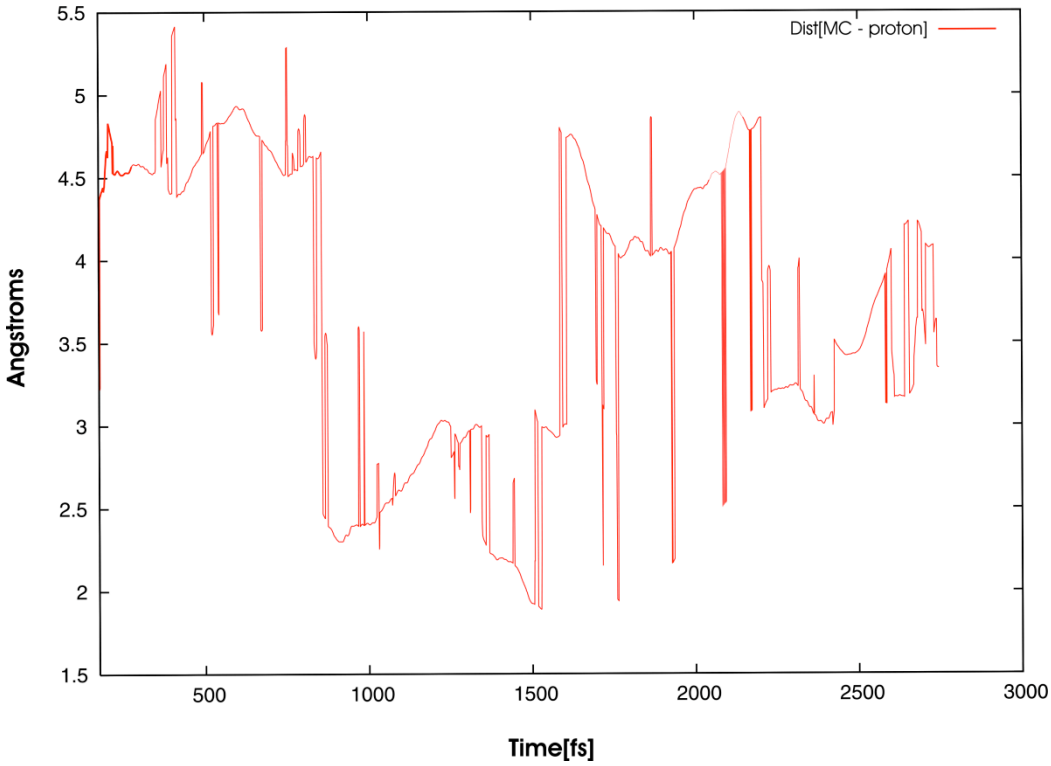


Figure 4.16Distance between the center of mass and the proton in a protonated aqueous solution of 48 water molecules.

The simulation ends with the proton located almost in the same region where it was located atthe initial starting point. In conclusion, during this simulation the proton makes a movement from the center of the cluster to the extremities of this cluster but does not preferred one of these regions. Probably for a simulation of longer length of time, the proton will stabilize in one of these regions (inside or outside the cluster) but is not the case for this one discussed here.

4.8 Conclusions and perspectives

Using the Claisen reaction and intramolecular Dies Alder reactions we were able to validate the “Entropy” application for computing the thermodynamical parameters of the reaction such as internal energy, entropy and the free energy function of the chosen reaction path. We have also proved that, in order to have a good sampling for the system and to find the correct

thermodynamical parameters of the reactions, we must run several independent but equivalent simulations.

The only disadvantage of this approach is that by running more than one simulation, the computational time increases. This application was created with a control of errors through the different convergence criteria, which are set without difficulty by the user. Thus, he can control the optimum ratio between the accuracy and simulation time. But, in the same time, the build-up structure of this application permits the user to set-input the convergence criteria. As a result, the user obtains the optimum ratio between the accuracy and simulation time. This characteristic is the strong point of the application.

We wrote a Bias exchange metadynamics program to be run with cp2k code. It has two great advantages for the future developments. First, the Bias Exchange method by itself is ableto accelerate the metadynamic simulations when a large number of CVs is used, and the second advantage is represented by the fact that CP2K code is very well parallelized and so this interface should be quite useful.

The subject of following the behavior of the proton in a protonated water molecular cluster is a process under development. It is a part of a collaboration started this year between our group and two other groups from ENS Lyon and VU Amsterdam. Although this ambitious project is at a starting point, our preliminary results presented in this chapter are very promising.

Another round-up conclusion is that the strategy to find the thermodynamical parameters by metadynamics simulations and bias exchange metdaynamics simulations could become reliable tools, which can be used successfully in the chemical reactivity on complex media.

Relative bibliography at Chapter 4

(1) Pietrucci, F.; Laio, A. A Collective Variable for the Efficient Exploration of Protein Beta-Sheet Structures: Application to SH3 and GB1. *J. Chem. Theory Comput.***2009**, *5*, 2197-2201.

(2) Piana, S.; Laio, A. A Bias-Exchange Approach to Protein Folding. *J. Phys. Chem. B***2007**, *111*, 4553-4559.

(3) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics†. *J. Phys. Chem. B***2006**, *110*, 3533-3539.

(4) Bussi, G.; Gervasio, F. L.; Laio, A.; Parrinello, M. Free-Energy Landscape for β Hairpin Folding from Combined Parallel Tempering and Metadynamics. *J. Am. Chem. Soc.***2006**, *128*, 13435-13441.

(5) Michel, C.; Laio, A.; Milet, A. Tracing the Entropy along a Reactive Pathway: The Energy As a Generalized Reaction Coordinate. *J. Chem. Theory Comput.***2009**, *5*, 2193-2196.

(6) Fallis, A. G. The intramolecular Diels-Alder reaction: recent advances and synthetic applications. *Can. J. Chem.***1984**, *62*, 183-234.

(7) Ciganek, E. The intramolecular Diels-Alder reaction. *Org. React. (NY)***1984**, *32*, 1-374.

(8) Craig, D. Stereochemical aspects of the intramolecular Diels-Alder reaction. *Chem. Soc. Rev.***1987**, *16*, 187.

(9) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian 09, Revision A.02, Gaussian, Inc., Wallingford CT, 2009.

(10) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.***2005**, *167*, 103-128.

(11) Lin, Y.; Houk, K. N. Intramolecular diels-alder reactions of 1,3,8-nonatriene and 1,3,9-decatriene. *Tetrahedron Lett.***1985**, *26*, 2269-2272.

(12) Diedrich, M. K.; Klärner, F.; Beno, B. R.; Houk, K. N.; Senderowitz, H.; Still, W. C. Experimental Determination of the Activation Parameters and Stereoselectivities of the Intramolecular Diels−Alder Reactions of 1,3,8-Nonatriene, 1,3,9-Decatriene, and 1,3,10-Undecatriene and Transition State

Modeling with the Monte Carlo-Jumping Between Wells/Molecular Dynamics Method. *J. Am. Chem. Soc.***1997**, *119*, 10255-10259.

(13) Goldstein, E.; Beno, B.; Houk, K. N. Density Functional Theory Prediction of the Relative Energies and Isotope Effects for the Concerted and Stepwise Mechanisms of the Diels−Alder Reaction of Butadiene and Ethylene. *J. Am. Chem. Soc.***1996**, *118*, 6036-6043.

(14) Cayzer, T.; Wong, L.; Turner, P.; Paddon-Row, M. N.; Sherburn, M. S. Optimising Stereoselectivity in Intramolecular Diels–Alder Reactions of Pentadienyl Acrylates: Synthetic and Computational Investigations into the "Steric Directing Group" Approach. *Chem. Eur. J.***2002**, *8*, 739-750.

(15) Cayzer, T. N.; Paddon-Row, M. N.; Moran, D.; Payne, A. D.; Sherburn, M. S.; Turner, P. Intramolecular Diels−Alder Reactions of Ester-Linked 1,3,8-Nonatrienes. *J. Org. Chem.***2005**, *70*, 5561-5570.

CONCLUSION GÉNÉRALE ET PERSPECTIVES

Dans le domaine de la réactivité chimique, au cours des dix dernières années, les simulations de la métadynamique est devenu un outil fiable pour décrire les "événements rares" par simulation de la dynamique. L'objectif de cette thèse est l'étude détaillée des structures moléculaires complexes à partir d'un point de vue théorique. J'ai aussi créé et développé de nouvelles applications qui donnent la possibilité de trouver de meilleures approches dans l'avenir pour l'étude de la dynamique de la réactivité chimique des molécules avec leur environnement. De cette façon, nous serons en mesure de calculer les données thermodynamiques, en particulier la contribution entropique. Nous pouvons voir tout mon travail comme un puzzle ou toutes les pièces de ce puzzle ont leur place bien établie, leur but spécifique de contribuer à la grande image, et en particulier, mettant l'accent sur la tendance et les besoins dans le domaine des simulations de la réactivité chimique. Ca implique l'étude du soluté avec le solvant explicite par la dynamique, la description des interactions faibles, des événements rares, processus de association et dissociation et de la contribution entropique.

Dans le deuxième chapitre, sur la base et avec l'aide des méthodes présentées dans la première partie, nous avons souligné l'importance de l'énergie de dispersion. Ayant comme point de départ des derniers développementsdecesinteractionset de la méthode DFT nous avons réussi avec succès l'étude d'un couple de systèmes complexestels que le 1,1'-diméthyl-4, 4'-bipyridinium et ferrocène-bis-pyridinium. Les principaux sujets étudiés dans ce chapitre nous a permis d'entrer directement en contact avec les expérimentateurs. Nos calculs nous ont permis de proposer un modèle géométrique de ces structures, quasi-impossible à cristalliser expérimentalement. Nous avons également été en mesure de quantifier la différence d'énergie entre les différentes formes (ouverte, ferme) des nano-objets et de reproduire le processus en question.

Dans le chapitre trois, nous avons développé en utilisant l'approche du Grimme, les paramètres pour les atomes de platine. Utilisation cette méthode appelée DFT-D pour un complexe de platine nous avons étudié une réaction de cycloaddition [2+1] au le complexé du platine avait été utilise comme un composé catalytique dans cette réaction. Les implémentations de ces paramètres dans le code Gaussian09 permettra des futures études, en particulier les systèmes qui dépendent principalement par le terme de dispersion de London et qui contiennent des atomes de platine. Nous avons également étudié les interactions faibles entre les cryptophane et quelques petites molécules. Ces interactions sont gouverne par les contributions de dispersion et il est indispensable de utiliser fonctionnelles corrige pour la dispersion. Le fonctionnelles plus approprié pour décrire correctement ces interactions entre ceux que nous avons testé sur notre problème était le B97D et wB97XD.

Dans le dernier chapitre nous avons développé une méthode à suivre l'entropie pour un chemin réactionnaire par simulations des métadynamiques. Après la mise en œuvre de l'énergie potentielle comme une variable collective dans la nouvelle version du codeCP2K, qui effectue des simulations métadynamique, nous avons valide cette approche étudier une réaction de Claisen et une réaction Diels-Alder intramoléculaire. Ces développements sont engagés à mettre l'accent sur l'utilité de cette nouvelle approche dans l'étude d'une réactivité chimique et de trouver un protocole pour les critères de convergence de simulations des Métadynamiques. Cette nouvelle implantation est très utile pour déterminer les paramètres d'un système entropique et est indispensable lorsque le caractère dynamique de la simulation doit être pris en compte.

Une autre évolution importante est faite dans les extensions des simulations du métadynamique ou l'approche « Bias-Exchange » est mis en œuvre dans un code Python. Ces nouvelles perspectives, développées dans cette partie du ma thèse, sera un grand pas en avant pour les futures recherches utilisent cette nouvelle méthode du métadynamique. Cela signifie que ce que j'ai commencé à montrer dans ma thèse peut être considéré comme une plate-forme de lancement pour des simulations plus sophistiquées dans le domaine de la réactivité chimique, en particulier pour les réactions dans les quelles les solvants explicites sont indispensables.

GENERAL CONCLUSIONS AND PERSPECTIVES

In the field of chemical reactivity, during the last ten years, Metadynamics simulations became a reliable tool to overcome the so-called “rare events” issue in dynamics simulation. The goal of this thesis was the detailed study of complex molecular structures from a theoretical point of view. I have also created and developed new applications that give the opportunity to find better approaches in the future for the study of the dynamics of the chemical reactivity of the molecules with their environment. In this way, we will be able to compute the thermodynamic data, especially the entropic contribution.We can see all my work as an Jigsaw Puzzle where all the pieces of this puzzle have their well established place, their specific purpose to contribute to the big picture, emphasizing the trend and the needs in the area of simulations of chemical reactivity of the solute with the explicit solvent: dynamics, description of weak interactions, rare events, association/dissociation process and the entropic contribution, which comes along.

In the second chapter, based on and with the help of the methods presented in the first part we emphasized the importance of the dispersion energy. Having as a starting point the latest developments for these interactions and the DFT method we have managed to successfully study a couple of complex systems such as 1,1’-dimethyl-4,4’-bipyridinium and Ferrocene-bis-pyridinium. The main subjectsstudied in this chapter allowed us to enter directly in contact with the experimentalist limitations regarding the crystallizations of the reduced form of these complexes earlier mentioned. Our calculations permitted us to propose a geometry model of these structures, quasi-impossible to be crystallized experimentally. We were also able to quantify the energy difference between the different forms of the nano-objects and reproduce the process involved. Based on the theoretical calculations, I was also able to propose other substituent terminal groups of these compounds with the purpose of tuning these molecular machineries to desired functionalities.

In chapter three we developed using the Grimme’sapproach, the parameters for the Platinum atoms with the scope of using this method called DFT-D for a platinum complex used as a

catalytic compound in a [2+1] cycloaddition reaction. The implementations of these parameters in the code of Gaussian09 suit of package will permit to the future studies and calculations to relay on them for all systems, especially those systems which are predominantly dependent by the London dispersion term and which contain Platinum atoms. We also have studied the weak interactions between the Cryptophane and some small molecules. These interactions are governed by the dispersion contributions and it is indispensable to used functionals corrected for dispersion. The most suitable functionals to describe correctly these interactions among the ones we tested on our problem were the B97D and wB97XD functionals.

In the last chapter we developed a method to follow the entropy for a chemical reactivity pathway *via* metadynamics simulations. After implementation of potential energy as a new collective variable in the new version of CP2K code, that performs metadynamics simulations, we have expanded thisstudy to the Claisen reaction and to a Dies Alder intramolecular reaction. These elaborations are undertaken to emphasize the utility of this new approach in the study of a chemical reactivity and to find a protocol to the convergence criteria of a Metadynamic simulations. This new implementation is very useful to determine the entropic parameters of a system and is indispensable when the dynamically character of a simulation must be accounted.

Another important development is made in the extensions of the Metadynamics simulations were the Bias-Exchange approach is implemented in a new external Python code. Using Python as programming language for these applications is an advantage for the flexibility of the cod. These new perspectives, developed in this application, combined with the discovered utility of the available CP2K code concerning metadynamics simulations (new collective variables, and functions which combines collective variables), will be a big step forward for the near-future researches with this relatively new method namely Metadynamics. This means that what I have started to demonstrate in my thesis can be seen as a launching platform for more sophisticated simulations in the field of chemical reactivity, especially for those reactions in which explicit solvents are indispensable.

My highest hopes concerning this complex research field of Metadynamics are that I will continue to re-discover and implement, by demonstrated calculations and computer-based simulations, it’s capabilities in my future publications and collaborations.