QM and QM/MM studies on organic and bioinorganic systems

PhD Thesis

Alfons Nonell-Canals

Supervised by Feliu Maseras i Cuní and Jordi Villà i Freixa

INSTITUT CATALÀ D'INVESTIGACIÓ QUÍMICA (ICIQ)

Tarragona, July 2008

Els sotasignants Dr. Feliu Maseras i Cuní, director de grup de recerca a l'Institut Català d'Investigació Química i professor titular del Departament de Química de la Universitat Autònoma de Barcelona, i Dr. Jordi Villà i Freixa, professor agregat del Departament de Ciències Experimentals i la Salut de la Universitat Pompeu Fabra i director del grup de recerca en bioquímica i biofísica computacional de l'Institut Municipal d'Investigació Mèdica,

CERTIFIQUEM

que n'Alfons Nonell i Canals ha realitzat sota la nostra direcció a l'Institut Català d'Investigació Química el treball que porta per títol:

QM and QM/MM studies on organic and bioinorganic systems

que es presenta en aquesta memòria per optar al grau de Doctor.

I per tal que consti a efectes legals, signem aquest certificat l'1 de juliol de 2008.

Feliu Maseras i Cuní   Jordi Villà i Freixa

Agraïments

Hi ha qui diu que una tesi doctoral és com un viatge, una experiència personal plena de moments difícils, però també de grans i inoblidables estones. Són, sovint, més de quatre anys compartint despatx amb persones que fan les mateixes coses incomprensibles que tu.

Visitants temporals, algun alumne en pràctiques, *pre-docs* i *post-docs*. Enumerar-los a tots seria una tasca difícil i, segurament injusta, tinc por de deixar-me'n algun. Vull parlar però, del David, ara de post-doc a Le France, company de despatx i de pis durant una temporada. Del Pere, el meu amic "Miró", un dels que fa les coses més rares, però que a la vegada sempre està a punt per tot. L'Eva, companya de penes, cursos i *burrocràcia*, amb qui sovint hem compartit les inquietuts del present i del futur, on ens tornarem a trobar?. L'Ata, assessor en LATEX i en allò que faci falta, sempre disposat a aturar la feina per donar-te un cop de mà. Al Torstein, pels moments còmics dels que sovint ens fa gaudir. A la Núria, per estar sempre pendent de nosaltres. I també, tot i que només vam compartir departament durant uns mesos, al Marc, tècnic i amic.

Tot plegat però, no hauria estat possible sense l'oportunitat que, en el seu dia, em van donar els meus directors de tesi, en Jordi i en Feliu. Quan no tenia massa clar quin tren agafar, l'empenta del primer em va ajudar a escollir, i la confiança del segon, em va donar l'oportunitat de començar l'aventura (anar en tren sempre és una aventura!). Gràcies als dos per la vostra empenta, constància i paciència.

Fer la tesi m'ha permès viatjar, a vegades sortides puntuals per un congrés, però també estades més llargues en algún racó de món. Vull donar les gràcies al Professors Morokuma i Siegbahn per haver-me acollit als seus grups, a Atlanta i Estocolm, respectivament. En aquestes estades he tingut la sort de coincidir

amb grans persones, especialment en Rajeev, l'Stephan i en Holger, gràcies per haver-me ajudat a sentir-me un xic més com a casa.

A tots els amics que han patit els meus moments d'histèria amb estoica paciència. Al Jordi, filòleg de capçalera i company de lluites. Al Jordi (un altre), company de facultat i aviat, també doctor, ànims!. A la Maria, *vecina*, oi que m'enténs?. A la Laura i el Guillem, l'Ariadna, la Isabel, la Carme i el Joan, i l'Elena i l'Albert.

Avui però, tampoc estaria escrivint això si no fos pels que em van donar la possibilitat d'arribar allà on em digués el meu propi límit. Gràcies als meus pares, germanes i germans. I al Pep, tiet i mestre, que em va introduïr en el món de la recerca.

Han estat poc més de quatre anys. Com he dit, moltes persones hi han estat presents. Però hi ha qui sempre ha estat "sempre a punt", disposada a donar-ho tot. Qui ha mirat de compendre l'incomprensible per aconsellar-me i donar-me un cop de mà. A qui m'ha escoltat i animat quan ha calgut. Gràcies M.Elena!

Amb aquestes ratlles arribo al final de trajecte. Sé que tinc un nou tren que m'espera i al que ben aviat pujaré. Projectes nous i engrescadors. Però, encara que sigui un tren carregat amb més experiència i noves il·lusions, sé que serà diferent. Ja ho diuen, una tesi és una experència personal. Gràcies a tots i totes els que ho heu fet així.

Barcelona, 1 de juliol de 2007

8

Contents

CONTENTS

12

Chapter 1

Introduction

The application of computational chemistry has been expanding in recent years, leading to work in a wide variety of fields. This thesis collects the contributions we have made in the last years through theoretical chemistry in two particular fields: asymmetric synthesis involving large organic systems, and selenoproteins. The unifying factor is the use of the same computational methods, DFT and DFT/MM methods, and the focus in problems of experimental interest.

This introduction is organized in four sections. The first one presents briefly an overview of current methods and applications of computational chemistry, with particular focus on the topics studied in the thesis. The other three sections summarize the available knowledge on each of the experimental topics that will be studied in detail in the following chapters: helicenes, guanidinium organocatalysts and selenoproteins.

1.1 Computational chemistry

Computational chemistry is becoming increasingly closer to understanding and solving key chemical problems. Current applications range from heterogeneous

catalysis to enzymatic biochemistry.[1] One of the main reasons for this progress is, without any doubt, the improvement of the computer resources, the continued increase of the computing power allows to dream with the accurate computational study of larger and larger systems.

But these computer improvements are not the only reason for the current situation. Technology has been helped by the development and implementation of new and more performing methods in more efficient software. New techniques help us to understand how chemistry works. Nowadays, computational chemistry is no longer seen as an oddity by most experimental research groups. In the community of biologists, computation is being increasingly labeled as the *in silico* approach, to complement the traditional classification between *in vivo* and *in vitro* studies.

In addition, some developing countries have computational chemistry and biology as a way to develop their own scientific programs.[2]

1.1.1   Modern methods of computational chemistry

Different approaches are available for the study of a given system. **Quantum mechanics**[3–5] (QM) methods offer detailed information of the molecule, as the electronic structure of the atoms or the nature of bond cleavage/formation processes, but these methods require high computational cost. Quantum mechanics methods are appropriate to study a gas-phase reaction in detail, or the processes at the active center of a protein. **Molecular mechanics**[6,7] (MM) methods offer less information about the system, but are more affordable. One cannot obtain the electronic structure of the molecules, but the electrostatic or steric effects of the "molecule environment" can be included. The strong point of these methods is that they can be applied with moderate computational resources to the study of large systems as proteins.

1.1. Computational chemistry

There is a third choice which consists in **hybrid QM/MM**[8–13] methods. With these methods one can apply a quantum molecular method to the most important part of the molecule, and a molecular mechanics method to study the rest of the system. In a protein we could apply a quantum mechanics method to describe the reaction in the active center, and a molecular mechanics approach to the rest.

Some popular versions of these methods will be briefly reviewed in what follows.

Density Functional Theory (DFT)

Most quantum chemistry methods focus on the description of the stationary states of a system of N electrons moving in a potential field created by M nuclei. These stationary states are defined by the solutions of the time-independent Schrödinger equation:

$$\hat{H}\Psi = E\Psi \tag{1.1}$$

As an exact solution to this equation is only available in very particular cases, the goal of computational chemistry is to find an approximation. One approach to the ground-state wave function is provided by the Hartree-Fock[14–16] method, which assumes that this wavefunction can be described by a single Slater determinant of single-electron functions called molecular orbitals. The method struggles when there is a little energy separation between the energy of occupied and empty orbitals, which is a particularly serious problem when working with transition metal systems. Results can be improved in a systematic way by the use of variational or perturbational techniques in what has been labeled as post-Hartree-Fock methods.[16] The computational cost associated to

15

post-Hartree-Fock methods increases sharply.

Density functional theory (DFT) methods[17] do not focus on the many-electron wavefunction but on the electron density. The computational advantage of this concept, is that while traditional methods based on the Hartree-Fock approach deal with a wave functions of 3N variables (one spatial variable for each of the N electrons of the system), the electron density is a function of only three variables. Density functional theory is based on the idea that the energy of the system can be expressed as a functional of the electron density. This theorem was proved by Hohenberg and Kohn in 1964:

$$E_0 = [\rho] \tag{1.2}$$

The most common implementation of DFT is the Kohn-Sham formalism, where the electronic energy can be expressed as:

$$E(\rho) = T(\rho) + E_{ne}(\rho) + E_{ee}(\rho) + E_{xc}(\rho) \tag{1.3}$$

Where $T(\rho)$ is the kinetic energy of the non-interacting electrons, $E_{ne}(\rho)$ the nuclei-electron interaction, $E_{ee}(\rho)$ the electron-electron repulsion and $E_{xc}(\rho)$, the exchange correlation energy. The exact expression for the last term, the exchange-correlation energy, is not known. We can only obtain an approximate value for it using one of the different exchange and correlation functionals that have been proposed.

The most simple functionals are based on the description of the uniform electron gas. In the Local Density Approximation (LDA) the energy depends exclusively on the value of the electron density, but not on its derivatives. This method works properly for some systems, but its use is limited in the study of systems with transition metal atoms because of intrinsic inaccuracy. The work with transition metal compounds requires functionals where the energy

depends explicitly also on the gradient of the local electron density. These methods are labeled as Generalized Gradient Approximation (GGA).

The better results are obtained from functionals that combine LDA, GGA and HF exchange energies. These functionals are called hybrid functionals. The density functional used in this work is Becke3LYP. This functional is a hybrid functional, that uses the hybrid 3-term exchange functional proposed by Becke[18] and the correlation functional proposed by Lee, Yang and Parr.[19]

One of the main limitations of DFT is the lack of a hierarchy allowing systematic improvement of the quality of the results.

Hybrid QM/MM

Work with systems including transition metal atoms, like metalloproteins, is not easy. The electronic complexity of transition metal atoms makes necessary to work, at least, with modern post-Hartree-Fock methods.[16] But these methods are computationally demanding and scale poorly with the increase of the size of the system. To solve this problem, model systems have been traditionally used.[20–22] These models include critical parts of the protein, catalyst, ... but are quite limited in size.

An affordable solution is provided by hybrid QM/MM methods. The system is separated in different regions and the quantum mechanics method is applied only to the core of the system, where the reaction takes place, and a molecular mechanics description is applied to the rest. In the example of the metalloprotein, a QM method can be applied to the metal and other atoms involved in the reaction and a cheaper MM method to the "environment", the protein. The application of these methods, originally designed for biochemical systems,[23] is increasing in all fields of computational chemistry involving

relatively large molecules, from metallocatalysts to metalloproteins.

A general equation to calculate the QM/MM energy for a system is:

$$E_{tot} = E_{QM_{QMregion}} + E_{MM_{MMregion}} + E_{interaction_{QM/MM}} \tag{1.4}$$

Where $E_{QM_{QMregion}}$ is the quantum mechanics energy of the QM region, $E_{MM_{MMregion}}$ the molecular mechanics energy of the MM region, and $E_{interaction_{QM/MM}}$, the energy of the interaction between both regions described at the corresponding level (QM or MM). There are lots of QM/MM methods[10,24–28] and the key difference between them remains in the calculation of this $E_{interaction}$ term. This interaction can be described both at the QM and the MM levels:

$$E_{interaction_{QM/MM}} = E_{QM_{QM/MM}} + E_{MM_{QM/MM}} \tag{1.5}$$

$E_{QM_{QM/MM}}$ is usually referred to as electronic embedding while $E_{MM_{QM/MM}}$ is labeled as mechanical embedding. Evaluation of this mechanical embedding term does not require to recalculate E_{QM} and is relatively easy to introduce if one applies the same force field of the MM region. The $E_{MM_{QM/MM}}$ term provides an estimation of the steric effects but it is less accurate in the description of the electrostatic interactions. These steric effects come from interactions between both non-bonded and bonded atoms. For the case of non-bonded atoms, the van der Waals terms are usually well described by MM. For bonded atoms, steric effects are usually related to strain which is correctly reproduced by the force field parameters (stretching, bending, torsion).

The electronic embedding, $E_{QM_{QM/MM}}$, involves the introduction of the electronic effects. It means that the QM calculation of the QM region must be done taking into account the presence of the MM region. This electronic embedding is usually introduced as a pure electrostatic embedding, and the MM region is seen for the QM calculation as a set of charges often placed at the same position

1.1. Computational chemistry

of the MM atoms.[29]

A significant complication appears when there are bonds across the QM/MM boundary. The particular situation when a QM atom is covalently bound to a MM atom requires a specific treatment during the QM part of the calculation. There are two possible solutions: local orbitals or link atoms. In the local orbitals approach the dangling bond is represented by a special orbital. These local orbital methods, like the local self-consistent field (LSCF) and generalized hybrid orbital (GHO) approaches, provide a QM description of the charge distribution around the boundary, but their implementation is far from trivial.

The link atom approach is more simple. In this case a link atom is used to saturate the dangling bond during the QM calculation. This atom is usually a hydrogen[9] although other proposals have been made.[30] During the QM calculation of the QM region, all MM atoms covalently bonded to a QM atom are replaced by another atom, like hydrogen. The position of this link atom depends also on the particular QM/MM method.

For the computational chemistry of transition metal complexes the most simple of the approaches described above are sufficient: one can use pure mechanical embedding and link atoms. This has been proved by a number of publications.[31–37] The two QM/MM methods most applied in this field are the IMOMM and ONIOM methods, developed by Morokuma and co-workers.[10,26,27,38,39] The simplest ONIOM formulation is presented in equation 1.6, which represents a calculation with pure mechanical embedding. In this equation, $E_{tot_{system}}$ is the total energy of the full system, $E_{QM_{QMregion}}$ is the QM energy of the QM region, $E_{MM_{system}}$ is the MM energy of the full system and $E_{MM_{QMregion}}$ is the MM energy of the QM region.

$$E_{tot_{system}} = E_{QM_{QMregion}} + E_{MM_{system}} - E_{MM_{QMregion}} \tag{1.6}$$

19

There are also refinements of ONIOM that propose the introduction of electrostatic embedding.[27]

1.1.2 Modern applications of computational chemistry

Computational chemistry is currently being applied to many fields of chemistry. We will focus on this section on their application to the two fields that will be discussed in this thesis: asymmetric synthesis and metalloproteins.

Asymmetric Synthesis

A chiral object is an object that cannot be overlapped with its mirror image. The property is called chirality. A classical example of chiral objects are our hands, two "objects" whose mirror image cannot be overlapped. In chemistry we can find examples of chiral and non chiral molecules. The smallest chiral aminoacid is alanine, its two enantiomers (*R* and *S*) cannot be overlapped. An example of a non-chiral molecule is methane, which is identical to its mirror image (Figure 1.1).
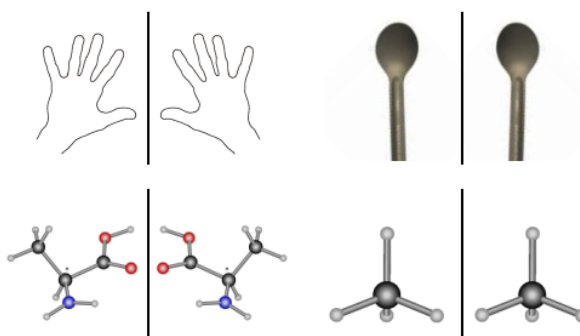

Figure 1.1: Examples of chiral (left) and non-chiral (right) objects and molecules.

Chiral molecules have the physical property of rotating the plane of polar-

1.1. Computational chemistry

ized light, and because of this they can be also referred to as optically active. Optically active compounds are prepared, usually, by **asymmetric synthesis**, this is the procedure to synthesize a chiral product from a non-chiral substrate. It allows to create a stereogenic center with a particular configuration.

The synthesis of chiral compounds is an increasingly important research area in chemistry and, also, in computational chemistry. Chiral compounds are important in pharmaceutical research[40–43] where different enantiomers can have different effects. Public health is important and the governments are developing laws concerning the optical purity of drugs.[44,45] The field of asymmetric synthesis is not limited to pharmaceutical research. It is also important in probes of biological functions,[46–49] new materials[50,51] and polymers.[52–55]

Two enantiomers have the same energy, and because of this, its synthesis is not trivial. If a reaction can follow two enantiomeric paths, these pathways will be isoenergetic and the product will be a 50% mixture of the two enantiomers, a racemate. During an asymmetric transformation, in some point the reaction has to diverge into two non equivalent pathways leading to the *R* and *S* products. One approach to solve the synthetic problem is to use as reactant a chiral compound with a well known and stable conformation. The use of a chiral compound as a substrate is the most simple enantioselective process.

In asymmetric synthesis, to obtain chirality from a non-chiral substrates we can use the chiral compounds in three different roles. A) As chiral reagents, like optically active oxaziridines[56] and peroxides.[57] B) As chiral auxiliaries, like menthol.[58,59] And, C), as chiral catalysts (asymmetric catalysis[60]). Chiral compounds are usually expensive or not easily accessible. Asymmetric catalysis does not require stoichiometric amounts of chiral compounds, and they are recycled for the following reaction. In 2001, the Nobel Prize for chemistry was awarded to Knowless,[61] Sharpless[62] and Noyori[63] "for their work on chirally

21

catalyzed hydrogenation and oxidation reactions".

A usual magnitude to quantify enantioselectivity is the enantiomeric excess (ee, eqn. 1.7). It is quite simple to give an estimation of the enantiomeric excess from computational data. Enantioselective reactions are, usually, under kinetic control, ee can be computed assuming that the final R/S ratio is given by the Boltzmann distribution of population of the two transition states, one for each enantiomer (eqn. 1.8). Easy combination of these two equations (eqn. 1.7 and 1.8) generates a third equation (eqn. 1.9), which we can use to compute ee from the difference of energies ($\Delta G^{\ddagger}_{R/S}$) between the pro-$R$ and pro-S transition states.

$$ ee = \frac{[R] - [S]}{[R] + [S]} \times 100 \tag{1.7} $$

$$ \frac{[S]}{[R]} = e^{-\Delta G^{\ddagger}_{R/S}/RT} \tag{1.8} $$

$$ ee = \frac{1 - e^{-\Delta G^{\ddagger}_{R/S}/RT}}{1 + e^{-\Delta G^{\ddagger}_{R/S}/RT}} \times 100 \tag{1.9} $$

These computational data can be obtained using the methods explained in the previous section (1.1.1). Quantum mechanics (QM) methods allow to know in detail the main aspects of the reaction. In particular, DFT is a good method to elucidate the reaction mechanism for metal-catalyzed reactions. QM methods are good when the electronic effect of the catalyst plays a key role. To obtain information about the steric effects we can use molecular mechanics (MM) methods which are not very computer demanding. But the use of these methods can be seriously complicated when we have to describe a reaction involving formation or breaking of chemical bonds. To work with large systems where both electronic and steric effects are required, hybrid QM/MM methods are often an appealing option.

In conclusion, in asymmetric catalysis we need a good method to describe the reaction center and its immediate surroundings (QM), but to include at the

1.1. Computational chemistry

same time all the steric effects of the system (MM). The combination of both methods (QM/MM) has already been successfully applied to the study of several catalytic processes, like asymmetric hydrogenation, hydroformylation, hydrosilylation, sulfoxidation and olefin dihydroxilation.

Metalloproteins

A general perception that organic chemistry is related with life processes, while inorganic chemistry is not, is far from the truth. There are lots of organic compounds which are relevant far from biological environments, like plastics, and there are a number of inorganic compounds which are fundamental for life. In this second group we can find the metalloproteins.[64] Metalloproteins participate in a number of biologically critical processes, like oxygen transport and cellular respiration. These macromolecules are called metalloproteins because, apart from the organic aminoacids, they contain a metal cofactor, often involving transition metal atoms.

Metalloproteins are only a small subset of the large number of existing proteins, and have some specific features complicating their theoretical study which remains to a certain extent a challenge for computational chemistry. The key is that it is more difficult to compute the energy of a system involving a transition metal center than a purely organic system of the same size.

Metalloproteins are large molecules and we could think in molecular mechanics as the main tool for their study. However, molecular mechanics calculations are not straightforward, because force fields are often poorly parametrized for transition metal systems. Semiempirical methods, often valid to describe reactivity in organic systems, are much more limited in transition metal chemistry. If one moves to methods based in quantum mechanics, things are not much bet-

23

ter. The Hartree-Fock[16] method is also often insufficient even for a qualitative description. The cheapest alternative is represented by generalized gradient approach (GGA) density functional theory[5] and it is already quite computationally demanding.

Another problem is the size of the system on metalloprotein calculations. It is still difficult to introduce a full protein in a QM calculation. In practice, the computational study of metalloproteins is often confined to the description of the reactivity of model systems of the active center at 0 K. The environment effects of the protein are often introduced as a dielectric continuum.
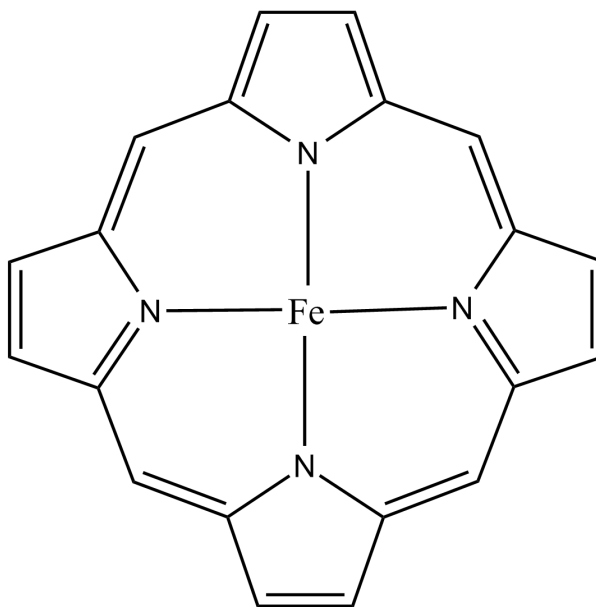
Figure 1.2: The heme group.

Despite the problems mentioned above, computational work in metalloproteins is interesting, because the reactivity of these systems is also richer than that of other proteins, and it is less known experimentally. A number of a reviews on computational studies of metalloproteins have been published in recent years.[22,65–67] An example of the high activity on this field is provided by

1.1. Computational chemistry

the special issue[68] that the *Journal of Computational Chemistry* dedicated in 2006 to the topic of "Theoretical Bioinorganic Chemistry" with 19 contributions, most of them dedicated to metalloproteins. A second example is the high number of publications involving computational studies on metalloproteins. These papers include work on cytochrome P450,[69,70] cytochrome C,[71] other cytochromes,[72] Ni superoxide dismutase,[73] methane monooxygenase,[74] hydrogenases,[75,76] nitrogenases,[77] galactose oxydase,[78] cobalamine and heme groups[79,80] (see Figure 1.2) and blue copper proteins.[81]

There is a small subfield in metalloproteins constituted by the selenoproteins which will be studied in this thesis (see section 1.4). In this field we have to mention the works of Prof. Morokuma and co-workers in the study of the mechanism of glutathione peroxidase (GPx).[31,82]

1.2 Chemistry of helicenes

1.2.1 Thermal stability of [4]helicenequinones

Helicenes are compounds based on ortho-condensed aromatic rings. The smallest helicene is [3]helicene, called phenanthrene, and it consists of three aromatic rings. Because of the intramolecular steric effect, helicenes with 4 rings ([4]helicenes), and higher homologues have a non-planar structure.[83–86]

These non-planar molecules are chiral and, if the interconversion barrier between the two structures (enantiomers) is high enough, is possible to isolate both non-superimposable structures. [3]Helicenes also contain this non-planar feature,[87] but the low interconversion barrier at room temperature does not allow to isolate the enantiomers.[88–90] Helicene enantiomers are called *P* or *PLUS*, if the direction of the helix from the front ring is clockwise and *M* or *MINUS*, if it is counter-clockwise (see Figure 1.3).
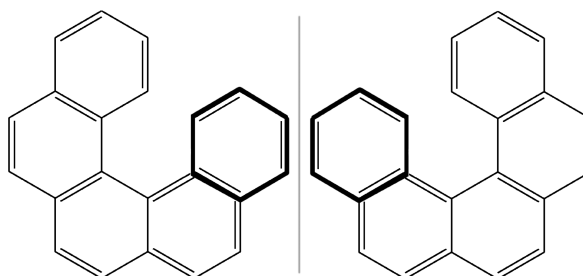
Figure 1.3: Nomenclature of helicenes. The left drawing represents the *P* or *PLUS* structure and the right drawing, the *M* or *MINUS* structure

The stability of both enantiomers depends of the number of rings of the helicene and of the substituents in these aromatic rings. Phenanthrene, [3]helicene, enantiomers are only separable at low temperatures when they are 4,5-disubstituted. On the other hand, enantiomers of [5], [6] and [7]helicenes are

1.2. Chemistry of helicenes

stable at room temperature and their isolated enantiomers can be stored during long periods of time with no significant loss of purity. These larger helicenes have a high interconversion barrier.[91,92]

The situation of [4]helicenes is special. They have a lower barrier than larger helicenes, and therefore it is more difficult to isolate each enantiomer. The eventual isolation of the enantiomers of [4]helicenes is highly dependent on the nature of the substituents on the terminal rings, especially those at the positions 1 and 12 (see Figure 1.4 for the labels). Small substituents allow racemization, but with large substituents the molecules can be solved into their corresponding P and M enantiomers.[93]

The chirality of [4]helicenes was first reported in 1948 by Newman[94] and co-workers. Eight years later, in 1956, the same author[95] prepared and solved the first configurationally stable 1,12-dimethyl-substituted [4]helicene. These 1,12-dimethyl-substituted structures are stable even at high temperatures. Since 1956, other [4]helicene structures have been resolved but their stability depends of the substituents. These data provide the evidence that the racemisation barriers are highly structure-dependent.
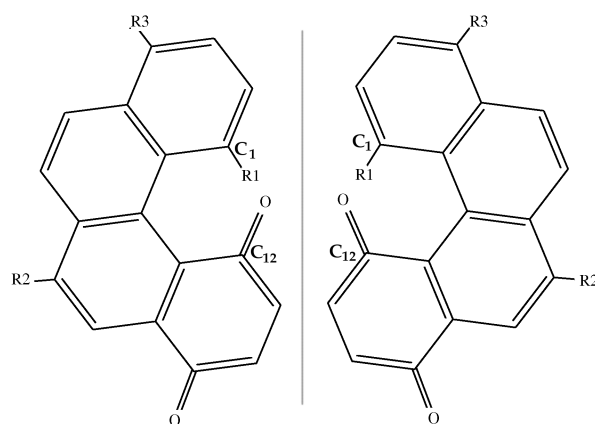

Figure 1.4: Skeleton for both configurations of a [4]helicenequinone.

Most applications of these [4]helicenes are related with their chirality. Some examples are asymmetric catalysis,[96] formation of chiral Langmuir-Blodgett (LB) films,[97,98] or synthesis of optically active bihelicenol.[99] It is thus important to obtain optically pure molecules and, in consequence, to understand how the nature of the substituents and its placement can affect the interconversion barrier in [4]helicenes.

Carreño and co-workers[100] reported the synthesis of a 12-($tert$-butyl)-substituted [4]helicene quinone. The $tert$-butyl substituent confers configurational stability to the system. New works are focused on the study of the 7,8-dihydro[4]helicenequinones (see Figure 1.4). Experimental data show that the inversion barrier of these [4]helicenes depends of the nature of the substituents, specially at position R_1. The chiral stability of these compounds will be investigated in this thesis.

1.2.2 Asymmetric synthesis of [5]helicenequinones

As mentioned above, the synthesis of enantiopure products is a key challenge in the chemistry of chiral molecules. Helicene chemistry is not different and the increase in the range of applications of functionalized helicenes requires efficient and versatile enantioselective synthesis approaches to both M and P enantiomers.

Some years ago, Carreño and co-workers reported asymmetric synthesis methods leading to both enantiomers of [5]helicenebisquinones. The experimental process is summarized in Figure 1.5.[100–103] The reaction begins with the Diels-Alder cycloaddition between [3]helicene, phenantrene, derivatives **1** and the enantiomerically pure sulfinylquinone **(SS)-2**. This produces an enantiomerically pure intermediate **3**, which is a dihydro[5]helicenequinone. **3** reacts afterwards with an oxidating agent and is dehydrogenated to produce the [5]helicenequinone **4**. Final oxidation of **4** leads to removal of the R_1 substituents,

28

1.2. Chemistry of helicenes

and obtention of [5]helicenequinone **5**. This method has been used for the synthesis of [5]helicenequinones and bisquinones using a functionalized vinyl dihydrophenanthrene in the cycloaddition.[103]
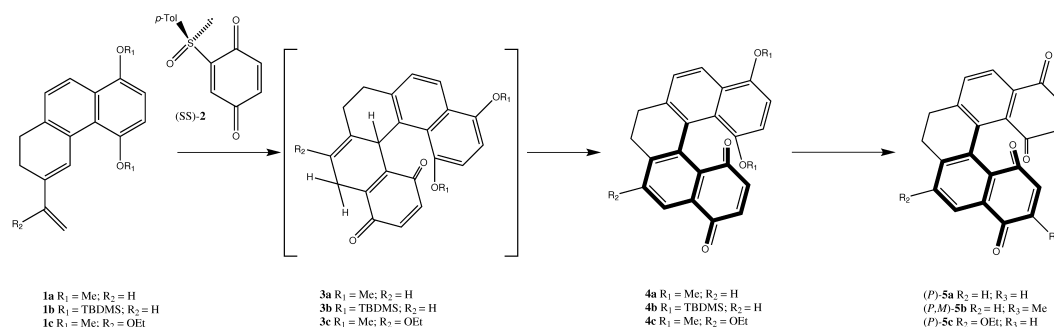
Figure 1.5: Pathway for the synthesis of [5]helicenes reported by Carreño and co-workers.[103] TBDMS is *tert*-butyldimethylsilyl, (^{t}-Bu)(Me)$_2$Si.

Compound **3** is known to be formed in all cases in an enantiomerically pure R form, regardless of the nature of the R_1, R_2, R_3 substituents.[100] However, the final enantiomeric excess of the reaction (product **5**), even the nature of the major enantiomer, depends heavily on the nature of the substituents, especially R_1. The selectivity seems to be decided in the step from **2** to **3**, the dehydrogenation of the dihydro[5]helicenequinone intermediate. The governing effect on enantioselectivity of the substituents was unclear, and we devoted a computational study to the topic.

Different oxidating agents, including **(SS)-2** can be used for the dehydrogenation step from **2** to **3**, but a powerful oxidating agent as 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) seems to produce good results. Remarkably, the nature of the oxidant can also affect the chiral outcome, but we did not investigate this topic.

29

1.3   Guanidinium salts as organocatalysts

Guanidinium is a functional group (Figure 1.6A) with special properties. In biological systems, guanidinium is important for the peptide chain stabilization in proteins, bonding to DNA, molecular recognition,... In chemistry, guanidine derivatives are used as strong bases, especially as asymmetric basic organocatalysts.[104,105]
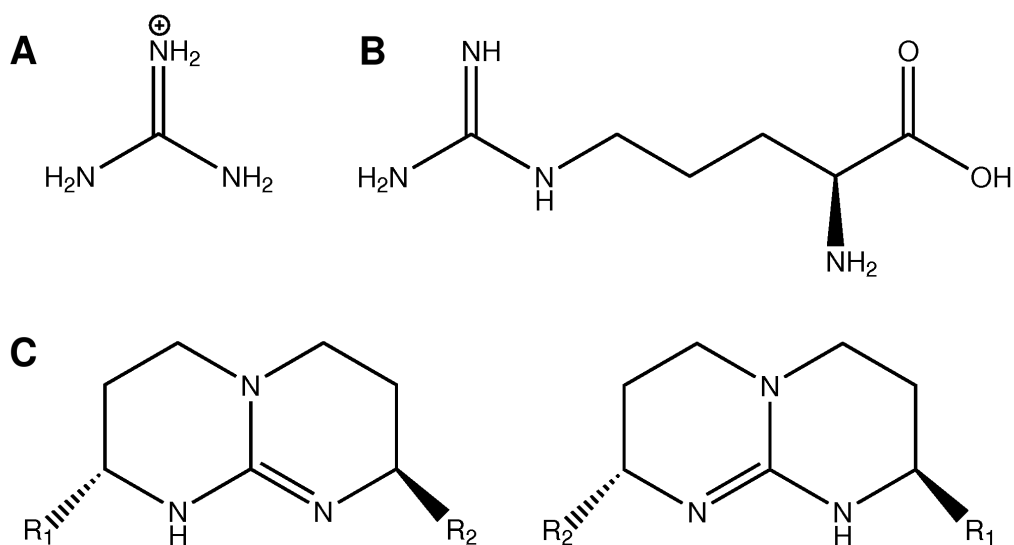
Figure 1.6: A, guanidinium cation. B, guanidine as a functional group in arginine. C, the two enantiomeric forms of a guanidine bicyclic derivative.

Present in the side chain of arginine (Figure 1.6B), the guanidinium cation forms strong ion pairs with oxoanions such as carboxylates or phosphates[106] in enzymes and antibodies. The capacity to bind oxoanions is associated to the Y-shaped, planar shape of the cation, which directs strongly the hydrogen bonding. The high pK_A value (around 12-13)[107] is also critical, because it makes this functional group to be protonated in a wide range of pH. Due to these properties, binding to oxoanions results from both ion-pairing and hydrogen

1.3. Guanidinium salts as organocatalysts

bonding. This interaction will be less likely in high polar solvents as water. Because of this, in proteins the guanidinium-oxoanion interactions usually occur inside hydrophobic pockets or in regions with a low dielectric constant.

Experimentally, in order to improve its solubility in non-polar solvents, guanidinium can be incorporated in a bicyclic framework (Figure 1.6C). These bicyclic guanidinium frameworks can be used for the enantioselective recognition of carboxylates and as asymmetric organocatalysts. The role of these derivatives in enantioselective recognition was first discovered in 1989 by de Mendoza and co-workers.[108]

Guanidinium salts have beeen recently used by de Mendoza and co-workers[109] to catalyze the enantioselective opening of the ring of azlactone derivatives in order to produce quaternary glutamic acid analogues. The reaction is shown in Figure 1.7. This process is part of the new protocols that have been recently described for the synthesis of enantiomerically pure α-aminoacids.[110–112] These quaternary aminoacids are an interesting synthetic target, because their introduction in biologically active peptides often increases their activity.[113] The conventional route to introduce chirality in aminoacids, asymmetric catalytic hydrogenation, cannot be applied in this case.
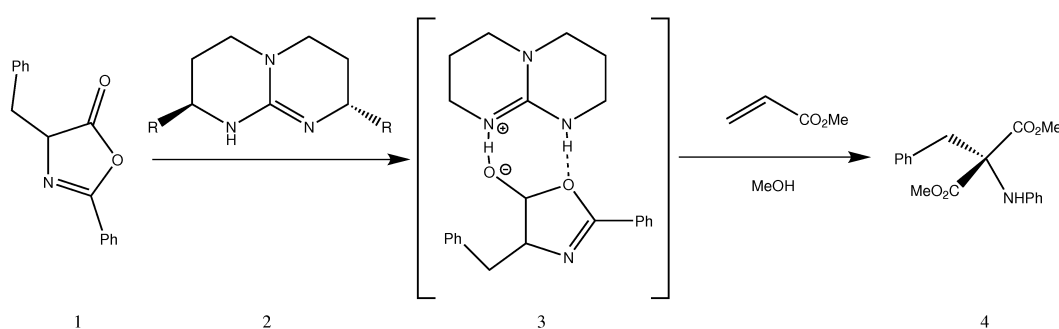
Figure 1.7: Synthesis of a quaternary aminoacid from azlactone using a guanidine derivative as catalyst.

The reaction has been performed using two different guanidine catalysts, **5** and **6**, Figure 1.8. **5**, using dicloromethane (DCM) as solvent at $25\,^\circ$C, produces product **4** with a yield of 46% and an enantiomeric excess of 57%. Both yield and enantiomeric excess are higher with **6**, the corresponding values being 60% and 68%, respectively. Optimization of the conditions with **6**, using toluene as solvent and a temperature of $-78\,^\circ$C produce a yield of 68% with a 87% ee.

Figure 1.8: The two bicyclic guanidine catalysts applied in the experimental study.

The catalytic role of the guanidinium in the ring-opening of azlactone is very likely the formation of an adduct with the substrate (Figure 1.7), which then is more easily attacked by the methyl acrylate. The presence of the chiral guanidinium will eliminate the equivalence between the two faces of the azlactone, and drive the attack towards one of them. The precise structure of the adduct is neverthcles not known, and because of this, we decided to investigate computationally this topic.

1.4 Selenoproteins

Selenoproteins are the proteins that contain an aminoacid called selenocysteine. This aminoacid is a special residue similar to cysteine but containing a selenium atom instead of a sulfur. Insertion of a selenocysteine in the peptidic chain requires a complex mechanism which is specific for this residue. The difficulty of this insertion and the role of selenium as essential nutrient makes these proteins an interesting field in computational biochemistry.

1.4.1 Selenium, the element and health implications

Selenium, in Greek "moon light", was discovered in 1817 by the Swedish physician and chemist Jöns Jacob Berzelius[114] when he was working in the study of the etiology of a strange disease among workers at a sulfuric acid plant in Gripsholm (Sweden). It is a non-metal element with atomic number 34 and Se as chemical symbol. It is chemically related with sulfur (see Table 1.1).

Table 1.1: Chemical properties of selenium and sulfur

Properties	Selenium	Sulfur
Atomic weight	78.96 uma	32.07 uma
Atomic radius	115 pm	100 pm
Covalent radius	116 pm	102 pm
van der Waals radius	190 pm	180 pm
Electronic configuration	[Ar] $3d^{10}4p^44s^2$	[Ne]$3s^23p^4$
Oxidation states	$\pm2,4,6$	$\pm2,4,6$

In contrast to sulfur, selenium was considered a poison until Schwaz and Foltz[115] identified its role as a micronutrient for bacteria, mammals and birds.

Some years after, in 1973, two bacterial enzymes, formate dehydrogenase[116] and glycine reductase,[117] were reported to contain selenium. In this context, the works in formate dehydrogenase carried by Andreseen and Ljungdahl during the 1970's are remarkable.[116, 118–120] During this same period, selenium was discovered in the active center of the antioxidant enzyme glutathione peroxidase (GPx).[121,122] This clearly established the important biological role of selenium in mammals.

Trace amounts of selenium are necessary for cellular functions in most eukaryote and prokaryote cells.[123,124] In contrast, it seems that plants do not require selenium for survival, although they incorporate it non-specifically into compounds that usually contain sulfur when growing in soils containing selenium. One can found selenium within the cells in separate molecules or in the residue selenocysteine as part of proteins.

In humans, there are several health problems related with selenium deficiency.[125] Selenium is needed for the proper functioning of the immune system, has an important antioxidant activity, and is required as a catalyst for the production of the active thyroid hormone. Elevated selenium intake reduces the risk for cancer[126–129] and cardiovascular diseases. Selenium deficiency can accelerate the development of different cancers[130] and decrease male fertility.[131,132]

Anticancer activity is, perhaps, the best known and more relevant role of selenium in health protection. It is probably related with the antioxidant activity of several selenium-containing enzymes, that can act against the oxidative stress inside the cells.[133–137] Glutathione peroxidases (GPx)[138,139] are particularly relevant in this context. GPx's reduce potentially damaging reactive oxygen species, and five of the known GPx's are selenoproteins, The role of selenium in cancer protection has also been associated with the regulation of different cell pathways that induce apoptosis and DNA reparation.[140–146]

Selenium requirements[147] vary according to the age, from the 15-20 μg/day for 0-3 years old to 55 μg/day for 19 years old and older. These requirements are increased up to 60 μg/day in pregnant women and up to 70 μg/day during breastfeeding. Vegetables and meat are the main source of selenium[148] and in a regular diet is difficult to have a selenium deficiency. This deficiency occurs only in regions where selenium content is abnormally low in the soil. Keshan's disease derives its name from Keshan, a region in China with a soil poor in selenium, where during 1935 appeared a heart fatal disease associated to selenium deficiency. The oddity itself of selenium diseases means that it is usually associated with other nutritional deficiencies.

1.4.2 Selenocysteine and selenoproteins

The major biological form of selenium is the aminoacid selenocysteine. An aminoacid is a molecule which contains amine and carboxyl functional groups. Proteinogenic aminoacids are the basic structural building units of the proteins. There are also non-proteinogenic aminoacids with important biologically functions as vitamins and neurotransmitters.
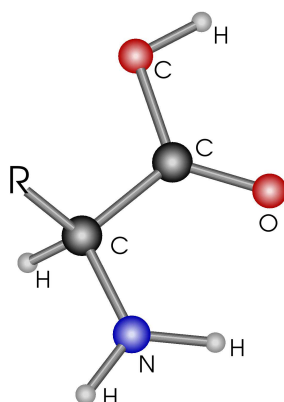
Figure 1.9: General structure of a proteinogenic aminoacid.

The structure of the proteinogenic aminoacids is shown in Figure 1.9. **R** represents the sidechain, which is specific for each aminoacid. The central atom, where R is attached, is a called C_α. C_α is a chiral for all aminoacids except glycine.

There are 20 "classical" proteinogenic aminoacids, which become 21 when we include selenocysteine. The aminoacids form the protein structure through peptide bonds between the amine group of one aminoacid and the carboxyl of another one. The aminoacid sequence is encoded in the genetic code corresponding to each protein. A combination of three letters (codons) from the DNA determines which aminoacid will be inserted. The sequence of the protein is terminated by a special codon named STOP codon. The translation takes place in the ribosomes. Seccys

Figure 1.10: Cysteine (left) and selenocysteine (right).

The selenocysteine structure is like that of the other proteinogeic aminoacids. Its R group is the same than that of cysteine but with a selenium atom replac-

1.4. Selenoproteins

ing the sulfur atom. For this reason, selenocysteine is usually compared with cysteine (see Figure 1.10).

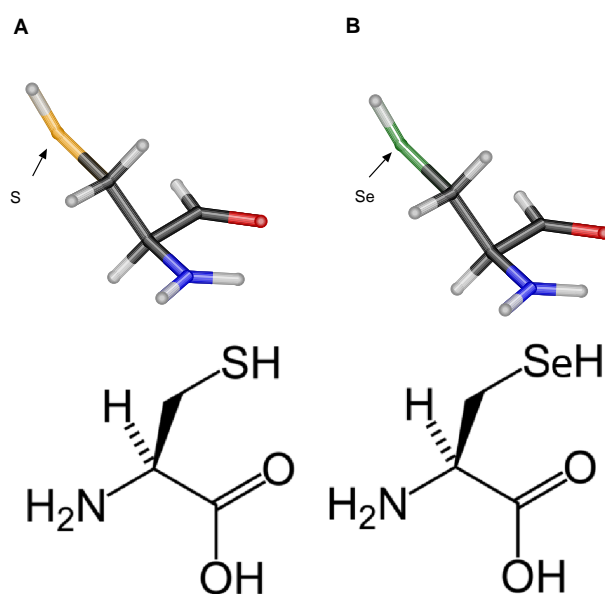The translation of selenocysteine is special because this aminoacid is encoded by a STOP codon UGA. Because of this, its insertion in the peptidic chain requires something else than a "normal" translation (mRNA, ribosome and tRNA). First of all, downstream from this UGA codon there has to be a conserved sequence called SECIS (SElenoCysteine Insertion Sequence) that, because of some internal bonds, has a loop structure. This sequence helps the specific tRNA to bind the mRNA, with the additional help of a protein complex. In eukaryotes, this complex contains two proteins, SBP2 (SECIS Binding Protein 2), which binds the SECIS element, and eEFSec (Sec Elongation Factor), which binds to SBP2 and to a specific tRNA (see Figure 1.11). In prokaryotes, there are some changes in the SECIS sequence, and there is only an additional protein, SelB, helping binding to tRNA.
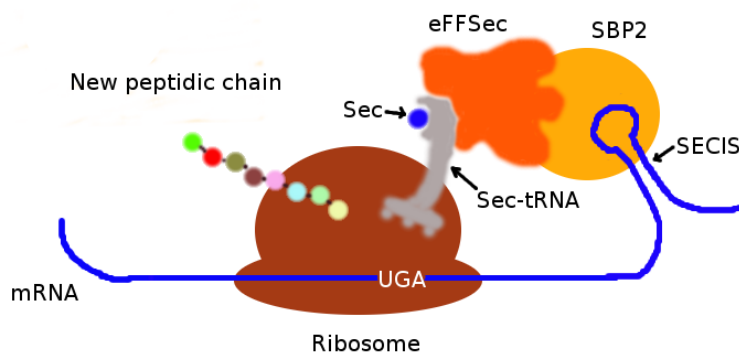

Figure 1.11: Translation mechanism of selenocysteine in eukaryotes.

The number of identified selenoproteins has increased substantially in the last years because of the improvement in the bioinformatic tools allowing the direct search of genome sequences.[139,149–151] The most remarkable selenoenzymes in prokaryotes are formate dehydrogenase,[152] hydrogenases[153] and glycine re-

ductase.[154] In eukaryotes, one can mention iodothyronine deiodinases,[155] thiore-doxin reductase,[156,157] selenoprotein P[158,159] and Methionine-R-sulfoxide reductase (MSR),[160,161] in addition to the glutatione peroxidases (GPx) mentioned above. A summary of selected selenoproteins and its function is shown in Table 1.2.

Table 1.2: Most relevant selenoenzymes and their known function

Enzyme	Catalytic reaction
Formate dehydrogenase	Formate oxidation to carbon dioxide
NiFeSe-hydrogenases	Oxidation of molecular hydrogen (H_2)
Glycine reductase	Conversion of glycine to acetyl phosphate
Iodothyronine deiodinase	Conversion of T4 to T3 and both degradation
Thioredoxin reductase	Reduction of different thioredoxins (TRx)
Selenoprotein P	Antioxidant, protection of endotelial cells
Methionine-R-sulfoxide	Reduction of oxidized methionine
Glutatione peroxidases (GPx)	Reduction of reactive oxygen species (ROS)
Selenoprotein W	Possible role in muscle metabolism

1.4.3   Formate dehydrogenase H (FDHh)

The family of formate dehydrogenases (FDH) is an heterogeneous group of enzymes present in both prokaryote and eukaryote organisms. All of these enzymes have the same function, to catalyze the oxidation of formate to carbon dioxide. The FDH in anaerobic prokaryotes, as is the case for FDHh, oxidize the formate produced from piruvate during anaerobic respiration. They serve as a major electron donor to a variety of respiratory pathways that use terminal acceptors different from molecular oxygen.[162,163]

1.4. Selenoproteins

The active center of FDH enzymes in anaerobic prokaryotes contains a transition metal atom, usually molybdenum or tungsten, and a molybdopterin guanine dinucleotide (MGD) cofactor. Some of these enzymes have been crystallized.[152,164–170]

Formate dehydrogenase H (FDHh) is an anaerobic enzyme that catalyzes the oxidation of formate to carbon dioxide in absence of oxygen. It is activated by the presence of formate and inhibited if there is oxygen in the environment. It contains a selenocysteine in the position 140, and it is thus a selenoprotein. The presence of molybdenum converts it also in a metalloprotein. The crystal structure of this enzyme was published by Boyington *et al* in 1997.[152]

Figure 1.12: Active center of formate dehydrogenase H.
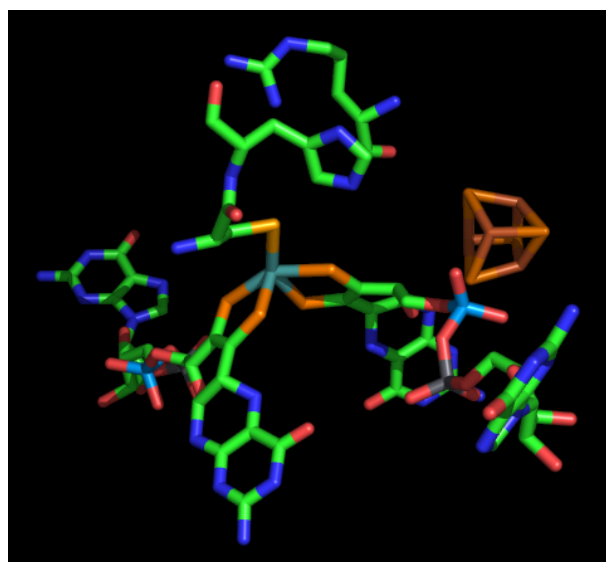
The structure of the active center of FDHh is shown in Figure 1.12. The selenium from selenocysteine is coordinated to the central molybdenum atom. The molybdenum is also bound to four sulfurs from the two molybdopterin guanine dinucleotide (MGD) cofactors. As an electron acceptor, the enzyme has a sulfur-iron box (Fe_4S_4) which receives, one by one,[171] the two electrons

generated in the catalyzed reaction. There are two important aminoacids more, a histidine (His[141]) that transfers the proton released by the reaction and an arginine (Arg[333]) that can stabilize reaction intermediates and transition states.

FDHh is part of a complex called formate-hydrogen lyase,[172] from *Escherichia coli*. This complex decomposes formic acid to hydrogen and carbon dioxide in anaerobic conditions. FDHh catalyzes the reaction shown in equation 1.10. This reaction releases two electrons and one proton which are used for the cell metabolism.

$$HCOO^- \longrightarrow CO_2 + H^+ + 2e^- \tag{1.10}$$
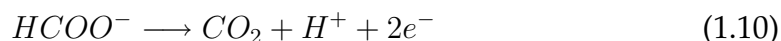
The presence of selenocysteine instead of cysteine in the protein active center implies a dramatic catalytic change. The reaction rate is 300 faster with selenocystein than with cysteine.[173] The reason for this sensitivity is not known, and is likely associated to the essential characteristics of selenocysteine as an aminoacid. Because of this, we decided to study the topic from a computational point of view.

Chapter 2

Objectives

The main aim of this thesis is to solve some organic and bioinorganic, biological and synthetic, questions using both QM and QM/MM computational chemistry tools. These tools have been applied to different systems and problems.

- **Thermal stability of [4]helicenes.** The thermal stability of chiral [4]helicenes depends on the nature of the substituents. The barrier for the interconversion between the P and M enantiomers will be computed for different systems, with the goal of finding the relationship between the nature of the functional groups and the chiral stability of the compounds.

- **Asymmetric synthesis of [5]helicenequinones.** The synthesis of [5]helicenequinones can yield either P or M products. Experimental data show that the configuration of the product depends on the nature of the substituents at the [3]helicene precusor. The mechanism controling the chirality is not well understood, and will be investigated through calculations on the selectivity determining transition states of the key reaction step.

- **Chiral guanidine derivatives as organocatalysts for asymmetric synthesis.** Chiral guanidine derivatives have been recently shown to catalyze

41

the ring opening of a prochiral azlactone to produce a chiral quaternary aminoacid with good enantiomeric excess. The first step of the reaction is complexation between the azlactone and the catalyst. The structure of the adduct, which decides the final chiral outcome, will be analyzed.

- **Role of selenocysteine in the active center of Formate Dehydrogenase H .** Experimental data show the critical role of selenocysteine in the active center of FDHh. We will characterize the FDHh catalytic mechanism and study the effect of the presence of selenocysteine in the active center as alternative to the more usual cysteine.

- **Effect of the nature of the metal in the behavior of different formate dehydrogenases.** Most FDH contains molybdenum as a central transition metal atom. Sometimes, this metal is replaced by tungsten. The reasons and effects of this change remain unknown and will be studied.

Chapter 3

Chemistry of Helicenes

3.1   Thermal stability of [4]Helicenequinones

Experimental data[174] show that the configurational stability of [4]helicenequinones represented in Figure 3.1 and Table 3.1 depends on the nature of the substituents of the tetracyclic skeleton. The influence is largest for R_1, placed in carbon C_1. Different compounds, defined in Table 3.1, have been studied experimentally and their enantiomeric stability has been measured as the ease of a pure enantiomer to racemize, in terms of temperature and time of racemization. For this series of helical compounds the following order for chiral stability has been experimentally found:

$$A, H << D < E << F < G << B, C \tag{3.1}$$

It is clear from this experimental ordering that the R_1 substituent has the largest effect. The racemization is easiest when R_1 is OMe and most difficult when R_1 is ^{t}-Bu, with Me and ^{i}-Pr lying in between. For R_2, the effect is smaller, although the racemization is favored when it is H with respect to OEt. Finally, there are not enough data to make a conclusive observation on the role of the
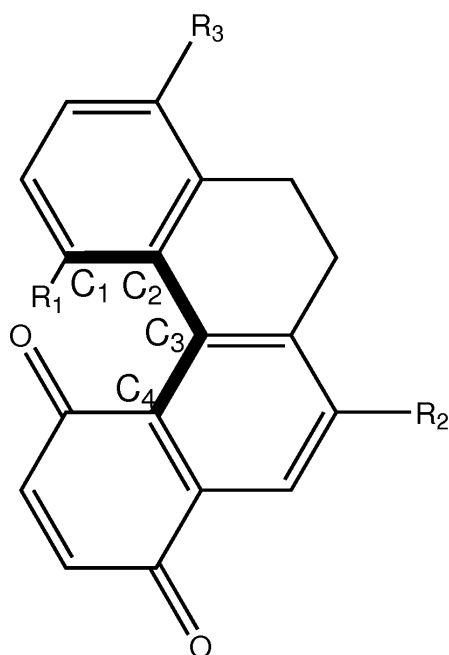
Figure 3.1: Generic formula of the tetracyclic skeleton of [4]helicenequinones under consideration

R_3 position. The systems with OMe in the R_1 position have also an OMe group in R_3 because of experimental synthetic constraints. In any case, the effect of R_3 is likely to be minor.

The energy barrier for the enantiomeric inversion of a given [4]helicenequinone was computed through the evaluation of the free energy difference between two structures, one of the enantiomeric minima, and the transition state connecting it to the other enantiomeric minimum. The relative planarity of these structures is measured from the value of the dihedral angle ϕ, defined as C_1-C_2-C_3-C_4, where the atom numbering is shown in Figure 3.1. The value of ϕ for a planar structure would be zero.

Table 3.1: Labeling of [4]helicene species according to the nature of the substituents

Molecule	R_1	R_2	R_3
A	OMe	H	OMe
B	^{t}-Bu	H	H
C	^{t}-Bu	OEt	H
D	Me	H	H
E	Me	OEt	H
Eb	Me	OH	H
F	^{i}-Pr	H	H
G	^{i}-Pr	OEt	H
H	OMe	OEt	OMe

3.1.1 Computational details

A preliminary set of geometry optimization calculations was carried out with the semiempirical AM1[175] method as implemented in the Gaussian03 package.[176] For each system, the minimum and transition state were computed, and the connection between them was checked. The conformational space was also explored in the cases where the substituents have conformers associated to the rotation of single bonds, as is the case for OEt, OMe and ^{i}-Pr. These AM1 optimized structures structures were used as starting points for full geometry optimizations with the Becke3LYP density functional method[18,19,177] as implemented in the Jaguar program[178] using the 6-31G(d) basis set.[179,180] The nature of the Becke3LYP optimized stationary points as minima or transition states was confirmed by frequency calculations. The reported free energies include zero-point energy corrections and were computed for a temperature of $25\,°C$ and a pressure of 1 atmosphere.

3.1.2 Overall energy profile

We first analyzed in detail the behaviour of a particular system, **E**, which is the one where a more detailed experimental study is available. This structure has a methyl group as R_1, an ethoxy group as R_2 and a hydrogen as R_3. Figures 3.2 and 3.3 present the optimized geometries of the minimum and the transition state, respectively, for this system. An analysis of the structure for the minimum shows a non planar geometry, the torsional angle ϕ is $42.7\,°$. Of course this was fully expected, because it is the non-planarity of this kind of molecules what allows the existence of P and M enantiomers. The transition state between the two degenerate enantiomers is also non-planar, and this was somehow unexpected. The optimized structure for the transition state is shown in Figure 3.3, and the torsional angle ϕ of $31.0\,°$ confirms the lack of planarity.
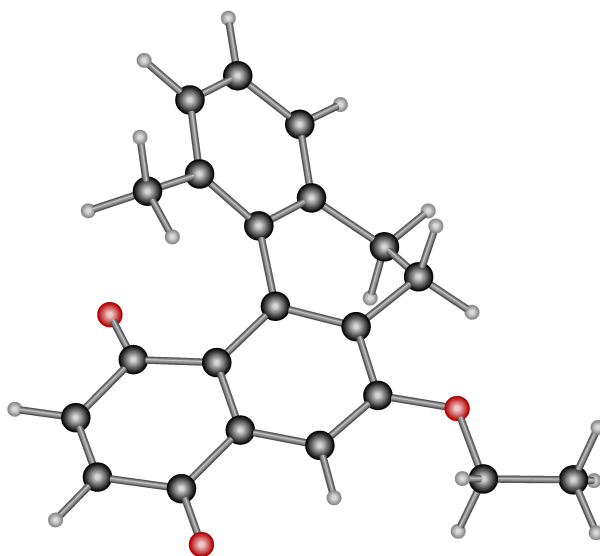

Figure 3.2: Becke3LYP optimized structure of the minimum for [4]helicene system **E**

Of course, a planar structure is the simple way to internconvert two enantiomers, but it will not exist as such if it is too constrained. An inspection of

3.1. Thermal stability of [4]Helicenequinones

the transition state structure shows the origin of this behavior. There are two sources of strain in this type of molecules which push them away from planarity. The first one is the steric repulsion between the carbonyl group of the quinone ring and the R_1 substituent (methyl in **E**). The second one is associated to the presence of two sp^3 carbon atoms in the cyclohexadiene ring where atoms C_2 and C_3 are included. Both strains would be maximized at the same time in a planar structure.
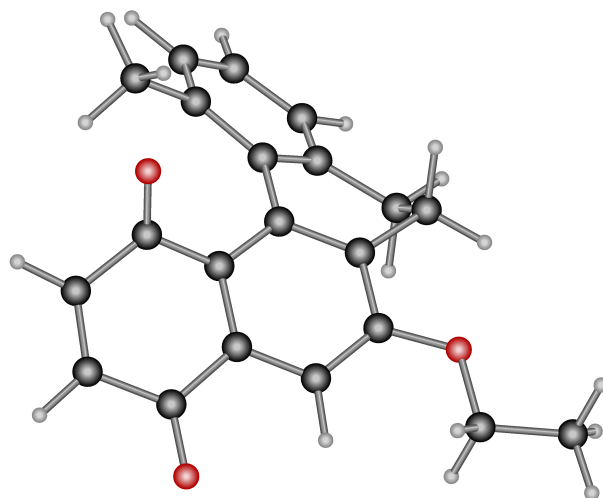
Figure 3.3: Becke3LYP optimized structure of the transition state for the [4]helicene system **E**

In the optimized transition state the steric strain between carbonylic oxygen and the methyl at C$_1$ is at its highest point, but the sp^3 carbons are practically in the same arrangement than in the minimum. A hypothetical C_s structure would be too sterically constrained because the strain in both sp^3 carbon should be added to the strain of the R_1 substituent. It would have a higher energy and more than one negative eigenvalue, not being thus a true transition state. The presence of a non-planar structure implies a lower energy barrier.

Remarkably, the overall symmetry of the system means that there must exist

another enantiomeric transition state with an identical energy to that reported here. We did not explore it further because of its lack of chemical relevance.

We repeated these calculations of minimum and transition state for all sytems **A** to **H** presented in Table 3.1 and we obtained the results summarized in Table 3.2. The most relevant result is that the order of the computed values for $\Delta G^{\ddagger}$ matches exactly the ordering of barriers observed by experiment. This confirms the validity of our computational approach, and of the assumption that the enantiomeric stability is associated to the barrier of this particular interconversion mechanism. Concerning the values for ϕ, it is worth noticing that in general the higher barriers are asociated to both the more planar transition states and the less planar minima.

Table 3.2: Values for the dihedral angle ϕ (degrees) form minimum and transition state, and free energy barrier (kcal/mol) for each [4]helicene computed system.

Molecule	R_1	R_2	R_3	ϕ**min**	ϕ**TS**	Δ**G**
A	OMe	H	OMe	39.0	35.3	22.8
B	^{t}-Bu	H	H	50.5	7.2	39.2
C	^{t}-Bu	OEt	H	50.5	19.1	42.0
D	Me	H	H	43.2	30.6	25.2
E	Me	OEt	H	42.7	31.0	25.9
F	^{i}-Pr	H	H	45.0	16.4	28.3
G	^{i}-Pr	OEt	H	46.0	15.1	29.3
H	OMe	OEt	OMe	38.9	36.4	23.0

3.1.3   Effect of the R_1 substituent

We will analyze in what follows the effect of the nature of the R_1 substituent on the interconversion barrier by comparing the behavior of selected pairs of molecules.

OMe *vs* Me

We analyze first the comparison between the methoxy and methyl substituents. For the rest of the cases, the highest barriers correspond to the bulkiest groups, which suggests a simple explanation based on steric effects. Things are not so clear for the methoxy/methyl pair, because the interconversion is faster for the systems containing the methoxy group.

We compare first the behavior of systems **E** and **H**. **E** has been discussed above, and **H** differs from **E** in the nature of the substituents at the R_1 and R_3 positions. **E** has methyl and hydrogen in these positions, respectively; while **H** has two methoxy groups. The R_2 substituent is OEt in both cases. Figures 3.4 and 3.5 present the optimized geometry of the minimum and the transition state, respectively, for **H**.

As shown previously in Table 3.1, the barrier for **H**, containing OMe, is lower (23.0 kcal/mol) than for **E**, containing Me (25.9 kcal/mol). This was in principle unexpected, because the strain in the transition state is obviously associated to the bulk of R_1, which in the transition state gets closest to one of the carbonyl groups of the quinone ring. The observation of the structures in Figures 3.4 and 3.5 provides however a first clue to the explanation of this behavior. The methoxy group has certainly more atoms than methyl, but is able to place its methyl substituent away from the core of the molecule. The sterically active part in **H** is thus only the oxygen atom, smaller than the methyl of **E**, whith is
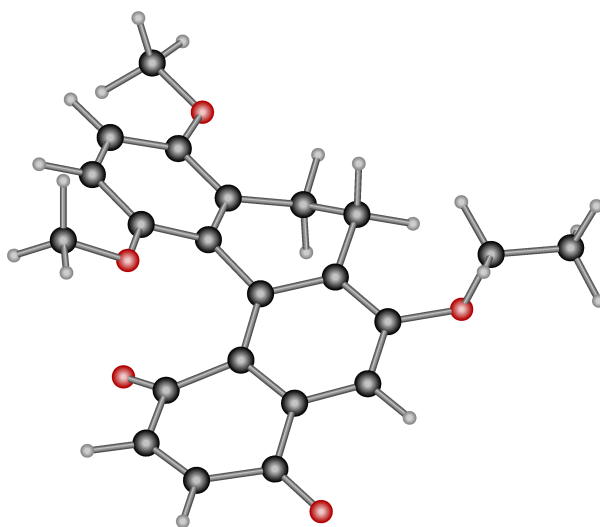
Figure 3.4: Becke3LYP optimized structure of the minimum for the [4]helicene system **H**

three hydrogen substituents.

A similar behavior is observed when comparing structures **D** and **A**. The R_1/R_3 pattern is the same that for the **E**/**H** pair, the difference is that in this case the R_2 position is occupied by hydrogen. The barrier is again lower for the molecule with OMe in the R_1 position (**A**, 22.8 kcal/mol) than for the system with Me in this position (**D**, 25.2 kcal/mol). Both examples shows how OMe presents less steric constrain than Me. In systems **A** and **H**, the OMe substituent (in R_1) puts its methyl away from the carbonylic oxygen while in **D** and **E** the closest contacts are associated to the hydrogen atoms in the methyl group attached to C_1.

The lower steric constrain of methoxy with respect to methyl is also reflected in the geometrical parameters collected in Table 3.3. The local minima containing methyl in the R_1 position, **D** and **E**, are less planar, as shown by the higher values for the dihedral angle ϕ, 43.2° and 42.7°, respectively. The corresponding values for the species containing methoxy are 39.0° and 38.9°. The deviation

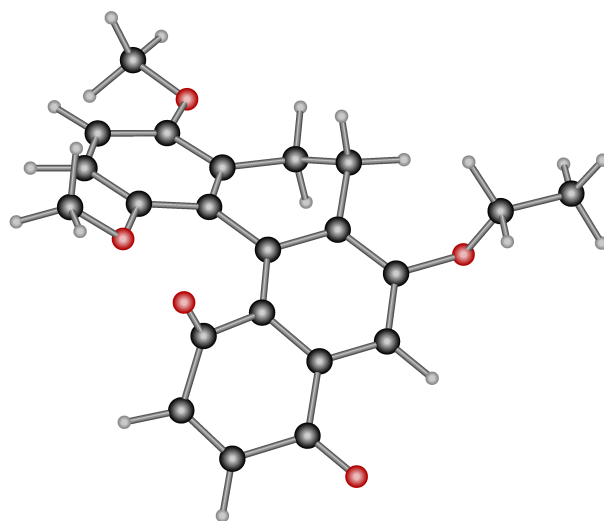## 3.1. Thermal stability of [4]Helicenequinones

Figure 3.5: Becke3LYP optimized structure of the transition state for the [4]helicene system **H**

from planarity of the minima is directly related to the steric bulk of the substituent at R_1. This confirms that, at least for this particular set of molecules, methyl is bulkier than methoxy.

Table 3.3: Computed values for the dihedral angles ϕ (degrees) and free energy barrier (kcal/mol) for [4]helicene systems differing in the nature of the R_1 substituent.

Molecule	R_1	R_2	ϕ_{min}	ϕ_{TS}	$\phi_{TS} - \phi_{min}$	ΔG
A	OMe	H	39.0	35.3	-3.7	22.8
D	Me	H	43.2	30.6	-12.6	25.2
E	Me	OEt	42.7	31.0	-11.7	25.9
H	OMe	OEt	38.9	36.4	-2.5	23.0

The geometry of the transition states follows the opposite trend, they are more planar (smaller ϕ) for the systems with methyl. There seems to be a general trend that the bulkier systems have the higher barriers and the more planar

51

transition states. This is likely related to the fact that in the reaction coordinate both the ϕ torsion and the angles around the sp^3 carbons are involved. And the ϕ torsion is likely to play a larger role for bulkier ligands.

Me *vs.* ^{i}**-Pr**

When the methyl at R_1 is replaced by a larger substituent, like an isopropyl, an increase of the interconversion barrier should be expected due to a larger steric constrain. This can be analyzed comparing the behavior of systems **E** and **G**. Both molecules differ only in the R_1 substituent, which is a methyl in **E** and an isopropyl en **G**. They share the same R_2 and R$_3$ substituents, ethoxy and hydrogen, respectively. The barrier for the **G** system, 29.3 kcal/mol, is indeed higher than that for **E** 25.9 kcal/mol, as expected. However the increase is relatively small, 3.4 kcal/mol. This is in fact quite similar to the values for the methoxy/methyl exchange of 2.9, 2.4 kcal/mol reported above.
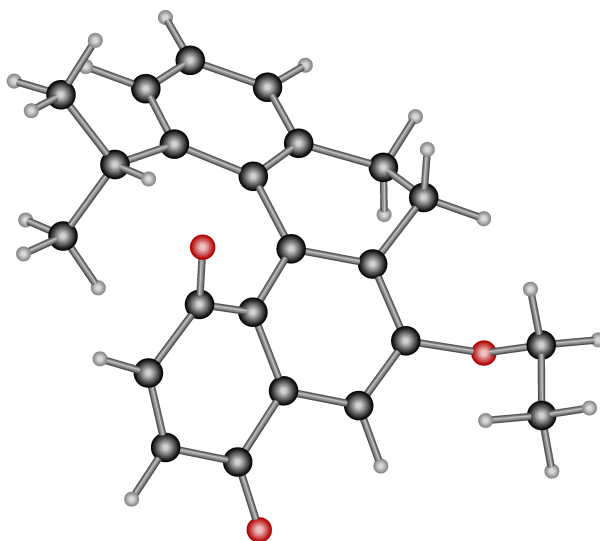

Figure 3.6: Becke3LYP optimized structure of the transition state for the [4]helicene system **G**

3.1. Thermal stability of [4]Helicenequinones

This lower than expected difference in barriers can also be appreciated in the comparison between **D** and **F** which differ also only in the replacement of methyl by isopropyl. In this case, the change only brings a barrier increase of 3.1 kcal/mol. The unexpectedly low effect of the replacement of methyl by isopropyl can be explained by the topology of the substituents. Figure 3.6 presents the optimized transition state for system **G**. The two methyl substituents of the isopropyl are away from the sterically constrained position, which is occupied by a hydrogen atom, and because of this the increase in steric pressure is relatively small.

^{i}-Pr *vs.* **^{t}-Bu**

The replacement of one hydrogen substituent by a methyl in isopropyl, converting it into *tert*-butyl, produces a high increase in the interconversion barrier. The effect of this replacement can be observed by comparing the behaviors of the **F**/**B** and **G**/**C** pairs of molecules. Comparing **G** with **C** the replacement of isopropyl by *tert*-butyl increases the energy barrier to interconversion by 12.7 kcal/mol (from 29.3 to 42.0 kcal/mol). The comparison between systems **F** and **B** yields a similar result, with an increase in barrier of 10.9 kcal/mol.

The reason for this big steric effect is clear from Figure 3.7, where the transition state for system **C** is shown. In the case of *tert*-butyl unlike previous substituents, there is no possible conformation to put all methyl groups away from the more sterically constrained position. As a result, systems **B** and **C** have much higher barriers to interconversion. This fits well with the observation that neither **B** nor **C** present enantiomeric inversion under the experimental conditions.

Table 3.4 summarizes the effect of the R_1 substituent in the barrier for in-
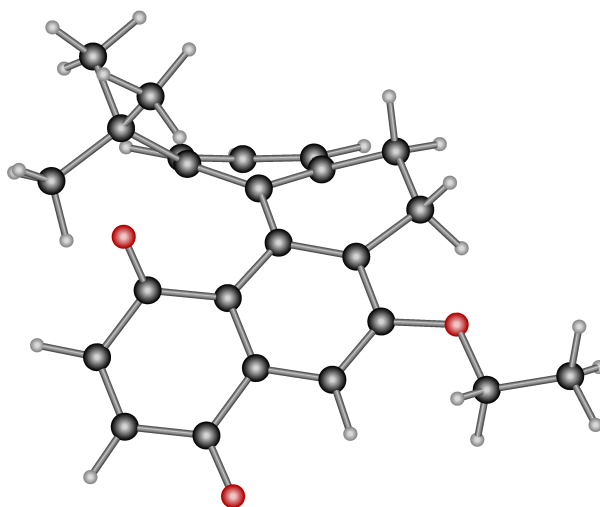
Figure 3.7: Becke3LYP optimized structure of the transition state for the [4]he-licene system **C**

terconversion. The barriers for different pairs of molecules, differing only in the nature of the R_1 substituent are compared. The columns on the left present the results for systems with R_2= OEt, while the columns on the right present the results for systems with R= H. The similarity in the trend for both series is remarkable, and it is summarized in the following equation

$$OMe < Me <^{i\text{-}} Pr <<^{t\text{-}} Bu \qquad (3.2)$$

Systems with the larger steric constrain in the R_1 substituent have the higher barriers. According to these results, methoxy has a smaller steric constraint of all the systems. *Tert*-butyl, the bulkiests substituent, stands out, with a barrier higher than 10 kcal/mol than that of all other systems. Our computational study succeeds in explaining both the placement of methoxy in the series and the much larger effect of *tert*-butyl with respect to the other substituents. In both cases, the key is the eventual availability of conformations where the large groups can be placed away from the sterically sensitive region of the system.

54

3.1. Thermal stability of [4]Helicenequinones

Table 3.4: Relative effect (kcal/mol) of the nature of the R_1 substituent on the barrier for interconversion in [4]helicene systems. The case with R_1= Me is taken as origin of differences.

R_1	Molecule	$\Delta(\Delta G^{\ddagger})$	Molecule	$\Delta(\Delta G^{\ddagger})$
Me	**E**	0	**D**	0
OMe	**H**[a]	-2.9	**A**[a]	-2.4
^{i}-**Pr**	**G**	3.4	**F**	3.1
^{t}-**But**	**C**	16.1	**B**	14.0

[a] Also differ in R_2

3.1.4   Effect of the R_2 substituent

Experimental data show that the replacement of the hydrogen in R_2 by OEt means an increase of the interconversion barrier. The effect is smaller than that of R_1, but it is clearly reproduced in the different pairs of systems. This experimental effect is well reproduced by the calculations. This increase is 0.2 kcal/mol in (**A/H**), 2.8 kcal/mol in (**B/C**), 0.7 kcal/mol in (**D/E**) and 1.0 kcal/mol in (**F/G**).

The geometries for minima and transition states show that the ethoxy substituent has a very similar conformation on both structures. It is not clear wheather the effect of the R_2 is steric or electronic.

This was clarified by a supplementary calculation on a new system **Eb**, where the OEt group in the R_2 position was replaced by OH. The idea was that for this type of molecules, the electronic properties of OH are similar to those of OEt, but its steric properties are closer to those of H. The results were conclusive. The computed inversion barrier for this new molecule, **Eb**, was 25.9 kcal/mol, only 0.02 kcal/mol lower than **E**, but 0.6 kcal/mol higher than **D**. The effect of the R_2 substituent is thus essentially electronic, and the presence of OEt increases slightly the barrier for interconversion, and thus the enantiomeric stability. Ap-

parently, the presence of an oxygen atom attached to the aromatic system produces an electronic effect which increases its resistance to distortion, and thus makes more difficult the interconversion.

As a summary, we can say that our computational study on [4]helicene systems has been able to reproduce the experimental data on chiral stability. This study has been also able to produce simple rules to explain the differential effect of substituents. The barrier to inversion, and thus the chiral stability, is significantly increased by the presence of substituents with large steric bulk in the R_1 position. A minor increase is also obtained through the replacement of hydrogen by ethoxy in the R_2 position, and this second effect is mostly electronic.

3.2   Asymmetric synthesis of [5]helicenequinones

Experimental data show[103] that the configuration, M or P, of the final product in the multistep synthesis of [5]helicenequinones is decided in the step of oxidation from dihydro[5]helicenequinone to [5]helicenequinone. Figure 3.8 shows this oxidation step, which is an aromatization of the B ring (highlighted in bold in the Figure) of the dihydro[5]helicenequinone intermediate. This aromatization involves the transfer of the axial hydrogens of both carbon 5 and 14 (marked by red arrows in the Figure) to the oxidating agent. The full reaction has been explained in section 1.2.2 in the introduction chapter.
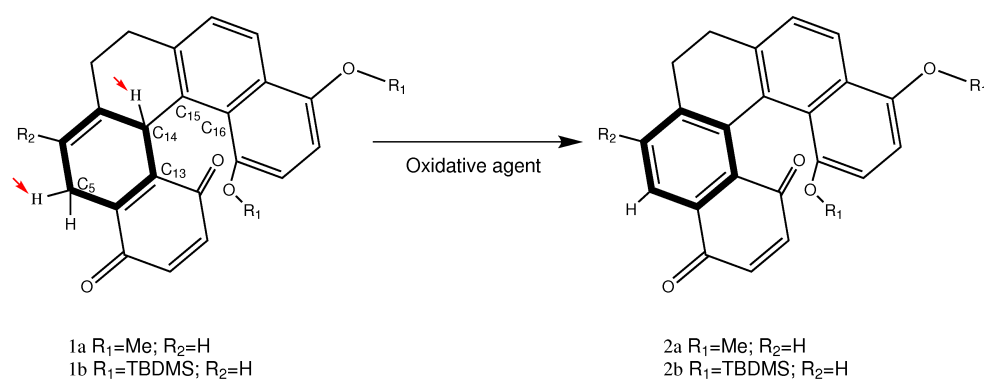

1a R_1=Me; R_2=H
1b R_1=TBDMS; R_2=H

2a R_1=Me; R_2=H
2b R_1=TBDMS; R_2=H

Figure 3.8: Aromatization of the intermediate 5-helicene

The [5]helicenequinone product has two possible enantiomeric forms, M and P, depending on the helical arrangement. The experimental data show that the proportion of the two enantiomeric forms depends on a number of factors, including the nature of the R_1 substituents and of the oxidant used. It has been proposed that the configuration of the product is related to the nature of the conformation of the dihydro[5]helicenequinone intermediate. This intermediate can have in principle two conformers, **I** and **II**, shown in Figure 3.9. **I** would be the precursor of the P form of the product and **II** of the M form. However, this hypothesis has not been proved.
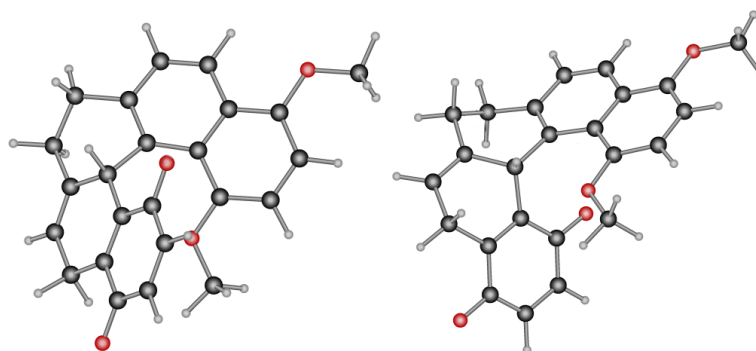
Figure 3.9: Conformers I (left) and II (right) for the R_1 methyl disubstituted 5-helicene intermediate

We carried out a computational study to understand the role of the R_1 substituents (see Figure 3.8) in the outcome of the reaction. Experimental data indicate that the reaction with the 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) oxidant produces sharply different results when R_1 is modified. When both R_1 substituents are methyl the major product is M, with an enantiomeric excess of 44%. When both substituents are *tert*-Butyldimethylsilyl (TBDMS), the main product is P, with an enantiomeric excess of 96%. Both systems were computationally analyzed. Because of the complexity of the system, we studied also two more simple model reactions. The first one was the aromatization of 1,4-cyclohexadiene by p-benzoquinone. The second one was the same experimental dihydro[5]helicenequinone but with p-benzoquinone as oxidant.

3.2.1 Reaction of 1,4-cyclohexadiene with p-benzoquinone

Our first computational study on this oxidation reaction was carried out on the model system 1,4-cyclohexadiene plus quinone. This is the smallest model of the full system still conserving the main electronic features of the B ring. Moreover, similar systems had been previously studied computationally.[181,182] Those

3.2. Asymmetric synthesis of [5]helicenequinones

studies are relevant to the present work because they analyzed the validity of a variety of computational methods, and gave an approach to the reaction mechanism. The methodological calibration was important because of the presence of open-shell species in the mechanism in spite of the closed-shell nature of reactants and products. Chan and Radom[182] found that the BS-DFT approach[183] produced sufficiently accurate results.

Because of this, we carried out full geometry optimizations with the unrestricted Becke3LYP density functional[18,19,177] method as implemented in Gaussian03 package using the 6-31+G(d) basis set. Open-shell states were obtained in the pertinent reaction steps by manipulating the orbital guess mixing the HOMO and LUMO. SCF convergence was very slow in a number of cases and quadratically convergent (QC) SCF procedure[184] was applied using a tighter convergence in solving the linear equation systems.
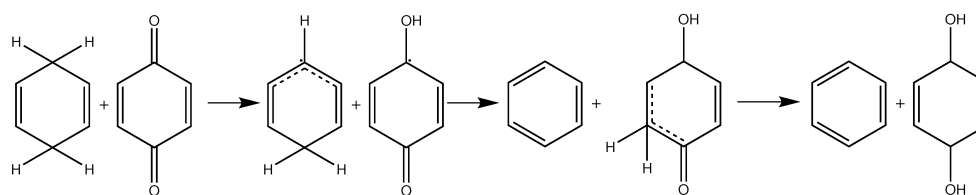
Figure 3.10: Computed mechanism for the reaction between 1,4-cyclohexadiene and quinone

The previous study mentioned above[182] was systematic from a methodological point of view, but fell short of considering the full reaction, as it focused only on the initial hydrogen transfer step. We repeated these previous calculations with our computational method and calculated the full reaction. The computed mechanism is shown in Figure 3.10. The two hydrogens from 1,4-cyclohexadiene are transferred sequentially to the quinone. In the first step, a hydrogen atom is transferred from the cyclohexadiene molecule to one of the

oxygens of the quinone. The resulting intermediate consists of two loosely bound open-shell doublets, with the unpaired electrons mainly placed on the carbon centers involved in the hydrogen transfer. In the following step, the second hydrogen is transferred to the quinone. However, it does not go directly to the second oxygen of the quinone, but to one of the alkenic carbons. The benzene product is then already formed. The other molecule reorganizes through a low barrier keto-enol isomerization to the final 1,4-dihydroquinone product.
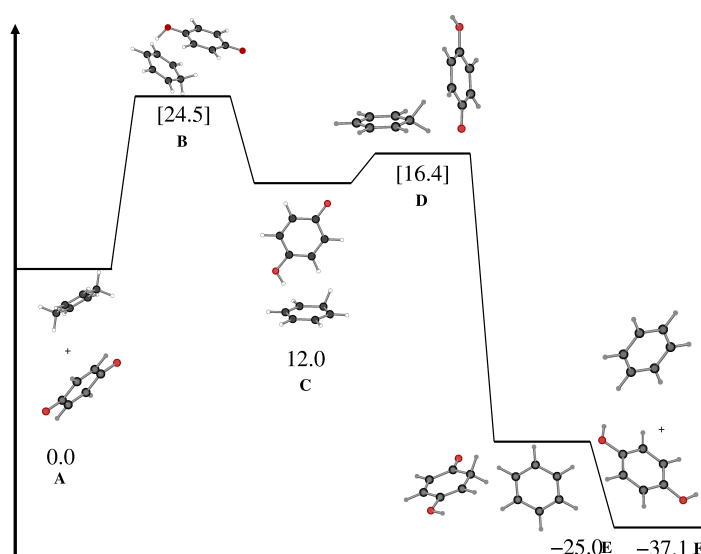
Figure 3.11: Computed energy profile for the reaction between 1,4-cyclohexadiene and quinone. Energies in kcal/mol

Computed energies and structures for the reaction are summarized in Figure 3.11. The overall reaction is quite exothermic. Final products **F** are -37.1 kcal/mol below reactants **A**. There are two main barriers, associated to transition states **B** and **D** (Figure 3.12), corresponding to both hydrogen transfer steps. The first transfer, from **A** to **C** is endothermic, the intermediate is 12.0 kcal/mol above the reactants. In transition state **B**, 24.5 kcal/mol above **A**, the rings of the two reacting fragments are in nearly parallel planes. However, this changes

3.2. Asymmetric synthesis of [5]helicenequinones

in intermediate **C**, where the rings are alsmot perpendicular. The second hydro-
gen transfer, involving transformation from **C** to **E** is clearly exothermic, **E** is 38
kcal/mol below **C**. **E** evolves to the final products through a simple keto-enol
isomerization, which takes place because the enol form of 1,4-dihydroquinone
is 12.1 kcal/mol more stable than the ketonic form. Most of the species involved
in the mechanism are closed-shell singlets, with the exception of intermediate **C**
and transition states **B** and **D**, which contain open-shells. They were computed
through the broken-symmetry DFT approach, as mentioned above.
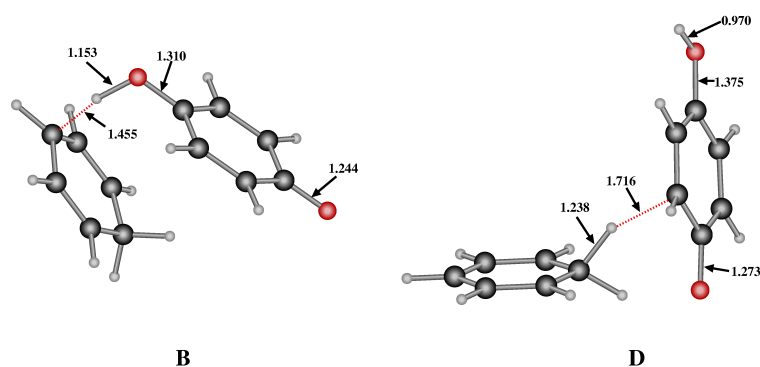
Figure 3.12: Structure of transition states **B** and **D** for the hydrogen transfer
steps of the reaction between 1,4-cyclohexadiene and *p*-quinone. Distances in Å

As expected, the global reaction is quite exothermic, by nearly 40 kcal/mol,
it is a clearly favorable oxidation. The barrier, over 20 kcal/mol, is however
moderately high. There are two critical transition states, corresponding to both
hydrogen transfer steps. The first transition state, **B**, has the highest barrier,
24.5 kcal/mol with respect to the reactants, and corresponds for this system
to the rate-determining step. The second transition state, **D**, has an energy 8.1
kcal/mol below **B**. However, we will consider both of them when considering
the reaction between the real [5]-helicene and quinone systems.

61

3.2.2 Computational model for the study of the real systems

The system is quite large, and because of this a QM/MM approach was used. The calculations were done using the hybrid QM/MM ONIOM method implemented in Gaussian03 package. The QM layer was computed using the unrestricted Becke3LYP density functional[18,19,177] and the 6-31+G(d) basis set. The broken-symmetry approach was used in the same steps where found necessary in the model system, and the quadratically convergent (QC) SCF procedure[184] was applied using a tighter convergence in solving the linear equation systems. The MM layer was described using the UFF force field.[185] The QM/MM partition is shown in Figure 3.13. The QM layer involved the **B** ring of the studied [5]helicene and the full oxidant agent. The rest of the helicene constituted the MM layer.
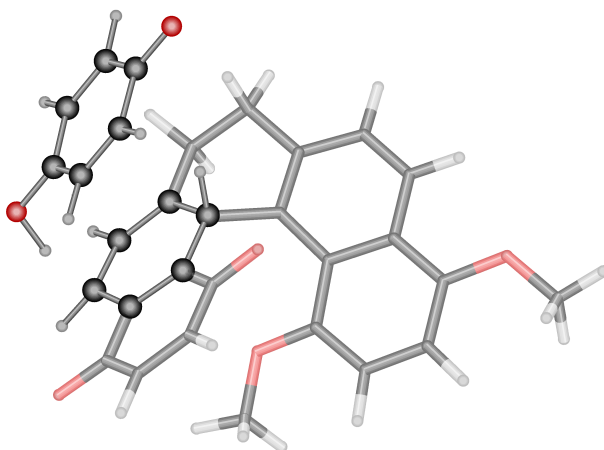

Figure 3.13: QM/MM partition used in the ONIOM calculation of the oxidation of dihydro[5]helicenes. The tube representation shows the MM part and the ball-and-stick represents the QM part.

Two different helicene systems were considered. These contained in the R_1 position 3.8 either methyl or TBDMS. Two different oxidants were considered,

3.2. Asymmetric synthesis of [5]helicenequinones

the simplest quinone model and the experimentally applied derivative DDQ. The quinone model was applied to simplify the analysis. For each of the four systems computed, the two conformers, **I** or **II**, shown in Figure 3.9 had to be computed. Minima and transition states for each case were computed in order to compare the barriers leading to the P or M products. From the computed barriers, the enantiomeric excess can be computed and compared to the experimental values, and the origin of enantioselectivity can be analyzed.

3.2.3 Reaction of dihydro[5]helicene systems with methyl substituents

Conformers of the reactant
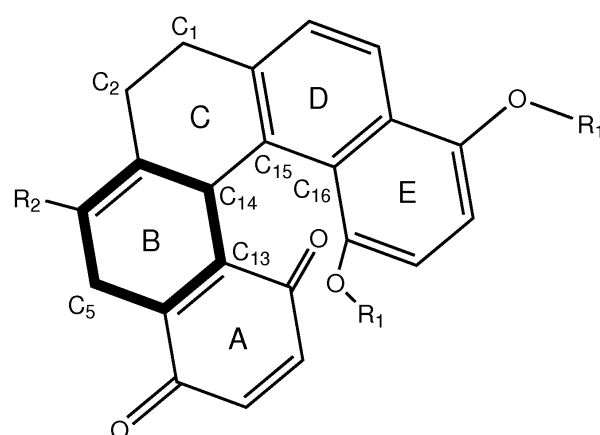
1a R_1=Me; R_2=H
1b R_1=TBDMS; R_2=H

Figure 3.14: Labeling of the of the [5]helicene systems

First we analyzed the structure and relative stability of the two conformers of the dihydro[5]helicene starting compound. Figures 3.15 and 3.16 present the

63

optimized structure of both conformers, **I** and **II**. They have the same configuration in the asymmetric carbon C_{14}, but they differ in the arrangement of the ring **C**. Ring **C** has two possible conformations, that are clearly defined by the dihedral angle ϕ (C_{13}-C_{14}-C_{15}-C_{16}, defined in Figure 3.14). This angle collected in Table 3.5 is positive (7.0°) for conformer **I**, but negative (-77.5°) for conformer **II**.

Figure 3.15: ONIOM(Becke3LYP:UFF) optimized structure for conformer **I** of the methyl disubstituted dihydro[5]helicene

Table 3.5: Computed parameters (degrees) and relative energies (kcal/mol) for the two conformers of dehydro[5]helicene with methyl substituents.

Helicene	ϕ_{C1}	ϕ_{C2}	ϕ	Total energy	QM energy	MM energy
I	111.1	112.4	7.0	7.6	-7.0	14.6
II	109.8	108.3	-77.5	0.0	0.0	0.0

3.2. Asymmetric synthesis of [5]helicenequinones
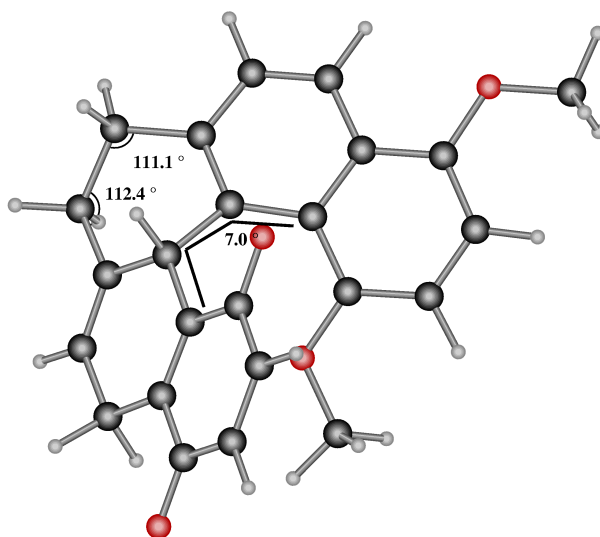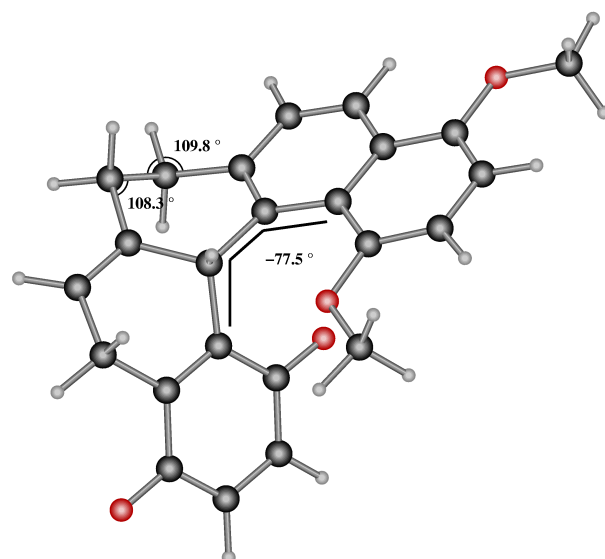

Figure 3.16: ONIOM(Becke3LYP:UFF) optimized structure for conformer **II** of the methyl disubstituted dihydro[5]helicene

Conformer **II**, which is expected to lead to the M product, is more stable, by 7.6 kcal/mol than conformer **I**, as shown in Table 3.5. The partition between the QM and MM parts of the energy difference is very informative. The stability of the **II** conformer comes exclusively from the MM part, because the QM part alone would in fact favor the **I** form by 7.0 kcal/mol.

There are two possible sources of steric strain in this kind of molecules. One is the repulsion between the carbonyl oxygen of the A ring and the R_1 substituent in the E ring, in this case, a methyl. The other strain is associated to the presence of some carbons with sp^3 hybridization in the cyclic skeleton, C_1, C_2, C_5 and C_{14}. C_1 and C_2 are especially affected by the conformation of the helicene.

Inspection of Figures 3.15 and 3.16, and of Table 3.5 clarifies the origin of the energetic difference between both conformers. The steric effect of the R_1 substituents seems minor at most for these systems. Differences in steric strain

65

on the sp^3 carbons C_1 and C_2 are however clear. The ideal relaxed tetrahedral angles should be close 109.5°. The values for conformer **II** are reasonable close, 109.8° and 108.3°. In contrast, the values for conformer **I** are 111.1° and 112.4°, quite far from the ideal values. The origin of this distortion can be traced down to the rotation around the C_{14}-C_{15} bond, which rules the planarity of the whole system. The different configuration at the sp^3 carbon C_{14} forces the system to be much more planar (ϕ close to 0) in **I**, and this leads to higher steric strain overall in the system. The QM energy is in contrast more favorable to the **I** isomer. In this conformer, the B ring has a more similar arrangement to that of unconstrained 1,4-cyclohexadiene. However this is not sufficient to overcome the large steric advantage of conformer **II**.

Reaction with *p*-quinone

We first analyzed the oxidation of the dihydro[5]helicene by the simple oxidant *p*-quinone. Although this simple oxidant had not been experimentally used, this study will allow us to better understand the mechanistic complexity of the process and to focus on the key step when the real oxidant is considered.

As seen in the preliminary study with 1,4-cyclohexadiene, the aromatization reaction consists of two sequential hydrogen transfer steps. However, the complex topology of dihydro[5]helicene increases largely the number of possible reaction paths for both hydrogen transfers. The different possibilities are schematically shown in Figure 3.17. For the first hydrogen transfer, the complication emerges from the fact that the two hydrogens to be transferred from the helicene are not equivalent. The first hydrogen transferred can come from either C_5 or C_{14} and this creates two manifolds of reaction paths, which have been labeled as internal and external. For the second transfer, the complication

66

3.2. Asymmetric synthesis of [5]helicenequinones

appears in the quinone part. Calculations on 1,4-cyclohexadiene showed that the hydrogen was transferred to a carbon *meta* to the carbonyl group initially hydrogenated. For the real system, there are two non-equivalent carbons that can receive this transfer, defining paths that we have labeled as **A** and **B**. We also considered the possibility that the second hydrogen transfer goes directly to the oxygen center (path **C**).
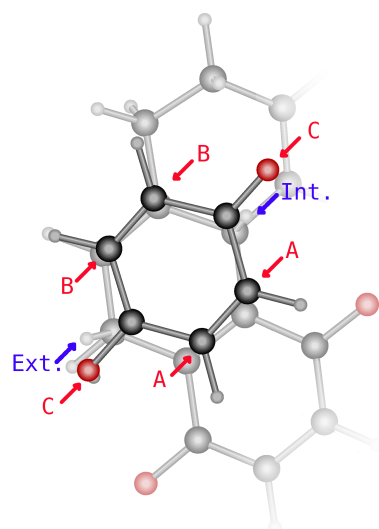
Figure 3.17: Different possible paths for the reaction between dihydro[5]helicene and *p*-quinone. In blue, the two possible origins (internal or external) of the first hydrogen transfer. In red the three possible destinations (A, B or C) for the second hydrogen transfer.

The combination of two possible conformers, **I** and **II**; two possible origins, internal or external, for the first hydrogen transfer; and three possible destinations, **A**, **B** and **C** for the second hydrogen transfer, gives rise to twelve possible reaction paths. The two transition states for each of them have been computed and the resulting energy barriers are collected in Table 3.6. All values are relative to that of the separate reactants, with the dihydro[5]helicene in the most

67

stable conformation **II**.

Table 3.6: Relative energies (kcal/mol) for each transition state in the reaction between the methyl-disubstituted dihydro[5]helicene and p-quinone.

Conformer	1st TS	1st TS energy	2nd TS	2nd TS energy
I	internal	27.3	A	19.1
I	internal	27.3	B	18.9
I	internal	27.3	C	15.4
I	external	27.3	A	24.4
I	external	27.3	B	22.8
I	**external**	**27.3**	**C**	**5.9**
II	internal	29.6	A	17.5
II	internal	29.6	B	17.2
II	internal	29.6	C	13.8
II	external	24.9	A	22.6
II	external	24.9	B	21.3
II	**external**	**24.9**	C	**10.7**

In all cases, the first transition state has a higher barrier, and because of this, we predict the first hydrogen transfer to be the rate-limiting step. The lowest barrier for this first step corresponds to the external arrangement of conformer **II** with a barrier of 24.9 kcal/mol. Therefore, the major product should be *M*. The lowest barrier leading to the *P* product, thus coming from **I** conformers could be from either the internal of external rearrangement, which are practically degenerate, with a barrier of 27.3 kcal/mol. If we accept Curtin-Hammett equilibrium conditions[186] between both starting conformers, this would lead to a significant enantiomeric excess of 97.63% at a temperature of 273 K. It is worth

3.2. Asymmetric synthesis of [5]helicenequinones

noticing that arrangement **C**, direct transfer to oxygen, is favored in all cases for the transition state of the second step. This is different from what happened in the 1,4-cyclohexadiene model system. It is likely due to the larger steric effects in the real system, but we did not study it further because it has no relevance on the final outcome.

We will briefly analyze in what follows the structure of the most favored rate-determining transition states leading to the M and P products. They are shown in Figures 3.18 and 3.19, respectively. Key geometrical parameters as well as the relative energy decomposition in QM and MM contributions are shown in Table 3.6. For the reaction leading to the P product, conformer **I**, where there are two practically degenerate transition states, we have chosen the structure with transfer of hydrogen in the external position. This is more similar to that leading to the M product, and this facilitates the comparison.

Both transition states are clearly associated with hydrogen transfer, as proved by the C-H and O-H distances involved in the atom being transferred. For the transition state leading to the M product, the C-H and O-H values are of 1.429 and 1.175 Å, respectively. The corresponding values for the structure leading to the R product are 1.450 and 1.166 Å. As mentioned above, the transition state leading to the M product is 2.4 kcal/mol more stable than that leading to the P product. The decomposition of the energy difference in QM and MM contributions shows that the difference is concentrated in the MM contribution (12.2 kcal/mol), with the QM contribution actually favoring the P product by 9.8 kcal/mol. This result reproduces that reported above for the two conformers of the dihydro[5]helicene reactant, and the explanation is the same. The transition state connected to conformer **I** is more planar ($\phi = 8.1°$), which leads to more strain on C_1 and C_2, which are in the center of angles (110.9°, 112.3° somehow far from the ideal tetrahedral value. The transition state connected to

69

Figure 3.18: ONIOM(Becke3LYP:UFF) optimized structure of the rate-determining transition state for the path leading to the M product, related to conformer **II**, in the reaction between p-quinone and the dihydro[5]helicene with two methyl substituents.

conformer **II**, and leading to the M product is more stable because although its low planarity (ϕ= -75.6°) leads to a higher QM energy, it has smaller strain on C_1, C_2 (bond angles of 109.7°, 107.9°) which leads to a much lower MM energy, and to a lower total energy overall.

It is however worth remarking that the energy difference between confor-mations **I** and **II** is smaller in the transition state (2.4 kcal/mol) than it was in the reactants (7.6 kcal/mol). This is in fact reflected in the less distorted angles around C_1 and C_2, and can be correlated with the fact that removal of one hy-drogen of dihydro[5]helicene starts to relaxes the strain associated to planarity in the ring adjacent to that being dehydrogenated.

70

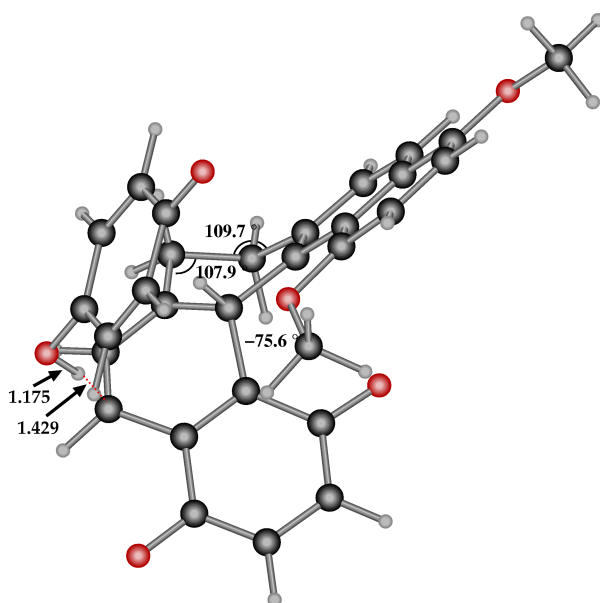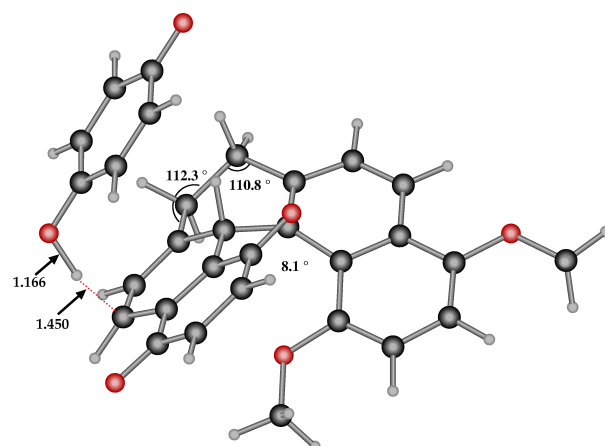## 3.2. Asymmetric synthesis of [5]helicenequinones

Figure 3.19: ONIOM(Becke3LYP:UFF) optimized structure of the rate-determining transition state for the path leading to the *P* product, related to conformer **I**, in the reaction between *p*-quinone and the dihydro[5]helicene with two methyl substituents.

It is clear in any case that our computational study predicts the *M* enantiomer to be the major product in the reaction between *p*-quinone and the dihydro[5]helicene system with two methyl substituents. The computed enantiomeric excess would be 97.6% at 273K. The origin of the enantiomeric excess is in the different stabilities of the conformers of the reactant leading to each of the enantiomeric forms of the product.

Reaction with DDQ

After studying the reaction of the methyl disubstituted dihydro[5]helicene with *p*-quinone we moved to study the reaction of the same substrate with 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ). This reaction has been experimentally carried out, and found to produce the *M* product with an enantiomeric excess of 44%. The use of DDQ introduces an additional source of complexity

71

Table 3.7: Computed parameters (degrees) and energy barriers (kcal/mol) for the key transition states in the reaction between p-quinone and dihydro[5]helicene with methyl substituents.

Path	ϕ_{C1}	ϕ_{C2}	ϕ	Total energy	QM energy	MM energy
I	110.9	112.3	8.1	27.3	17.4	9.8
II	109.7	107.9	-75.6	24.9	27.3	-2.4
II-I	–	–	–	-2.4	9.8	-12.2

in the system, because this oxidant is less symmetric than p-quinone. There are two possible arrangements depending on the orientation of the chloride substituents with respect to the helicene. Both of them are shown in Figure 3.20. We have labeled the approach where the chlorine groups are oriented towards the helicene ring A as *cisoid* (Figure 3.20, left) and the opposite one as *transoid* (Figure 3.20, right).
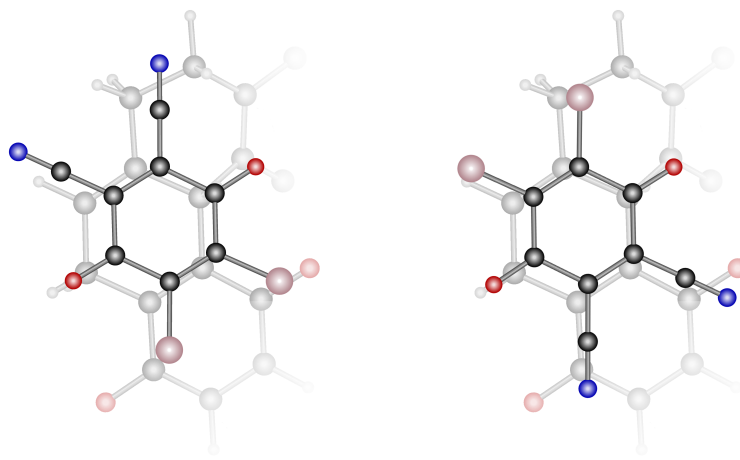

Figure 3.20: The *cisoid* (left) and *transoid* (right) approaches in the reaction of DDQ with dihydro[5]helicene.

We computed four transition states, corresponding to the consideration of both the *cisoid* and *transoid* arrangements for the analogous structures to the

3.2. Asymmetric synthesis of [5]helicenequinones

two key transition states for the reaction with *p*-quinone. The results are summarized in Table 3.8, and the two key structures leading to the *M* and *P* products are presented in Figures 3.21 and 3.22, respectively.
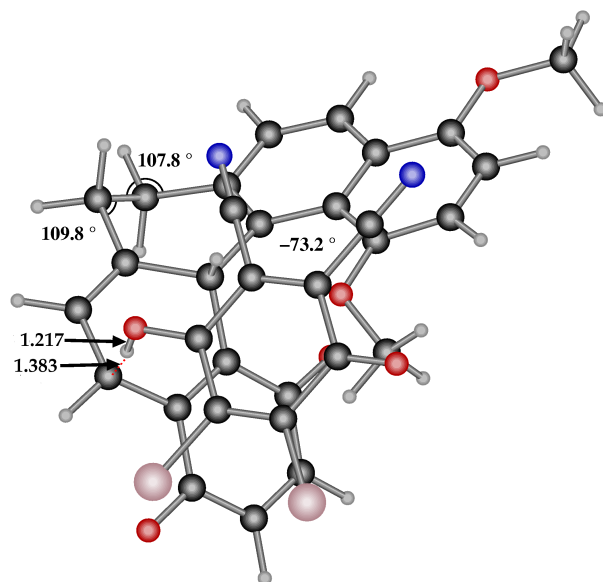

Figure 3.21: ONIOM(Becke3LYP:UFF) optimized structure of the rate-determining transition state for the *cisoid* path leading to the *M* product, related to conformer **II**, in the reaction between DDQ and the dihydro[5]helicene with two methyl substituents.

The energetics summarized in Table 3.8 indicate that the barrier for the reaction with DDQ is lower than it was for the reaction with *p*-quinone. The values for DDQ are in the 14-18 kcal/mol range, while the values for *p*-quinone were in the 24-28 kcal/mol range. This result is not surprising, DDQ is known to be a better oxidant. As far as selectivity is concerned, the pattern is quite similar to that observed for *p*-quinone, and that the *cisoid-transoid* dichotomy brings only minor corrections. Both transition states derived from conformer **II**, leading to the experimentally observed major product *M* have the lowest energies, separated by only 0.4 kcal/mol. The *cisoid* approach is slightly favored. The two
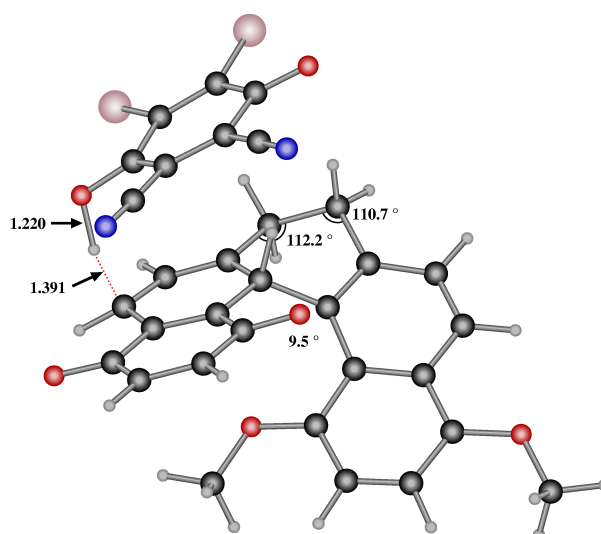
73

Figure 3.22: ONIOM(Becke3LYP:UFF) optimized structure of the rate-determining transition state for the *transoid* path leading to the *P* product, related to conformer **I**, in the reaction between DDQ and the dihydro[5]helicene with two methyl substituents.

transition states derived from conformer **I**, leading to the minor product *P* have energies around 4 kcal/mol (3.8 and 4.3 kcal/mol) above the lowest transtion state. The only minor novelty is that for the structures leading to the *P* product the *transoid* approach is favored by 0.5 kcal/mol. The decomposition of the total energy in QM and MM contributions shows the pattern already discussed above. The QM term would favor strongly the *P* product, but it is overridden by an even larger different in the MM term favoring the *M* product.

The MM stabilization is related to the planarity of the more saturated ring of the dihydro[5]helicene. In the type **I** transition states, the ϕ torsional angles are 8.8° and 9.5°. In the type **II** transition states, the corresponding values are -73.2° and -74.0°. The planarity has a translation in the bond angles centered in the

3.2. Asymmetric synthesis of [5]helicenequinones

Table 3.8: Computed parameters (degrees) and energy barriers (kcal/mol) for the key transition states in the reaction between DDQ and dehydro[5]helicene with methyl substituents.

Path	ϕ_{C1}	ϕ_{C2}	ϕ	Total energy	QM energy	MM energy
I-*cisoid*	110.7	112.2	8.8	18.3	8.8	9.5
I-*transoid*	110.7	112.2	9.5	17.8	8.8	9.0
II-*transoid*	109.7	107.8	-74.0	14.4	17.7	-3.3
II-*cisoid*	109.8	107.8	-73.2	14.0	17.2	-3.1

C_1 and C_2 carbons, more strained in the type **I** transition states leading to the *P* product. The difference in MM energy is similar to that previously computed for the *p*-quinone systems, with values around 12 kcal/mol.

The QM part of the energy, is more favorable to the conformer **I**, as it was also the case for the reaction with *p*-quinone. However, he difference is now reduced to values below 9 kcal/mol, while the value for the *p*-quinone system was 9.8 kcal/mol. This indicates that there are some effects associated to the presence of the substituents in DDQ that favor the reaction of conformer **II**.

The net result is that in the case of the reaction between the dihydro[5]helicene system with DDQ the energy difference between the two most stable transition states leading respectively to the *M* and *P* products is 3.8 kcal/mol. This would correspond to an enantiomeric excess of 99.8% at 0K. The sense of selectivity is correct, with the *M* enantiomer being the major product, but the computed enantiomeric excess is much larger than the experimental value of 44%. There are a number of possible reasons that can explain the discrepancy. On one hand, the computational method is improvable: the description of open-shell singlets with BS-DFT approaches is not perfect, the UFF force field is not the most accurate one, no solvation effects or entropic corrections have been introduced.

There are also more basic possible sources of error in the form of questionable assumptions, like the equilibrium between the different adducts of the reactants necessary for the Cutin-Hammett approach to be applicable, or the possible role of the second hydrogen transfer in the cases where we have not computed it. In any case, we consider the results encouraging, because they furthermore provide a simple qualitative explanation based on the relative stability of the different conformers of the reactant, which seems very reasonable.

3.2.4 Reaction of dihydro[5]helicene systems with TBDMS substituents

The second dehydro[5]helicene system was that with two *tert*-butyldimethylsilyl (TBDMS) in the R_1 positions. The reaction of this compound with DDQ has been shown to produce the *P* product with an enantiomeric excess of 96%. The study will follow the same pattern of that presented above for the system with dimethyl substituents: discussion of the conformers of the reactant, reaction with *p*-quinone and reaction with DDQ.

Conformers of the reactant

The two conformers of the dihydro[5]helicene system were computed with the ONIOM(Becke3LYP:UFF) method, and the results are summarized in Figures 3.23 and 3.24, as well as in Table 3.9.

The structures and stabilities are quite similar to those previously observed with the methyl substituents. Conformer **II** is again the most stable one, now 5.5 kcal/mol below conformer **I**. This value is somehow smaller than the 7.6 kcal/mol reported above for the methyl systems, but it still indicates a strong preference.

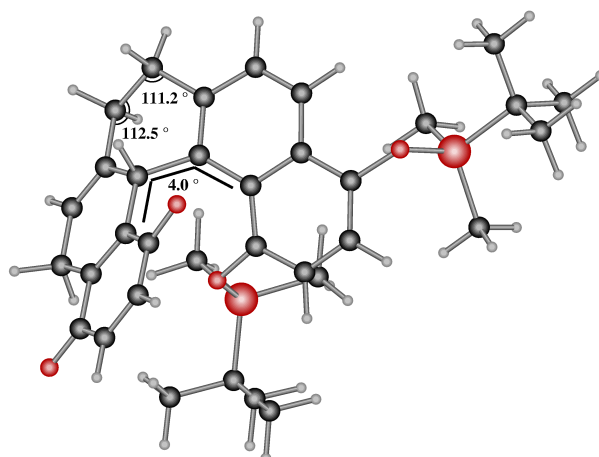## 3.2. Asymmetric synthesis of [5]helicenequinones

Figure 3.23: ONIOM(Becke3LYP:UFF) optimized structure for conformer **I** of the TBDMS disubstituted dihydro[5]helicene

Table 3.9: Computed parameters (degrees) and relative energies (kcal/mol) for the two conformers of dehydro[5]helicene with TBDMS substituents.

Helicene	ϕ_{C1}	ϕ_{C2}	ϕ	Total energy	QM energy	MM energy
I	111.2	112.5	4.0	5.5	-6.1	11.6
II	110.14	108.3	-71.6	0.0	0.0	0.0

The general structure for both conformers, shown in Figures 3.23 and 3.24 is similar to that with methyl substituents. In fact, the bulky TBDMS substituents are able to place themselves away from the core of the molecule, and this explains the similarities between both systems. Conformer **I** has a more planar structure (ϕ is 4.0°), than conformer **II** (ϕ is -71.6°). Compared with the methyl systems, the new larger TBDMS substituent in the R_1 position leads to slightly more planar conformers. This difference of planarity between **I** and **II** affects the angles around the C_1 and C_2 sp^3 carbons in the same way that for the methylated system. These centers are more constrained in **II**, which because of this as
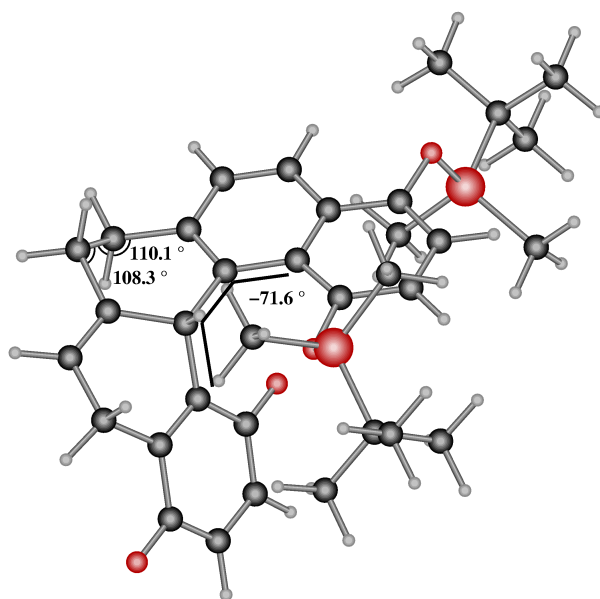
77

Figure 3.24: ONIOM(Becke3LYP:UFF) optimized structure for conformer **II** of the TBDMS disubstituted dihydro[5]helicene

a much higher MM energy, and an overall higher total energy.

Reaction with *p*-quinone and DDQ

Transition states for the oxidation by *p*-quinone and DDQ of the TBDMS disubstituted dihydro[5]helicene were computed. Following the results reported above for the dimethyl system, only the transition state for the first hydrogen transfer was computed. Moreover, only the conformational approaches that were found more productive have been evaluated. This means that two transition states, **I** and **II**, were evaluated for the reaction with *p*-quinone; and four, **I**-*cisoid*, **I**-*transoid*, **II**-*cisoid*, and **II**-*transoid*, were evaluated for the reaction with DDQ. The results were quite similar for both oxidants, and because of this they will be discussed as a whole in this subsection.

78

3.2. Asymmetric synthesis of [5]helicenequinones

Figure 3.25: ONIOM(Becke3LYP:UFF) optimized structure of the rate-determining transition state for the path leading to the *P* product, related to conformer **I**, in the reaction between *p*-quinone and the dihydro[5]helicene system with two TBDMS substituents.

Table 3.10 summarizes the results for the reaction between *p*-quinone and the dihydro[5]helicene system with TBDMS substituents. The overall barriers with respect to the reactants are in the same range (around 25 kcal/mol) to that computed above for the system with methyl substituents. In contrast, the selectivity trend is just the opposite. For the methyl system the barrier was 2.4 kcal/mol lower for conformer **II**, leading to the *M* product. For the TBDMS system the barrier is 3.7 kcal/mol lower for conformer **I**, leading to the *P* product. This result goes furthermore in the sense of reproducing the experimental trend in the oxidation with the DDQ.

Figures 3.25 and 3.26 represent the ONIOM(Becke3LYP:UFF) optimized geometries for the transition states associated to conformers **I** and **II**, respectively. The highlighted geometrical parameters follow a similar pattern to that of the
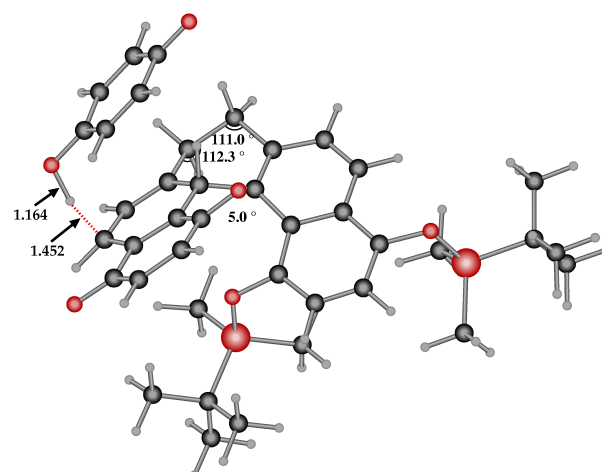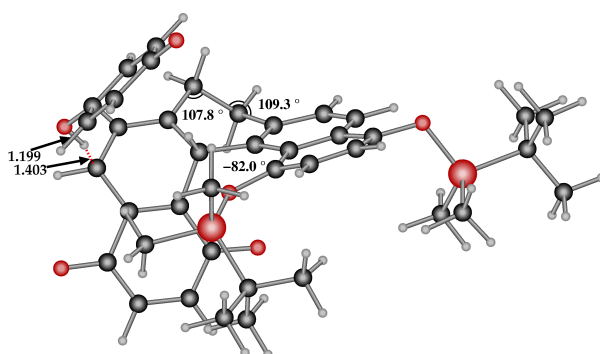
Figure 3.26: ONIOM(Becke3LYP:UFF) optimized structure of the rate-determining transition state for the path leading to the M product, related to conformer **II**, in the reaction between p-quinone and the dihydro[5]helicene system with two TBDMS substituents.

Table 3.10: Computed parameters (degrees) and energy barriers (kcal/mol) for the key transition states in the reaction between p-quinone and dehydro[5]helicene with TBDMS substituents.

Path	ϕ_{C1}	ϕ_{C2}	ϕ	Total energy	QM energy	MM energy
I	111.1	112.3	4.9	24.8	18.5	6.4
II	109.3	107.8	-82.0	28.6	29.6	-1.0
II-I	–	–	–	**3.7**	**11.1**	**-7.4**

systems with methyl substituents. **I** is more planar than **II**, ϕ is 4.9° for **I** and -82.0° for **II**, and this leads to more strain in the carbon centers C_1, C_2. The big difference between both systems is however in the relative position of the TBDMS groups with respect to p-quinone. In the structure related to conformer **I**, both TBDMS substituents are away from the entering oxidant, and are only marginally affected by the overall molecular distortion from the reactants. On the contrary, in the structure related to conformer **II** one of the TBDMS groups points exactly towards the region where the p-quinone has to approach the di-

3.2. Asymmetric synthesis of [5]helicenequinones

hydrohelicene. The molecule is quite flexible, and can keep the two groups relatively far, but this crowding is sufficient to increase significantly the energy of the associated transition state, and invert the selectivity of the reaction.

The difference between the methyl and TBDMS systems cannot be tracked down to the partition between QM and MM contributions to the total energy because it has significant contributions from both terms. The QM term is affected because of geometrical changes in the arrangement of the quinone with respect to the helicene to relax the steric strain. The MM term is affected by the direct steric repulsion term between the TBDMS substituent and the quinone.

Table 3.11: Computed parameters (degrees) and energy barriers (kcal/mol) for the key transition states in the reaction between DDQ and dehydro[5]helicene with TBDMS substituents.

Molecule	ϕ_{C1}	ϕ_{C2}	ϕ	Total energy	QM energy	MM energy
II-*transoid*	109.6	107.7	-80.1	19.5	20.2	0.7
II-*cisoid*	109.8	107.8	-79.7	19.1	20.2	-1.0
I-*cisoid*	110.9	112.2	5.6	14.9	9.8	5.0
I-*transoid*	110.8	112.2	6.4	14.5	9.9	4.6

Table 3.11 summarizes the results for the reaction betweeen the dihydro[5]helicene with TBDMS substituents and DDQ. This system has been experimentally studied, with an enantiomeric excess of 96% in favor of the *P* product. The barrier with respect to the reactants is lower than with the *p*-quinone oxidant by *ca* 10 kcal/mol, and similar to that for the reaction of the dihydrohelicene with methyl substituents and DDQ. Again, the *cisoid* and *transoid* approaches produce similar barriers, although it is not obvious a priori which one will have

the lowest energy. The most significant and important result is that the *P* product is preferred, contrary to what happened for the dihydro[5]helicene with methyl substituents, and in agreement with experiment. The most stable transition state is **I**-*transoid*, leading to the *P* product, and the lowest energy path going to the *M* product goes through **II**-*cisoid*, 4.6 kcal/mol above.

The structures for the transition states of the reaction with DDQ are quite similar to those for the reaction with *p*-quinone. The two most stable ones leading to each enantiomer, **I**-*transoid* and **II**-*cisoid* are shown in Figures 3.27 and 3.28, respectively. The role of the bulky TBDMS is the same as for the *p*-quinone system, it increases the energy of structures associated to **II** because of the repulsion between one of the substituents and the entering oxidant.
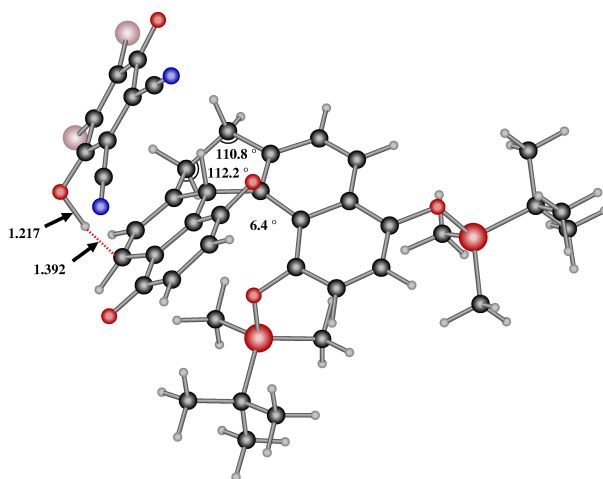

Figure 3.27: ONIOM(Becke3LYP:UFF) optimized structure of the rate-determining transition state for the *transoid* path leading to the *P* product, related to conformer **I**, in the reaction between DDQ and the dihydro[5]helicene system with two TBDMS substituents.

The key result is that the reaction leading to the *P* product is favored with respect to that leading to the *M* product for the reaction of the helicene with

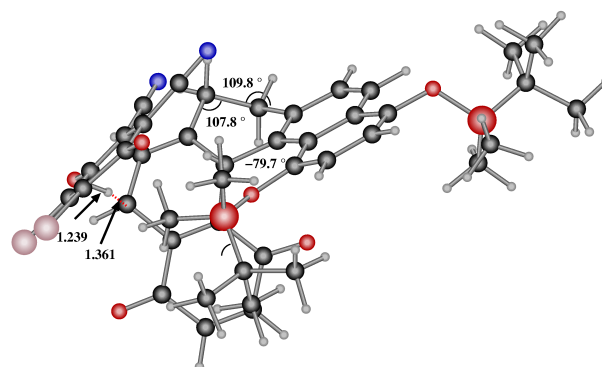## 3.2. Asymmetric synthesis of [5]helicenequinones

Figure 3.28: ONIOM(Becke3LYP:UFF) optimized structure of the rate-determining transition state for the *cisoid* path leading to the *M* product, related to conformer **II**, in the reaction between DDQ and the dihydro[5]helicene system with two TBDMS substituents.

TBDMS substituents and DDQ. The computed difference in energy barriers of 4.6 kcal/mol would lead to an enantiomeric excess of 99.9% at 298.15 K in favor of the *P* product. This value compares quite favorably with the experimentally reported value fo 96%.

Taken as a whole, our calculations on the asymmetric synthesis of [5]helicenequinones succeed in reproducing the dependence of the identity of the major product on the nature of the substituents at the R_1 position, and provide a simple qualitative explanation. The *M* product is favored when R_1 is methyl, and the *P* product is favored when R_1 is TBDMS. The selectivity can be explained by the direct connection between the nature of the reacting conformer and the preferred product.

Conformer **II** of the reactant is always more stable because it has less constrains in ring **A**. When the R_1 positions are occupied by methyl groups, this energy preference is kept in the rate-determining transition state for the transfer of the first hydrogen atom, and the resulting product is *M*. When the R_1

positions are occupied by bulky TBDMS groups the relative stabilities of the key transition states are inverted, because in conformer **II** the bulky substituent interferes with the entering oxidant. As a result, the transition state associated to conformer **I** has a lower energy, and the *P* product is preferred.

Chapter 4

Guanidine derivatives as an organocatalysts

A new method for the synthesis of quaternary aminoacids has been recently proposed based on the enantioselective ring opening of an azlactone catalyzed by a bicyclic guanidine derivative. The selectivity-determining step of this process, which is described in section 1.3, is shown in Figure 4.1. In this step, one of the nitrogen centers of guanidine abstracts a proton from a C-H bond in the azlactone ring, and an adduct is formed between the cationic guanidinium and the anionic derivative of azlactone. This adduct suffers then the attack of methyl acrylate, and the configuration of the final product is defined at this point.

The attack by methyl acrylate will likely take place from the least hindered face of the sp^2 carbon center in the adduct, and because of this, the structure of this adduct is critical. The structure of this critical substrate-catalyst complex was analyzed computationally. There are a priori two possible arrangements between the catalyst and the substrate, and they are shown in Figure 4.2. In the first arrangement, labeled as **A**, the protonated guanidine is complexed to both oxygens of the azlactone, the terminal one and the one in the ring. In this
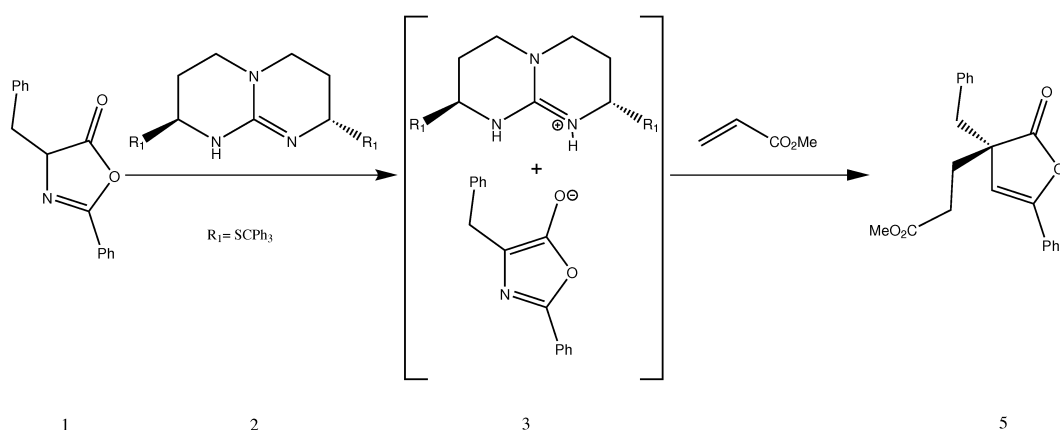
Figure 4.1: Reaction between azlactone and methyl acrylate catalyzed by bicyclic guanidine.

classical arrangement, the rings in the reactant and catalyst are approximately coplanar.

In the alternative arrangement **B** the complexation is between the guanidine and the azlactone enolate function. This means that the two acidic protons of guanidinium coordinate the terminal oxygen and the associated double bond in the azlactone ring. This arrangement forces the rings in the two fragments to be in approximately perpendicular planes.

In the chiral guanidinium derivatives used in experiment, the two nitrogens, and their respective protons, are not equivalent. This adds a new source of structural complexity, and results in the four possible adducts presented in Figure 4.2. They are labeled according to the expected chirality, R or S, of the final product. hen, there are a total of four possible complexes to study. These four adducts are: **AR**, **AS**, **BR** and **BS**.

86
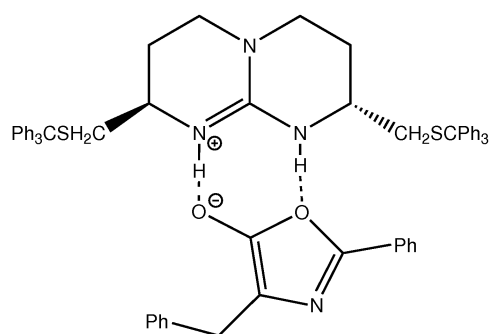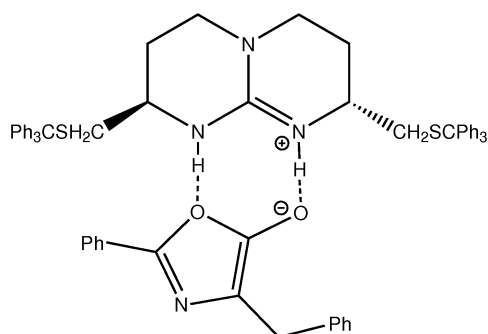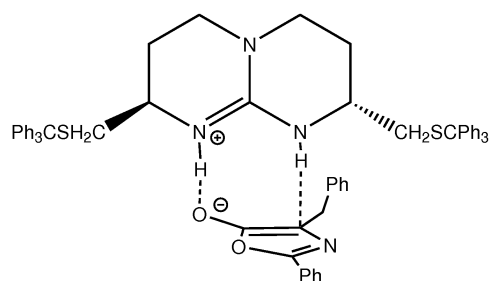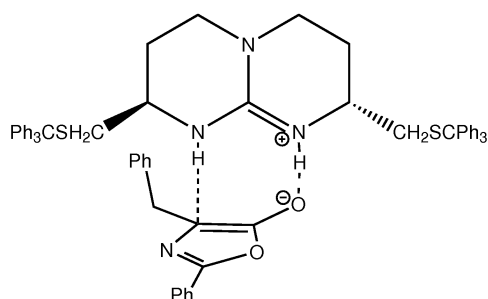
Figure 4.2: Four initial models for complexation between the guanidinium derivative and the deprotonated azlactone.

4.1 Structure and energy of the adducts in a model system

A first set of calculations was carried out in a model system. This model, shown in Figure 4.3, consisted of an unsubstituted azlactone and the small guanidinium fragment. The guanidinium model is not chiral, and thus only the two arrangements **A** and **B** indicated in the Figure are available. In the first one, Figure 4.3A, both subunits are in the same plane and the protonated guanidinium interacts with both oxygens of azlactone. In the second approach, Figure 4.3B, the subunits are in perpendicular planes, and the protonated guanidinium in-

teracts with the terminal oxygen and the double bond in the ring closer to this oxygen.
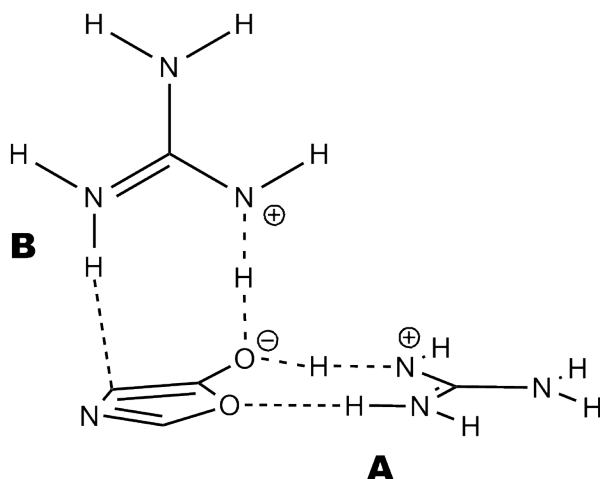
Figure 4.3: The model system used in the initial study of the guanidinium-azlactone adducts, and the two arrangements considered.

Full geometry optimizations in gas phase were done with the Becke3LYP density functional[18,19,177] method as implemented in Gaussian03 package[176] using the 6-31G(d) basis set. Because of questions on the reliability of DFT for the reproduction of hydrogen bonds, additional set of single point energy calculations, also in gas phase, were carried out with the MP2 and MP4 methods. The presence of charge separation hints to important solvation effects, and single point energy calculations on the Becke3LYP optimized geometries were carried out using PCM method. The solvent was that applied in the experimental procedure, toluene, (ϵ=2.379). These PCM calculations were carried out using different methods, Becke3LYP, MP2 and MP4, with always the same 6-31G(d) basis set.

Figure 4.4 shows the optimized structure for the **A** adduct. As expected, both fragments remain in the same plane and the guanidinium group interacts with

4.1. Structure and energy of the adducts in a model system

Figure 4.4: Becke3LYP optimized geometry for the **A** adduct in the model reaction between a guanidinium and an azalactone ring
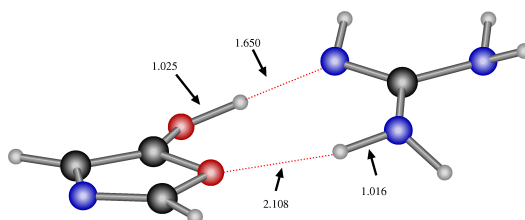
both oxygens of the azlactone. However, in this gas phase optimized geometry the expected charge separation is absent. Both fragments remain neutral, and there is a hydroxyl group on the azlactone side. In any case, this hydroxyl group is stabilized by a strong hydrogen bond with one of the nitrogens of the guanidinium group and the H-N distance is 1.650 Å. The other nitrogen-oxygen interaction seems to be weaker. The hydrogen remains bound to the nitrogen (1.025 Å) and the oxygen is further away (2.108 Å) from the hydrogen.

Table 4.1: Bonding energies (in kcal/mol) with respect to separated reactants for the guanidinium-azlactone adduct models **A**, **Abis** and **B**. Different computational descriptions are applied for gas phase (GP) and toluene solvation (PCM).

Molecule	B3LYP_{GP}	B3LYP_{PCM}	MP2_{GP}	MP2_{PCM}	MP4_{GP}	MP4_{PCM}
Reactants	0	0	0	0	0	0
A	-107.1	-42.2	-110.8	-45.7	-111.2	-46.1
Abis	-103.6	-44.2	-105.7	-46.5	-105.7	-46.5
B	-102.6	-42.9	-105.4	-45.9	-105.1	-45.7

The Becke3LYP bonding energy with respect to the separated reactants is -107.1 kcal/mol in gas phase and -42.2 kcal/mol in solution (Table 4.1). The bond-

ing energy is substantially lower in solution because of the the charged nature of the reactants. The protonated guanidinium and the anionic azlactonate are more stabilized by the PCM calculation, and the stabilization of the reactants decreases the bonding energy. The MP2 and MP4 results are also shown in Table 4.1). The absolute bonding energy is slightly higher (-110.8, -111.2 kcal/mol in gas phase; -45.7, -46.1 kcal/mol in solution), but the trend is the same as in the Becke3LYP calculation.
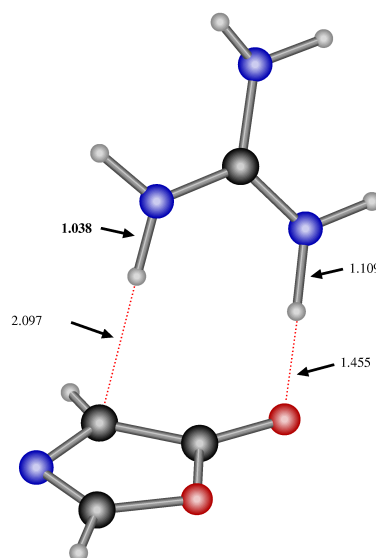
Figure 4.5: Becke3LYP optimized geometry for the **B** approach in the model reaction between a guanidinium and an azalactone ring

Figure 4.5 shows the optimized structure for the alternative approach **B**. As expected, the interaction between the guanidinium group and the enolate function of the azlactone forces the arrangement of both molecules in approximately perpendicular planes. In this case, the two fragments conserve the expected charges. The protonated guanidinium group is stabilized by a strong hydrogen bond with the terminal oxygen of the azlactone, being the distance H-O 1.455 Å. This hydrogen is at 1.109 Åfrom the corresponding nitrogen. The other N-H

90

4.1. Structure and energy of the adducts in a model system

group is involved in a weaker hydrogen bond (2.097 Å) to the sp^2 carbon of the enolate functional group of the azlactonate.

The Becke3LYP bonding energy for this **B** adduct is -102.6 kcal/mol in gas phase and -42.9 kcal/mol in solution (Table 4.1). The difference between the gas phase and the PCM calculation is again in the high stabilization of the charged reactants in solution. The computed MP2 and MP4 bond energies are -105.4, -105.1 kcal/mol in gas phase and -45.9, -45.7 kcal/mol in solution.

It is remarkable that according to these calculations the **A** form would be more stable in gas phase regardless of the method, by a value between 4.5 and 6.1 kcal/mol, but the relative energies would be much closer in solution. In toluene solution, **B** would be more stable than **A** by 0.7 kcal/mol at the Becke3LYP level and by 0.2 kcal/mol at the MP2 level, with **A** being more stable by 0.4 kcal/mol at the MP4 level.

Table 4.2: Solvation energies (in kcal/mol) of the guanidinium-azlactonate adducts in the model system for each method

Molecule	B3LYP	MP2	MP4
A	-6.8	-7.1	-6.9
Abis	-12.4	-12.9	-12.8
B	-12.0	-12.7	-12.5

This seemingly complex set of results can be easily interpreted from the results in Table 4.1, where the solvation energies of the adducts are presented. The solvation energy for a given adduct is quite stable with respect to the computational method. The Becke3LYP, MP2 and MP4 values for adduct **A** are -6.8, -7.1 and -7.9 kcal/mol; while the corresponding values for adduct **B** are -12.0,

12.7 and -12.7 kcal/mol. Adduct **B** is systematically more stabilized by solvation, and this is because of the presence of two charged subunits in this adduct. Things are different for adduct **A**, where the two subunits are neutral. One may wonder if the optimized structure in gas phase for **A** would correspond to the optimal structure in solution.
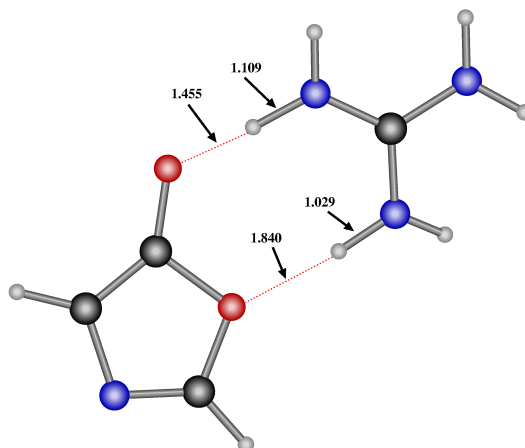
Figure 4.6: Becke3LYP optimized geometry for the **Abis** adduct in the model reaction between a guanidinium and an azalactone ring

To compare how the presence of this charge separation should affect the energy of the adduct **A**, a new system, **Abis** was computed. In this new adduct, the distances between the terminal oxygen of the azlactone and the hydrogen, as well as the distance of this hydrogen to the nitrogen were frozen. The distance O-H was frozen to 1.455 Å, and the distance H-N to 1.109 Å. These constrained distances were the same of the O-H-N optimized system of the **B** adduct.

Gas phase optimization of the constrained structure **Abis** produced the geometry shown in Figure 4.6. Now, both subunits remain charged, the azlactone is anionic while the guanidinium group is cationic. The hydrogen attached to nitrogen is stabilized by a strong hydrogen bond with the anionic oxygen of the azlactone. The forced charge separation between both molecules reduces

92

4.1. Structure and energy of the adducts in a model system

the distance between the subunits, especially the other O-H distance that now becomes shorter, 1.840 Å.

The presence of the charge separation in **Abis** affects the energy of the system. The gas phase bond energy at the Becke3LYP decreases when going from **A** to **Abis** (-103.6 kcal/mol, from -107.1 kcal/mol) because of the constrained nature of the latter calculation. The same trend is reproduced in the single point MP2, MP4 calculations on the Becke3LYP geometry. However, the trend is inverted in the PCM calculations in solution. In all cases, the **Abis** structure is more stable in solution than **A**, by values between 2.2 and 0.4 kcal/mol. Solvent effects, as expected, favor the structure with charge separation.

The results for the solvation of structure **Abis** are also consistent with those for **B** (see Table). Solvation is very similar for both structures.

The summary for the calculations on the model system is thus the following:

- Two interaction modes exist, **A** and **B**, as predicted by the qualitative analysis, and both of them produce local minima.

- The gas phase geometry optimization produces an adduct between two neutral fragments for structure **A**, but this is less stable than the ion pair form **Abis** in solution.

- **Abis** is the most stable structure, but the other structures are sufficiently close in energy in the model system that they are viable alternatives for the real system, and they have to be considered.

- The Becke3LYP, MP2 and MP4 methods produce similar results, especially in the relative energy of the different isomeric forms. Because of this, the Becke3LYP method will be used in the next section

4.2 Structure and energy of the adducts in the real system

After the study on the model system, the experimental system defined by a bicyclic guanidine and a substituted azlactonate ring was studied.

Because of the size of the system, it was very expensive to work at a pure quantum level with the whole molecule, and a QM/MM description was applied. A hybrid method was chosen because of the size. The QM layer was described with the Becke3LYP density functional method[18,19,177] as implemented in Gaussian03[176] package and the 6-31+G(d) basis set.[179,180] The MM layer was computed using the UFF force field[185] as implemented in the same software package. As one can see in Figure 4.7, the QM layer included the core of the bicyclic guanidine involving its substituents up to the sulfur atom; and the core of the azlactone, with its main ring and the phenyl substituent. The rest of the molecule was included in the MM layer. This included the three phenyl substituents of the catalyst and the benzyl substituent of the azlactone. This setting was found to provide the best price/quality ratio after test calculations with different partitions. Full geometry optimizations were carried out at the ONIOM(Becke3LYP:UFF) level. Solvation effects with toluene were evaluated through single point PCM calculations at the Becke3LYP level with a 6-31+G(d) basis set.

The hybrid QM/MM ONIOM method as implemented in Gaussian03 package was used.

Due to the free rotation of the phenyl rings there are lots of different conformations for the bicyclic guanidine, and this required the performance of systematic conformational searches. They were carried out with the Macromodel program.[187] The conformational search was carried out with the MonteCarlo

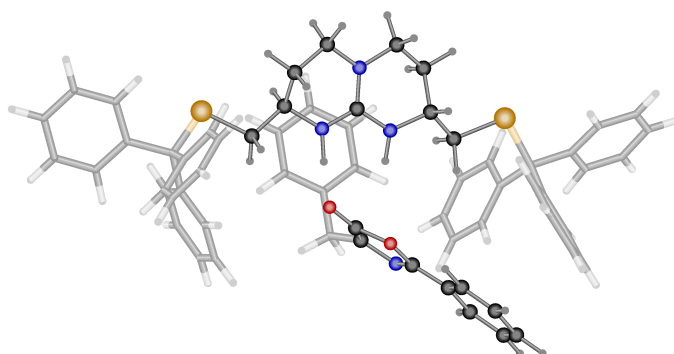## 4.2. Structure and energy of the adducts in the real system

Figure 4.7: QM/MM partition used in the ONIOM calculation of the guanidinium/azlactone adduct. QM layer is is ball-and-stick and MM layer is in tube representation.

Multiple Minimization (MCMM) method[188,189] as implemented in the program. The atoms of the bicyclic guanidine and the main ring of the azlactone were frozen to the positions similar to the computed model arrangements and the rest was optimized. The force field applied for this conformational search was OPLS2001.[190] The geometries resulting from the conformational search were then used as starting points for the ONIOM calculations.

The two approaches discussed above, **A** and **B** were used for the starting point of the calculations. The four proposed arrangements, **AR**, **AS**, **BR** and **BS** (see Figure 4.2) were computed using the full experimental complex (the CH_2SCPh_3 disubstituted bicyclic guanidine and the CH_2Ph substituted azlactone).

Unexpectedly, the free optimization calculations led to a structure that was neither **A** nor **B**, and that we labeled as **C**. The new type of structure is shown in Figure 4.8. In this new approach, the guanidinium is protonated, and makes two hydrogen bonds with the same terminal oxygen of the azlactone. The appearance of this new arrangement is likely associated to the steric strains be-
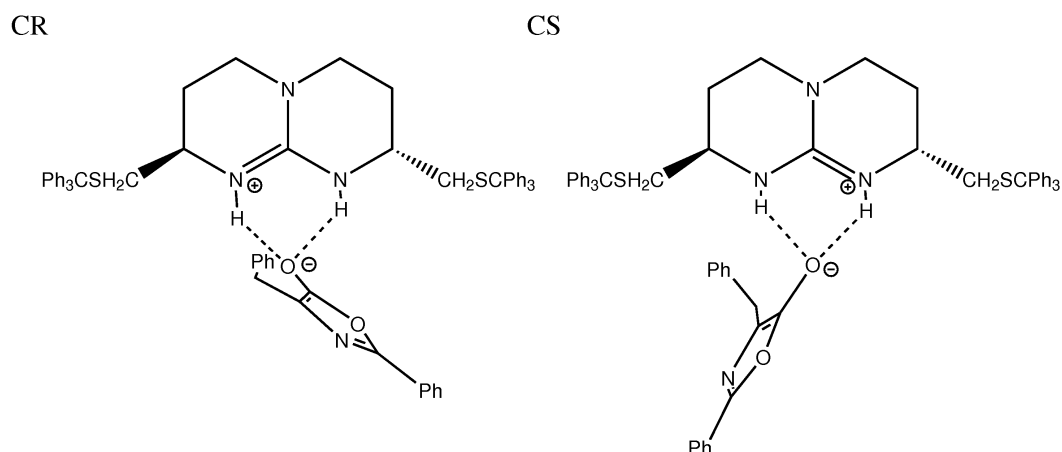
95

Figure 4.8: Arrangement **C** for the adduct between guanidinium and deprotonated azlactone.

tween the substituents of guanidinium and azlactone, that preclude the **A** and **B** approaches that were available in the model system. The defining feature of adduct **C** is the presence of the two hydrogen bonds pointing towards the same oxygen. Within this arrangement, only the two structures shown in Figure 4.8 are possible. They differ on the side towards which is tilted the azlactone plane. As a result of this tilt, one of the faces remains less hindered, and this is why the structures are labeled as **CR** and **CS**. In the two structures of Figure 4.8 the benzyl substituent of the azlactone is pointing backwards. The alternative structures with the benzyl pointing forward would be equivalent because of the C_2 symmetry of the guanidinium unit.

Figures 4.9 and 4.10 present the ONIOM(Becke3LYP:UFF) optimized geometries for **CR** and **CS**, respectively. Analysis of Figure 4.9, which corresponds to the **CR** adduct, confirms that this arrangement is neither **A** nor **B**. In this structure, the planes of the guanidinium ring and the azlactone are not approximately coplanar, as in **A**, and the azlactone plane is not approximately perpendicular to the guanidinium C_2 axis, as in **B**. The defining feature of this struc-

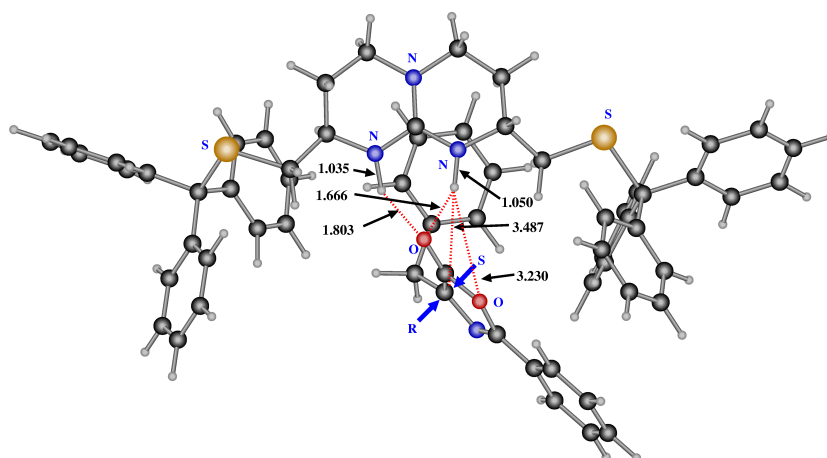## 4.2. Structure and energy of the adducts in the real system

Figure 4.9: ONIOM(Becke3LYP:UFF) optimized structure for the **CR** adduct.

ture is the presence of two hydrogen bonds between the terminal oxygen of the azlactone fragment and both nitrogens from the guanidine. These hydrogen bonds are not weak, being the O-H distances 1.666 Å and 1.803 Å.

Other characteristic features of structures **A** and **B** are also absent from **CR**. In the **B** approach, there is a short distance between the sp^2 carbon of the enolate group of the azlactone and one of the guanidinium hydrogens. This distance is 3.487 Å in **CR**. In a **A** arrangement, there must be a short distance between a proton and the oxygen of the azalactone ring. This distance is 3.230 Å in **CR**.

The labeling of this structure as **CR** can be understood from the geometry shown in Figure 4.9. The pro-**R** face of the azlactone is less hindered (see blue arrows in the Figure 4.9). Therefore this adduct will lead majoritarily to the R product.

The discussion of the **CS** optimized structure, which is represented in Figure 4.10, follows similar lines. Although the starting geometry in the conformational search was closer to **BS**, the optimization process drove it to an intermediate geometry. The arrangement fits neither with the **A** nor with the **B** labeling.
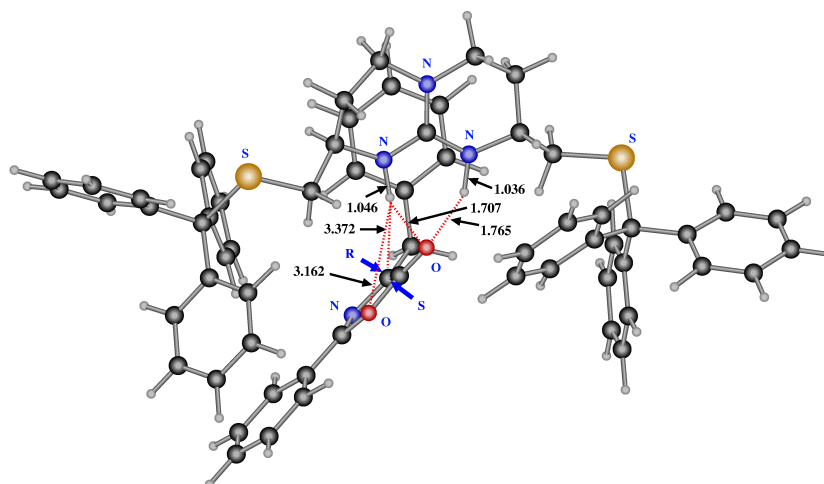
Figure 4.10: ONIOM(Becke3LYP:UFF) optimized structure for the **CS** adduct

In adduct **CS**, there are two hydrogen bonds between the guanidinium protons and the anionic terminal oxygen of the azlactone. The distances O-H are 1.707 Å and 1.765 Å. Interactions of the guanidinium protons with other targets in the azlactone ring are weaker. The shortest contact with the sp^2 carbon of the azlactone enolate is 3.372 Å and the corresponding value for the oxygen within the azlactone ring is 3.162 Å.

Table 4.3: Energy difference (kcal/mol) between the **CR** and **CS** adducts computed at different levels. A negative value corresponds to **CR** being the most stable species.

Method	Geometry	QM energy	MM energy	Total energy
ONIOM(B3LYP:UFF)	ONIOM(B3LYP:UFF)	-2.5	-0.5	**-2.9**
Becke3LYP	ONIOM(B3LYP:UFF)	-1.8	–	**-1.8**
Becke3LYP + solvent	ONIOM(B3LYP:UFF)	-3.4	–	**-3.4**

4.2. Structure and energy of the adducts in the real system

Table 4.3 collects the energy difference between the **CR** and **CS** adducts at different computational levels. The **CR** adduct is more stable in all cases, in agreement with the experimental observation of the *R* product. The energy difference is 2.9 kcal/mol in the ONIOM(Becke3LYP:UFF) calculation. The value is lowered to 1.8 kcal/mol in the full QM Becke3LYP single-point calculation. The value increases to 3.4 kcal/mol when solvent effects are introduced. The difference between the ONIOM and the full QM calculation is probably related to the fact that the UFF force field is not perfect, but it was the only available in the program used.

Solvation is significant for both structures, but quite similar in value (-12.3 kcal/mol for **CR**; -10.6 kcal/mol for **CS**). The charge separation between the two subunits explains this important absolute value, even in such a low polarity solvent as toluene. This result was already found in calculations on the model system. However, solvation is similar for both isomers, and the net result is a total energy of 3.4 kcal/mol in favor of the **CR** adduct.

The computed energies are clearly in favor of the **R** product, and this is in agreement with experiment. However, the energy difference is excessive. If one assumes that the Boltzmann distribution of the adducts at 195 K rules the selectivity, the values presented in Table 4.3 would lead to enantiomeric excesses of 99.9% in the PCM calculation, 99.8% in the ONIOM calculation and 98.1% in the Becke3LYP calculation. These numbers are notably higher than the experimentally reported value of 87%. There are however a number of causes that can explain this numerical disagreement. The discrepancies between the ONIOM(Becke3LYP:UFF) and pure Becke3LYP hint to possible inadequacies in the force field, which is used in the geometry optimization. The most critical source of error is however probably in the assumption that the energy of the adduct can rule the enantiomeric excess. This should be more directly corre-

lated with the energy of the selectivity determining transition states.

Since the MM energy difference between both optimized geometries is only 0.5 kcal/mol (**CR** is more stable than **CS**), the main energy difference in the ONIOM calculation provides the QM layer (2.5 kcal/mol). Comparing both ONIOM(B3LYP:UFF) optimized geometries, **CR** and **CS** one can see that the distances between the azlactone and the bicyclic guanidine are lower in the **CS** structure than in the **CR** geometry.

In summary, the combination of MM, QM and QM/MM techniques has allowed us to reproduce the experimentally observed selectivity in the ring opening of real systems are quite different from those that could be predicted a priori in model systems. The relative stabilities of these adducts, which are at the origin of enantioselectivity, depend on a subtle combination of steric and electronic factors in the interaction between the two subunits of the adduct.

Chapter 5

Selenoproteins

5.1 QM study of the catalytic mechanism of FDHh

To understand the effect of selenocysteine in a selenoprotein, the catalytic mechanism of the Formate dehydrogenase H (FDHh) was studied. This enzyme catalyzes the oxidation of a formate anion to carbon dioxide with release of one proton and two electrons.

5.1.1 Catalytic cycle

In 1997, when the crystal structure of this protein was resolved and published,[152] the authors suggested a possible catalytic cycle for the enzyme. This catalytic reaction would begin with coordination of a formate anion molecule to the molybdenum central atom. Then, the formate anion is oxidized to carbon dioxide releasing the hydrogen to the selenium atom, which later will transfer the proton to a nearby histidine, and two electrons to the central metal atom, which changes its oxidation state from Mo^{VI} to Mo^{IV}. Finally, molybdenum is reoxidated (in two steps) to the initial oxidation state (+6) to restart the reaction. A

101

catalytic cycle based in this proposal will be presented in this thesis. It consists of six steps, and is schematically shown in Figure 5.1.

The cycle begins in **A**, where a molecule of the substrate (formate anion) is coordinated to molybdenum. The first step is the oxidation of this formate anion to carbon dioxide. This oxidation implies a transfer of the hydrogen to selenium and two electrons from the ligand to the molybdenum center, which is formally reduced from Mo^{VI} to Mo^{IV}. The result of this reaction is complex **B** where the carbon dioxide product remains associated to the active center. Next step is the release of the product, carbon dioxide leaving the enzyme, producing system **C**.

Afterwards, the enzyme has to be regenerated to restart the catalytic reaction with a new formate anion. In **C**, the selenium is protonated while the molybdenum is reduced. The first step to recover the original enzyme is the transfer of the proton to an imidazole ring of the topologically close histidine (His[141]) obtaining **D**. In **D**, the selenium has already been deprotonated but molybdenum remains reduced.

The regeneration of the oxidation state of the molybdenum atom follows a two step mechanism,[171] where the release of the first electron yields intermediate **E**, with and unstable Mo^{V}. Then, a second electron leaves the active center of the enzyme while a hydroxyl is coordinated to the molybdenum atom and stabilized by a hydrogen bond with the protonated histidine. A new catalytic cycle can start when this system, **F**, releases a water molecule and a new formate anion joins the system.
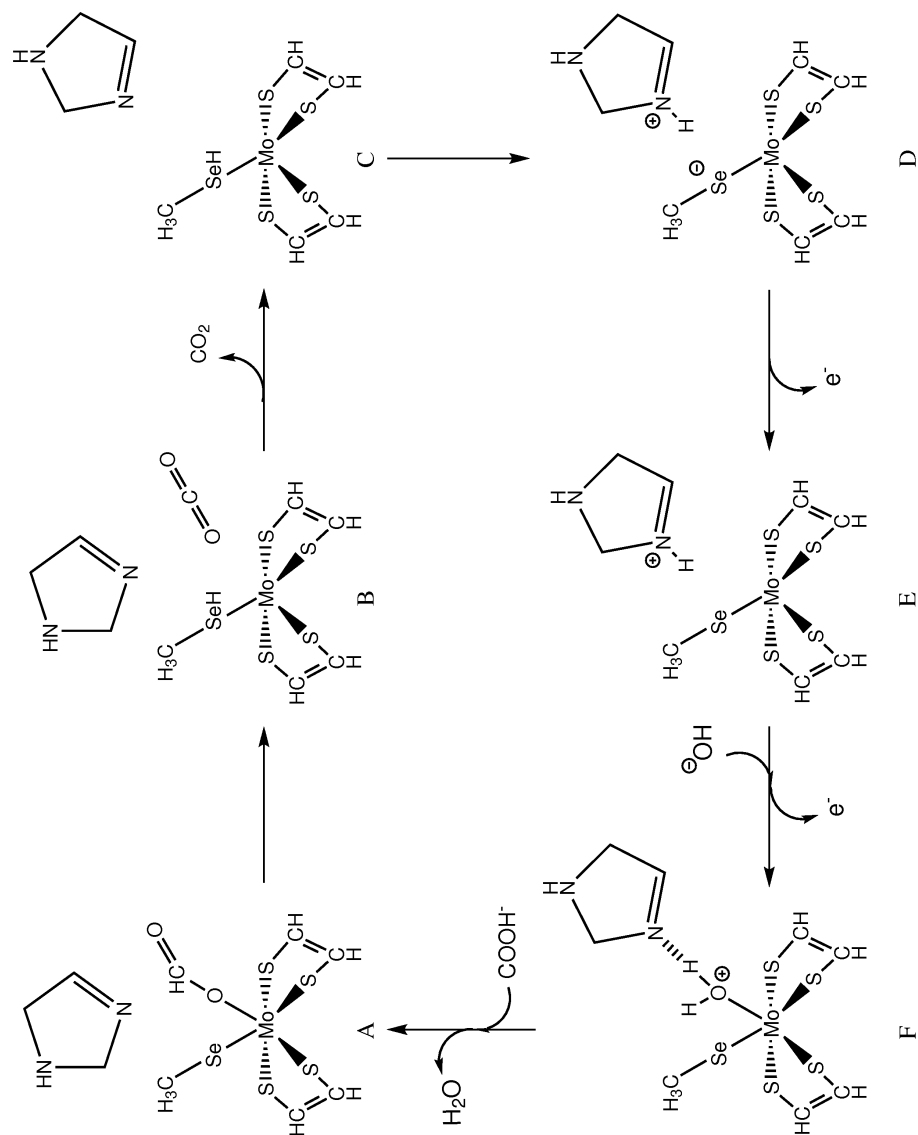
5.1. QM study of the catalytic mechanism of FDHh

Figure 5.1: Proposed catalytic cycle model for the FDHh enzyme

103

These six steps can be classified in two blocks: the oxidation of the formate anion and the regeneration of the enzyme. The process from **A** to **B** represents the oxidation of the formate anion. The transformation from **C** to **F** corresponds to the regeneration of the enzyme. This regeneration can be divided in two steps: deprotonation of selenium, from **C** to **D**; and reoxidation of molybdenum, from **D** to **F**.

5.1.2 Computational details

All calculations were carried out at the QM level using a model of the active center of the protein (see Figure 5.2).Although it is generaly acknowledged that the protein environment is critical for catalysis typically through the role of electrostatic interactions,[191] the use of a small model system in this study is justified, as we are mainly interested in the differential electronic effects of Cys and Sec. The size of the model was adjusted to include all necessary atoms for the reaction, while keeping the calculations feasible. Because of that, the model includes the central metal molybdenum atom, four sulfurs from both MDG cofactors, the selenium atom and a methyl which represents the selenocysteine (Sec_{140}) sidechain and an imidazole ring from the histidine$_{141}$. The starting point for geometry optimizations was taken from the PDB (Protein Data Bank), file accession code 1FDI. It corresponds to the oxidized form of the formate dehydrogenase H, published in 1997 by Boyington and co-workers.[152]

All minima and transition states of the catalytic cycle (section 5.1.1) were computed. In each step a free geometry optimization was carried out with the Becke3LYP density functional[18,19,177] as implemented in the Gaussian03 package[176] using the LANL2DZ basis set[192] for molybdenum, selenium and sulfur, and the 6-31+G(d) basis set[179,180] for other atoms. A d polarization shell was included for selenium and sulfur.[193] To simulate the effect of the solvent, single

5.1. QM study of the catalytic mechanism of FDHh
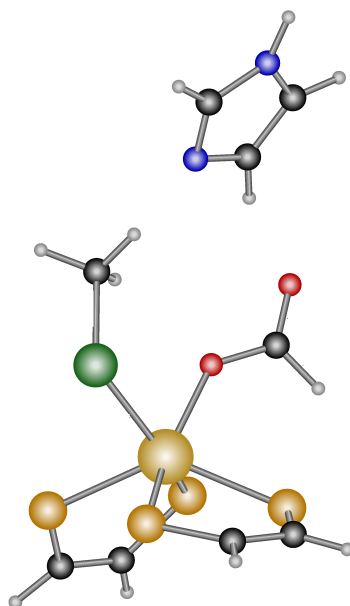
Figure 5.2: QM model for the active center of FDHh

point energy calculations were carried out using the PCM method[194] with each optimized geometry. A solvent with a dielectric constant ϵ close to 40 was used, this solvent was acetonitrile (ϵ=36.64). This choice of ϵ was made to simulate the effect of the hydrophilic region near the surface of a protein[?, 152]

To better evaluate the role of selenocysteine, an extra set of calculations was carried out replacing selenium by sulfur. In this way, the role of selenocysteine could be compared with that of cysteine.

Benzyl viologen

As mentioned above, the catalyzed reaction releases two electrons, is a redox reaction. The wild type enzyme has a sulfur-iron box (Fe_4S_4) as a primary electron acceptor. Computation of this kind of molecules is not easy, and they were not used in our study. Instead, this electron acceptor was replaced by another

one, easier to compute, which is also applied in biochemistry. It is benzyl violo-gen (1,1'-dibenzyl-4,4'-bipirydine). This is in fact used as final electron acceptor and also as indicator of the activity of the enzyme in experimental studies with this enzyme.[171–173]
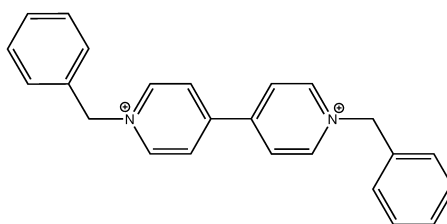
Figure 5.3: Bicationic form of benzyl viologen

Benzyl viologen (Bv) is an organic molecule which is used experimentally as an electron acceptor, especially in biochemistry. This molecule has different oxidation states, as neutral (Bv), cationic (Bv^+) and bicationic (Bv^2+). We computed the energy of each oxidation state and the energy differences between them were estimated as the necessary energy to remove one electron from our model of the active center. These energies are summarized in Table 5.1. Then, we will assume that the first electron that leaves the system (from **D** to **E**) is used to convert the bicationic form of benzyl viologen into the monocationic form, and that the second electron (from **E** to **F**) is used to convert the monoca-tionic benzyl viologen into the neutral form.

Including this electron acceptor, the overall net reaction becomes:

$$HCOO^- + OH^- + Bv^{2+} \longrightarrow CO_2 + H_2O + Bv \tag{5.1}$$

5.1. QM study of the catalytic mechanism of FDHh

Table 5.1: Computed energy (in kcal/mol) of the different oxidation states of benzyl viologen.

Molecule	Energy	Difference
Bv^{2+}	-360757.5	–
Bv^{+}	-360853.6	-96.1
Bv	-360927.9	-74.2

5.1.3 Overview of energy profiles

Figure 5.4 presents the energy profile (in kcal/mol) of the computed catalytic cycle (Figure 5.1) with selenium. The relative energies of intermediates and transition states are given with respect to intermediate **A**, where the formate anion is coordinated to the molybdenum atom. All energies are the PCM energies of the gas phase optimized structures.

The first remarkable thing is the exothermicity of the net reaction (eqn. 5.1) which is -26.4 kcal/mol, it is an exothermic reaction. According to the catalytic cycle shown, there are three barriers in this profile. They correspond to oxidation of the formate, deprotonation of selenium and reoxydation of the molybdenum atom. The barriers for oxidation of formate and reoxidation of molybdenum are moderately high, 25.3 kcal/mol and 13.7 kcal/mol, respectively. In contrast, the deprotonation of selenium has a very small barrier of 0.1 kcal/mol.

The barrier to oxidation of the formate is characterized by the transition state connecting **A** and **B**. Characterization of the barrier to reoxidation of molybdenum is more subtle. It corresponds to the energy of the high energy intermediate **E** in the path from **D** to **F**. This is because we assume the electron transfer processes from molybdenum and the electron sink to be barrierless. The raw data for the reaction with selenium suggest the oxidation of formate to be the
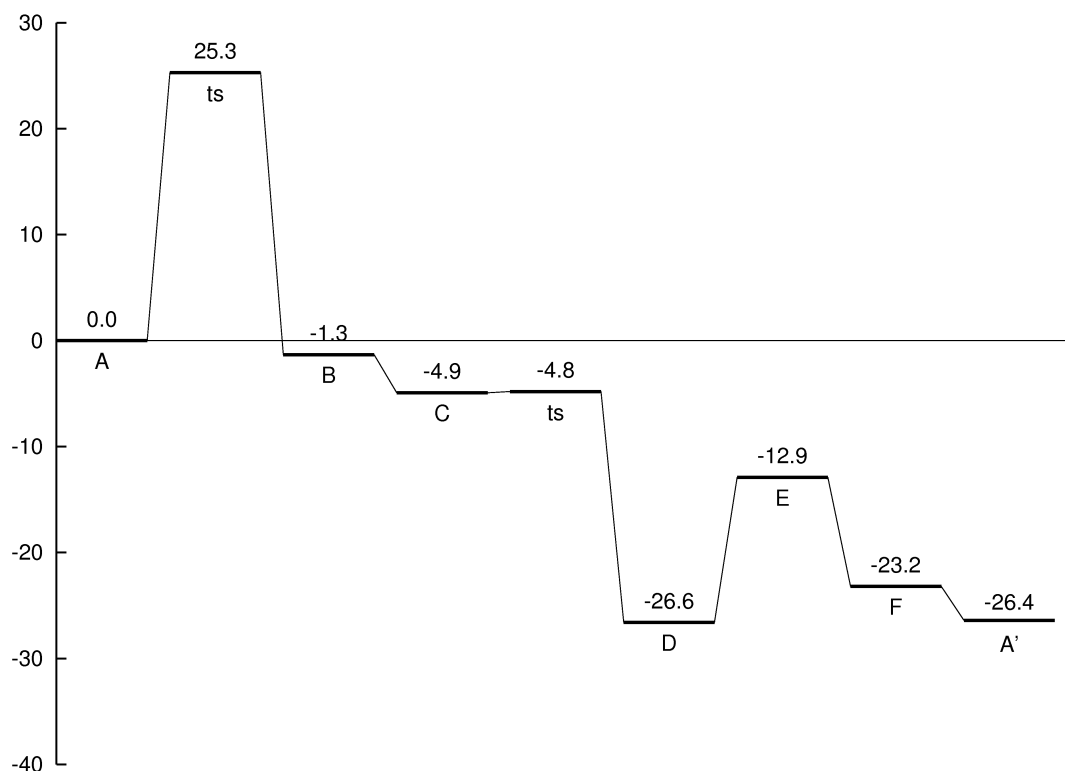
Figure 5.4: Computed energy profile (in kcal/mol) for the catalytic cycle of FDHh with the selenium system.

rate-determining step with a barrier of 25.3 kcal/mol (the energy of **TS(A-B)** minus that of intermediate **A**) compared with the 13.7 kcal/mol of the reoxidation of molybdenum (the difference between **E** and **D**).

The analogous catalytic cycle for the system where selenium is replaced by sulfur is shown in Figure 5.5. The overall pattern is similar. There are two moderately high barriers for the oxidation of formate (22.6 kcal/mol) and reoxidation of molybdenum (19.8 kcal/mol), and a very low barrier for the deprotonation of sulfur. In this case the barrier from **C** to **D** is formally negative (the

108

5.1. QM study of the catalytic mechanism of FDHh
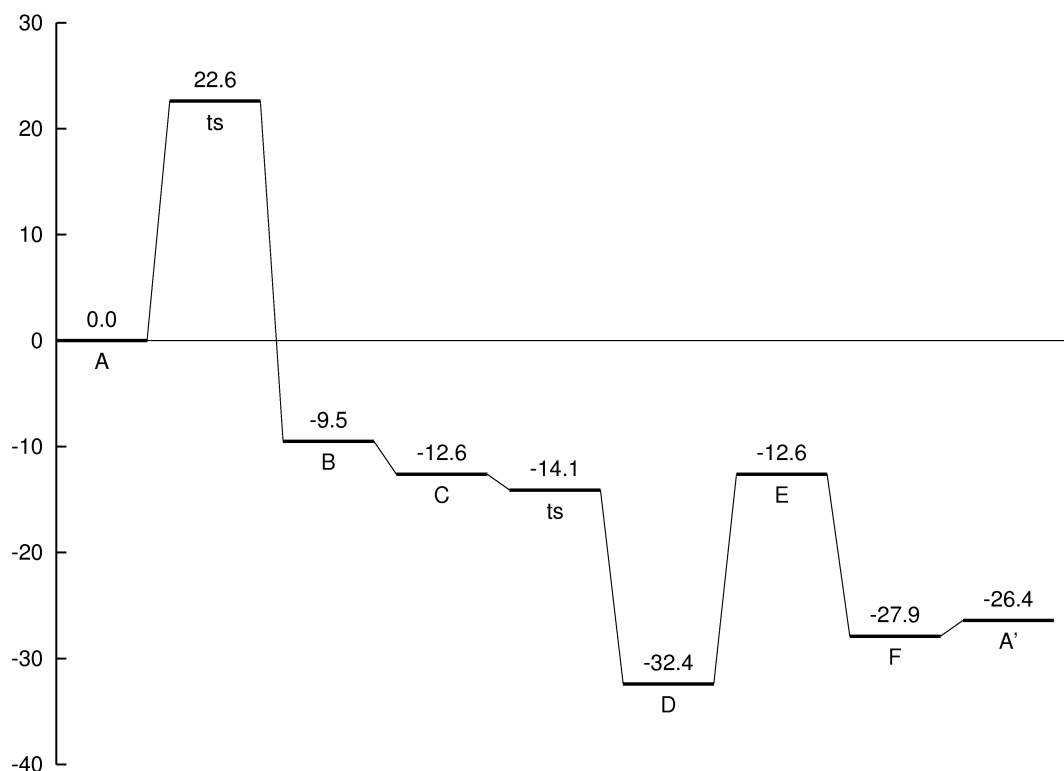
Figure 5.5: Computed energy profile (in kcal/mol) for the catalytic cycle of FDHh with the sulfur system.

energy of transition state is below the of **C**), but this is a computational artifact due to the fact that the geometry optimization is carried out in gas phase and the final energy includes environmental effects through PCM. A more refined calculation would be necessarily to properly evaluate this barrier, but it has not been done because this seems to have little relevance for the problem under discussion.

When we compare the raw computed values for the two main energy barriers for the two systems, we see that for each system the highest barrier corresponds to oxidation of the formate, and that the value for this barrier is higher for selenium than for sulfur. This is completely at odds with the experimen-

Table 5.2: Computed relative energies (kcal/mol) for minima and transition states involved in the catalytic cycle of FDHh with both the selenium and sulfur models.

Molecule	ΔE_{Se}	ΔE_S	Difference (Se-S)
A	0	0.0	0.0
ts	25.3	22.6	2.6
B	-1.3	-9.5	8.2
C	-4.9	-12.6	7.7
ts	-4.8	-14.1	9.3
D	-26.6	-32.4	5.8
E	-12.9	-12.6	-0.3
F	-23.2	-27.9	4.6
A′	-26.4	-26.4	0.0

tal observation that the reaction is orders of magnitude faster with selenocysteine than with cysteine. In spite of this seemingly disagreement the results are however very informative. For this, one has to take into account the relative accuracy of the calculations. It is quite reasonable to assume a similar accuracy in the description of the behavior of the selenium or sulfur systems in a given step, either formate oxidation or molybdenum reoxidation. Mostly because of the cancelation of systematic errors, which will be essentially the same. However, things are different when we compare the barriers for formate oxidation and molybdenum reoxidation for the same system. The difference between the two steps are evident, and so systematic errors may affect the numeric result in different directions. Because of this, we interpret the observation that the barrier for molybdenum reoxidation is lower for selenium as an indication that is the key step of the process. Selenocysteine exists thus in this particular enzyme because if favors the reoxidation of molybdenum, although it hinders the step

where formate is oxidized.

A detailed study of each step will be shown in what follows, and the differential effect of selenium and sulfur in each of them will be analyzed.

5.1.4 Oxidation of formate

This first step begins with intermediate **A**, where the central atom, molybdenum is Mo^{VI}. Besides the four *cis*-dithiolene sulfurs of the MGD cofactors, two more anionic ligands are coordinated to molybdenum. One is $Me\text{-}Se^-$, wich represents the selenocysteine of the real system, and the other one is the substrate of the reaction, a formate anion ($OCHO^-$). This formate anion is coordinated to molybdenum through one of its oxygens.

The product of this step, **B**, is the complex between the enzyme and the product of the oxidation of the formate anion, a molecule of carbon dioxide. In the transition state the hydrogen atom from the formate ligand is transferred to the selenium center. Carbon dioxide is produced and the molybdenum center of the enzyme is reduced to oxidation state +4.

Figure 5.6 collects the structures involved in this step for the selenium model. The Figure presents intermediate **A**, the transition state from **A** to **B**, and intermediate **B**. The structure **A** is the enzyme-substrate complex, where the substrate (formate) is coordinated to molybdenum. **B** is the enzyme-product complex, where formate is already oxidized to carbon dioxide.

Intermediate **A**, the enzyme-substrate complex, has short bond distances between molybdenum and both both the methylselenolate and formate ligands: 2.497 Å for Mo-Se and 2.032 Å for Mo-O. The Mo-Se distance increases in intermediate **B**, up to 2.689 Å. This increase of 0.192 Å is due the lower capacity of the ligand Me-Se-H (in **B**) to coordinate to molybdenum with respect to that of the ligand $Me\text{-}Se^-$ (in **A**).

The transition state shows a typical situation for hydrogen transfer. The Mo-Se and Mo-O distances increas with respect to **A**, up to 2.539 Å and 2.166 Å, respectively. The formate anion, which in **A** has the hydrogen oriented to the opposite side of selenium, is reoriented. In the transition state, the hydrogen is placed between the selenium and carbon atoms, with distances of 1.689 Å and 1.475 Å, respectively.
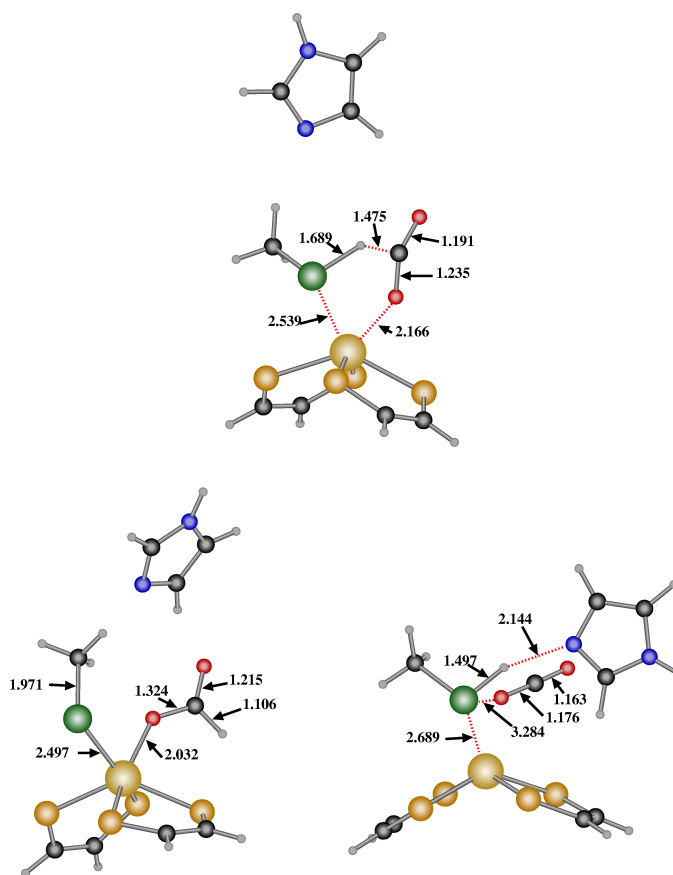

Figure 5.6: Minima and transition states for oxidation of formate in the FDHh model with selenium. The structures are: minimum **A** (bottom left corner), transition state (top) and minimum **B** (bottom right corner). Distances are in Å.

5.1. QM study of the catalytic mechanism of FDHh

Also, in the transition state, the intramolecular C-O distances within the formate subunit are more similar than in **A**. In **A** the distance between the carbon of the formate the carbonylic oxygen is 1.215 Å and carbon-oxygen distance with the oxygen coordinated to molybdenum is 1.324 Å. These distances decrease to 1.163 Å and 1.176 Å, respectively, in **B**, where the carbon dioxide is formed with the typical symmetrical linear structure.

Another remarkable feature of structure **B** is the reorientation of the histidine fragment. In this structure the aromatic nitrogen is oriented towards the protonated selenium. The N-H distance is 2.144 Å, while the Se-H distance is 1.497 Å. The N-H-Se arrangement is almost linear.

Similar structures can be observed in the system where sulfur replaces selenium. They are shown in Figure 5.7. As expected, the main change with respect to the model with selenium is in the distances directly involving the selenium or sulfur atoms. They become, in general, shorter in the model with sulfur. This is logical because of the different atomic radius (115 for selenium and 100 for sulfur, see Table 1.1).

In **A**, the Mo-S distance is 2.362 Å. The value in intermediate **B** is 2.586 Å, an increase of 0.224 Å. In the transition state, the hydrogen is situated between sulfur (at 1.588 Å) and carbon (at 1.439 Å). The S-H distance is reduced to 1.374 Å in **B**. Like in the model with selenium, the imidazole ring in **B** is oriented to receive a hydrogen bond from the protonated sulfur atom, with a N-H distance of 2.058 Å.

Table 5.3 summarizes the energies, for selenium and sulfur of this step of formate oxidation. The exothermicity of the reaction is lower for selenium than
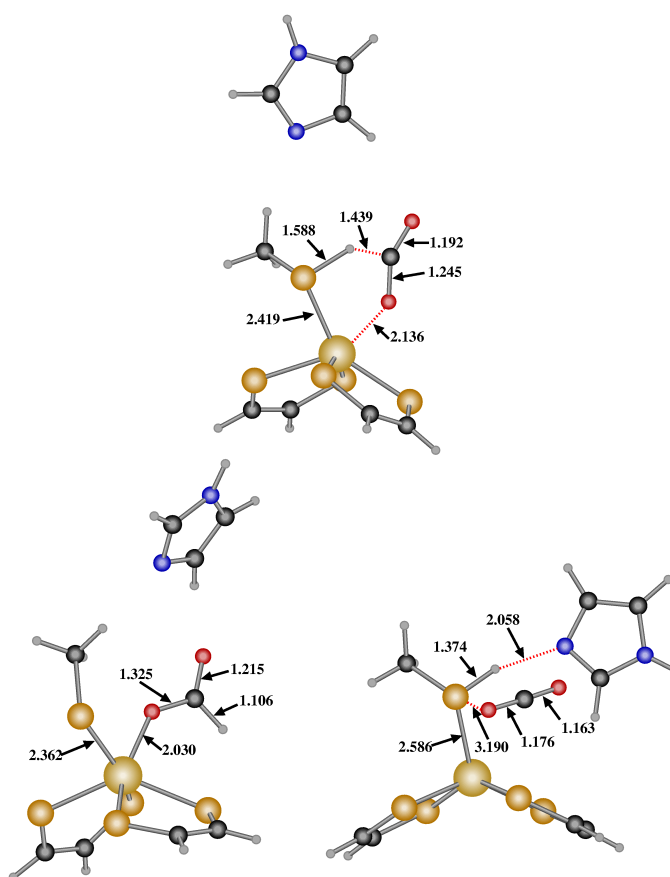
Figure 5.7: Minima and transition states for oxidation of formate in the FDHh model with sulfur. The structures are: minimum **A** (bottom left corner), transition state (top) and minimum **B** (bottom right corner). Distances are in Å.

for sulfur. The reaction is practically thermoneutral for selenium, where **B** is 1.3 kcal/mol below **A**. In contrast, with sulfur, the exothermicity of the reaction is larger, and **B** is 9.5 kcal/mol below **A**. The barriers follow the same trend as the exothermicity. The barrier is lower for the model with sulfur (22.6 kcal/mol) than for the model with selenium (25.3 kcal/mol).

The energetics of the process can be easily correlated with the relative basicities of selenium and sulfur centers and, ultimately, with the atomic electronegativities. Sulfur is more electronegative than selenium. As a result, sulfur cen-

114

Table 5.3: Computed relative energies (kcal/mol) for the species involved in the step of formate oxidation in FDHh.

Molecule	ΔE_{Se}	ΔE_{S}	Difference (Se-S)
A	0	0.0	0.0
ts	25.3	22.6	2.6
B	-1.3	-9.5	8.2

ters are more basic, and because of this they attract protons more easily. This explains why this step should be more favorable for a modified system with cisteine than for the original FDHh enzyme with selenocysteine.

5.1.5 Deprotonation of selenium/sulfur

This step consists in the transfer of the proton from the selenium atom to the imidazole ring representing the histidine aminoacid.

This step begins in **C** which differs from **B** in the fact that the carbon dioxide molecule has already left the active center of the enzyme. The departure of carbon dioxide reduces steric crowding near the metal center and allows the hydrogen bond of the protonated selenium with imidazole to become stronger. In structure **C**, shown in figure 5.8, the Se-H distance is 1.568 Å, and the N-H distance is 1.697 Å. The Mo-Se distance is reduced from 2.689 Å in **B** to 2.617 Å in **C**. Removal the carbon dioxide from the active center of the enzyme (the step from **B** to **C**) stabilizes the system by -3.6 kcal/mol with the selenium model.

Figure 5.8 presents the minima, **C** and **D**, and the transition state involved in the deprotonation of selenium. The Se-H-N arrangement is linear in the three structures, but the proton moves from selenium to nitrogen. In **D**, N-H and Se-H distances are 1.051 and 2.335 Å, respectively. A strong hydrogen bond remains in **D** because selenium has a formal negative charge, and the imidazole is
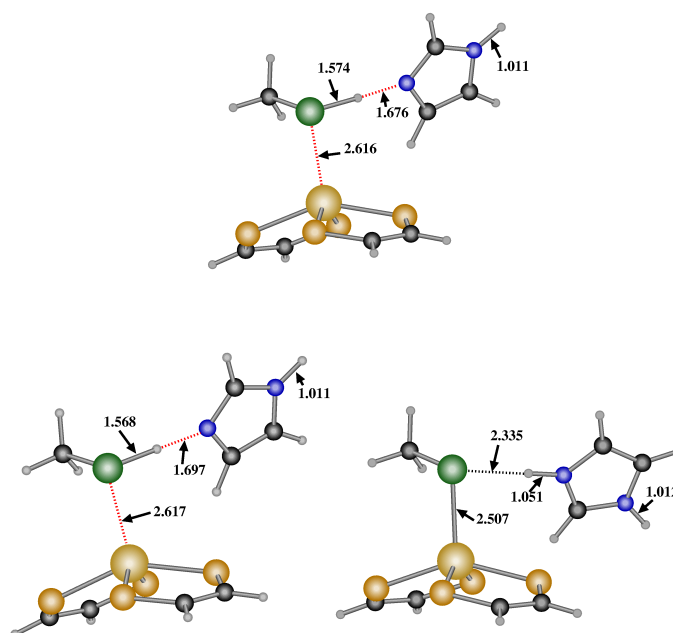
115

Figure 5.8: Minima and transition states for deprotonation of selenium in the FDHh model with selenium. The structures are: minimum **C** (bottom left corner), transition state (top) and minimum **D** (bottom right corner). Distances are in Å.

protonated. A side effect of the transfer of the proton from selenium to the imidazole ring is a reduction in the distance between selenium and molybdenum. The distance is reduced from 2.617 Å in **C** to 2.507 Å in **D**.

The transition state is very similar to intermediate **C**. which can be explained by the very low barrier of this step that will be discussed below. When going from **C** to the transition state there is a little increase of the Se-H distance, from 1.568 Å to 1.574 Å; and a decrease of the H-N distance from 1.697 Å to 1.676 Å.

Structures for the corresponding step in the systems with sulfur (Figure 5.9) reproduce the same trends. Again, the distances which involve the selenium or sulfur atom are shorter for the latter. There is a minor difference in the orientation of the imidazole ring which seems to have little chemical significance. In

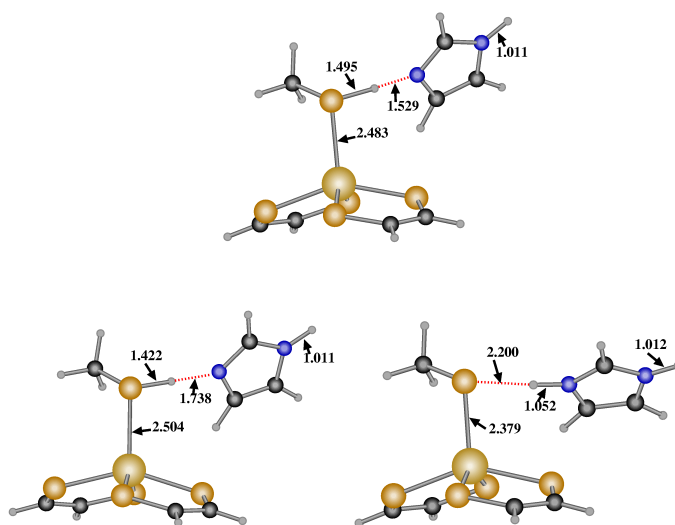## 5.1. QM study of the catalytic mechanism of FDHh

Figure 5.9: Minima and transition states for deprotonation of sulfur in the FDHh model with sulfur. The structures are: minimum C (bottom left corner), transition state (top) and minimum D (bottom right corner). Distances are in Å.

D, the non-reacting nitrogen is oriented upwards for the selenium model, and it is oriented downwards for the sulfur model.

Table 5.4: Computed relative energies (kcal/mol) for minima and transition states involved in the selenium deprotonation step of the FDHh catalytic cycle with selenium.

Molecule	PCM energy	Gas phase energy	Difference (PCM-GP)
C	0.0	0.0	0.0
ts	0.1	0.01	0.09
D	-21.7	-12.7	-9.0

The energetics for the selenium/sulfur deprotonation step are summarized in Tables 5.4 and 5.5. This deprotonation step is quite exothermic, **D** is 21.7

117

kcal/mol below **C** for the selenium system, and 19.8 kcal/mol below for the selenium system. Histidine is thus a sufficiently strong proton acceptor for the catalytic cycle to proceed. The barrier for the process is very low, as corresponds to acid/base reactions: 0.1 kcal/mol for selenium and -1.5 kcal/mol for sulfur. The presence of a formal negative barrier for the sulfur system deserves some comment. The proton transfer generates a charge separation in intermediate **D**, which as a result is stabilized with respect to **C** by the PCM single point energy calculation. This is shown in Tables 5.4 and 5.5, where gas phase relative energies are also presented. The effect of the application of the PCM method is remarkable, and it can be seen comparing the energy of **D** relative **C** in gas phase, -12.7 kcal/mol for selenium and -6.7 kcal/mol for sulfur; and with PCM, -21.7 kcal/mol for selenium and -19.8 kcal/mol for sulfur. The gas phase barrier is also reduced, and as a consequence the value for the sulfur system becomes negative in the single-point PCM calculation on the gas phase optimized geometry.

Table 5.5: Computed relative energies (kcal/mol) for minima and transition states involved in the sulfur deprotonation step of the FDHh catalytic cycle with sulfur.

Molecule	PCM energy	Gas phase energy	Difference (PCM-GP)
C	0.0	0.0	0.0
ts	-1.5	1.4	-2.9
D	-19.8	-6.7	-13.1

The low barrier for the step indicates little kinetic relevance, and because of this the comparison of selenium and sulfur here is not significant. It is however

118

worth mentioning the exothermicity of the step, that is likely irreversible in biological conditions.

5.1.6   Reoxidation of molybdenum

Once the selenium has been deprotonated and this proton transferred to the acceptor (at the moment, towards the imidazole ring of a near histidine), selenium becomes negatively charged while molybdenum remains with oxidation state +4. It is the situation in the intermediate **D**. Molybdenum has to return to the initial oxidation state +6 in order to restart the catalytic cycle. This process of reoxidation of the molybdenum has two steps and in each one an electron leaves the molybdenum. There is an intermediate, **E**, where the metal has the oxidation state +5.

The process starts with the transfer of the first electron to the external electron acceptor. In the real system, this electron acceptor is a sulfur iron box, but, due the difficulty in computing this kind of systems, an organic molecule called benzyl viologen was used, as discussed in section 5.1.2. This first transfer leads to intermediate **E**, containing Mo^V.

Figure 5.10 shows the structures involved in this oxidation process for the system containing selenium. In **E**, due the loss of one electron, the central metal atom binds more strongly to the coordinated selenium. The Mo-Se distance decreases from 2.507 Å in **D** to 2.445 Å in **E**. In turn the Se-H distance between the selenium atom and the protonated imidazole increases from 2.335 Å to 2.544 Å. At the same time, the H-N distance decreases to 1.031 Å.

In the next step of the reoxidation of the molybdenum process, a second electron leaves the system, and the final Mo^{VI} oxidation state is obtained, in intermediate **F**. The departure of the second electron is coupled with the entry of a hydroxyl group, necessary to produce the experimentally observed resting
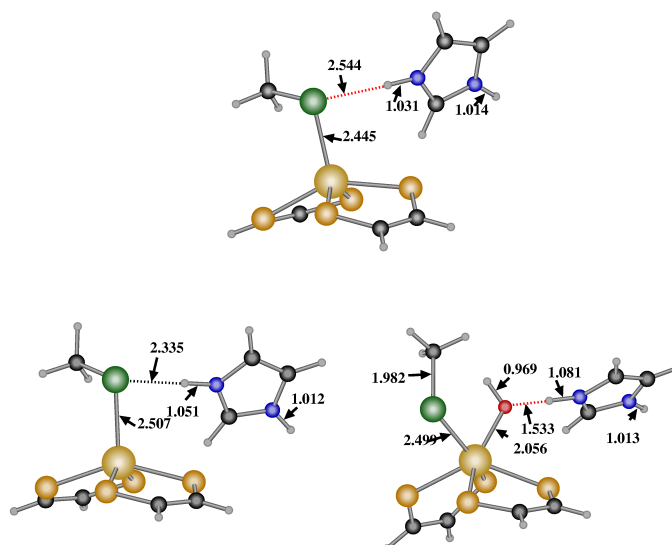
Figure 5.10: Minima for reoxidation of molybdenum the FDHh model with selenium. The structures are: minimum **C** (bottom left corner), minimum **E** (top) and minimum **D** (bottom right corner). Distances are in Å.

state. This hydroxyl group is coordinated to molybdenum, the distance Mo-O is 2.956 Å. It is furthermore stabilized by a hydrogen bond with the protonated histidine (O-H, 1.533 Å), and this induces a new increase of the N-H distance to 1.081 Å. In structure **F** the increase in the metal oxidation state brings a new shortening of the Mo-Se distance, to a value of 2.499 Å already very similar to the distance in **A** (2.497 Å).

As was the case for the other reaction steps discussed in previous sections, there are no important differences in the structures with the selenium model and the sulfur model. The latter are collected in Figure 5.10. All distances around sulfur are a little bit shorter, but there are not any outstanding changes.

The process of reoxidation of the molybdenum center (see Table 5.6) is an endothermic reaction. **F** is 3.4 kcal/mol above **D** in the selenium system and 4.6 kcal/mol in the sulfur system. We are not able to compute the transition states

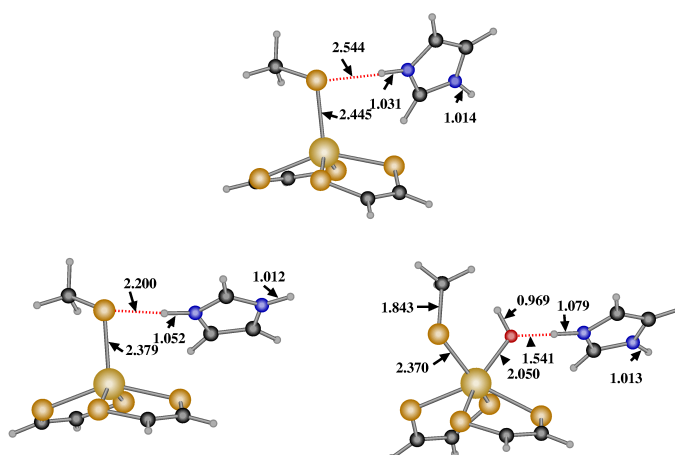## 5.1. QM study of the catalytic mechanism of FDHh

Figure 5.11: Minima for reoxidation of molybdenum the FDHh model with sulfur. The structures are: minimum **C** (bottom left corner), minimum **E** (top) and minimum **D** (bottom right corner). Distances are in Å.

for this step because we are considering neither the real iron sulfur box that acts as electron sink nor the path through which the electron reaches this sink. We will use the energy of intermediate **E** as an estimate for this barrier. **E** has an energy distinguishably higher than the other minima, **D** and **F**. The loss of the first electron (from **D** to **E**) implies an increase of energy of 13.7 kcal/mol for the system with selenium and 19.9 kcal/mol for the system with sulfur. There is thus a sizeable barrier difference favoring the reaction in the system with selenium.

We think that the presence of a significant energy difference of 6.2 kcal/mol in favor of the selenium system suggests that this is the rate-determining step for the overall process in the biological environment. This hypothesis is furthermore strengthened by the fact that this behavior can be easily explained by a qualitative reasoning. Selenium is less electronegative than sulfur, and will

121

Table 5.6: Computed relative energies (kcal/mol) for minima involved in the molybdenum reoxidation step of catalytic cycle of FDHh with both the selenium and sulfur models.

Molecule	ΔE_{Se}	ΔE_S	Difference (Se-S)
D	0	0.0	0.0
E	13.7	19.9	6.2
F	3.4	4.6	1.2

thus have a more covalent bond with molybdenum. Molybdenum will have more electron density when bound to selenium, and the removal of electrons associated to reoxidation will be easier in this case.

5.1.7 Incorporation of a new formate anion

It is the last step of the catalytic cycle that allows to restart the process. The last step discussed above, the reoxidation of the molybdenum, ended in **F**, where molybdenum had recovered the starting oxidation state +6 and a hydroxil anion was coordinated to the enzyme active center.

This last step, from **D** to **A'**, implies the exit of a water molecule and the entry of a new formate anion. The higher affinity of the proton (from the protonated histidine) to the hydroxil molecule generates a water molecule which leaves the system at the same time that a formate is coordinated to molybdenum. Since both **F** and **A** were involved in previous steps, there are no new intermediates involved in this step. Moreover a transition step of this process was not searched because it seems a simple ligand replacement process.

Structures for the selenium system are summarized in Figure 5.12, and for the sulfur system in Figure 5.13. The Mo-Se distance is practically unchanged, 0.002 Å shorter in **A** than in **F**; with sulfur, the Mo-S distance varies by only
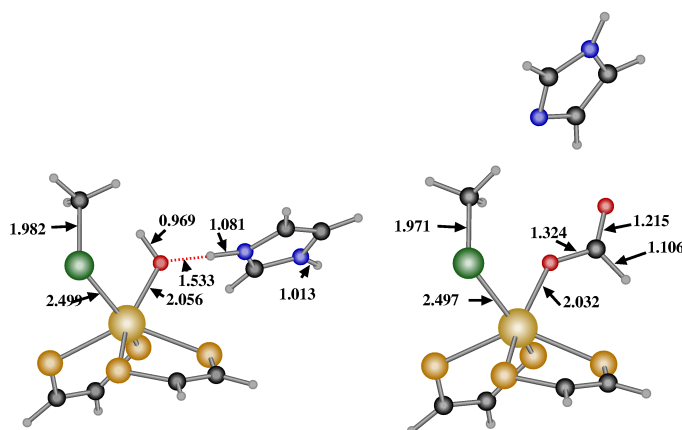
5.1. QM study of the catalytic mechanism of FDHh

Figure 5.12: Minima for incorporation of a new formate anion in the FDHh model with selenium. The structures are: minimum **F** (left), minimum **A** (right). Distances are in Å.

0.008 Å. Besides this negligible change in Mo-Se/S distances, the exit of the water molecule with the incorporation of a new formate provokes a movement of the imidazole ring. This ring had one of its nitrogens involved in a hydrogen bond, either as acceptor or as donor (when protonated), throughout most of the catalytic cycle. This caused a linear Se(S)-H-N arrangement from intermediates **B** to **F**. This hydrogen bond disappears in **A**, and the imidazole is only loosely connected to the system.

The replacement of hydroxyl by formate does not bring important energy changes. The process exothermic with selenium by -3.2 kcal/mol and endothermic with sulfur by 1.5 kcal/mol.

5.1.8 Summary

The geometrical structures of minima and transition states are quite similar for our computed models containing either selenium or sulfur. But the differences in energy barriers are remarkable. For both elements there are two main bar-
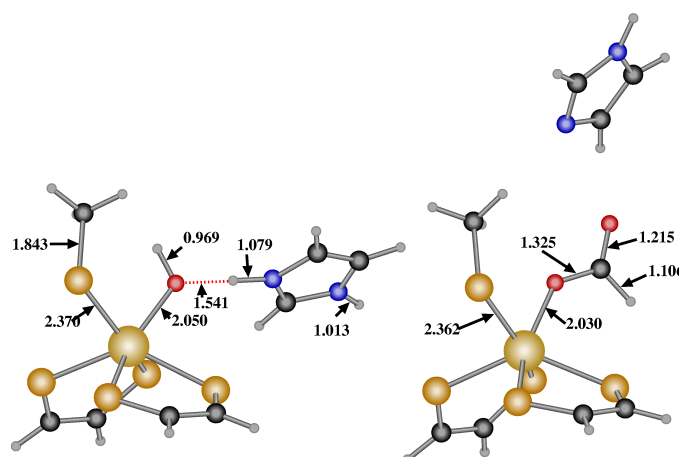
Figure 5.13: Minima for incorporation of a new formate anion in the FDHh model with sulfur. The structures are: minimum **F** (left), minimum **A** (right). Distances are in Å.

riers, related to formate oxidation and molybdenum reoxidation. For the selenium system, the computed barriers are 25.3 and 13.7 kcal/mol, respectively. For the sulfur sytem, they are 22.6 and 19.9 kcal/mol, respectively

The highest barrier is that for the step concerning oxidation of the formate substrate. The most simple explanation would thus say that this is the rate-determining step. However, the reaction is favored (by 2.6 kcal/mol) in the case of sulfur. And this result is in absolute contradiction with the experimental evidence that the reaction is 300 hundred faster with selenium than with sulfur. On the other hand, the reoxidation step is favored by 6.2 kcal/mol for the selenium system. This would fit well with the experimental preference of selenium over sulfur.

Confronted with this discrepancy, we decide to trust more the comparison between our results with selenium and sulfur for a given step than the comparison between our results for both steps with a given calchogen. There are a number of systematic errors in our computational approach that are likely

5.1. QM study of the catalytic mechanism of FDHh

mostly cancelled when replacing the element, but much more difficult to translate to the comparison between two such different steps.

Therefore, we conclude that the rate-determining step ought to be the metal reoxidation. Once this is accepted, the reaction is more favorable with selenocysteine, in agreement with experiment; and the preference can be easily explained in terms of the lower basicity of selenolate with respect to thiolate.

5.2   QM study of the catalytic mechanism of W-FDH

It has been discussed in the introduction chapter (section 1.4.3) that most formate dehydrogenases contain molybdenum in the active center. However, in some cases, this central metal atom can be different. This is the case of the tungsten-containing W-FDH formate dehydrogenase from *Desulfovibrio gigas*. Apart from the replacement of molybdenum by tungsten, the active center of this enzyme is similar to that of the previously discussed FDHh, sharing the same general features of the active center of different formate dehydrogenases. In addition, W-FDH is also a selenoprotein.

The catalytic mechanism of this tungsten-containing W-FDH formate dehydrogenase was computed in order to compare with that computed above for the FDHh from *Escherichia coli* (section 5.1).

5.2.1   Catalytic cycle

In agreement with all experimental observations, we will assume that the catalytic cycle of W-FDH is similar to that of FDHh. The reaction begins with the coordination of a formate anion to the central tungsten atom. Then, this formate is oxidized, transferring a hydrogen to selenium and reducing the central metal from tungsten +6 to tungsten +4. Afterwards, the reoxidation of the metal will recover the functional form of the enzyme to restart the catalytic cycle with a new formate molecule.

The catalytic cycle for W-FDH is shown schematically in Figure 5.14. The only difference with that of FDHh is the absence of intermediate **C**, that will be discussed below. In any case, to make the comparison easier, the same labels were conserved.
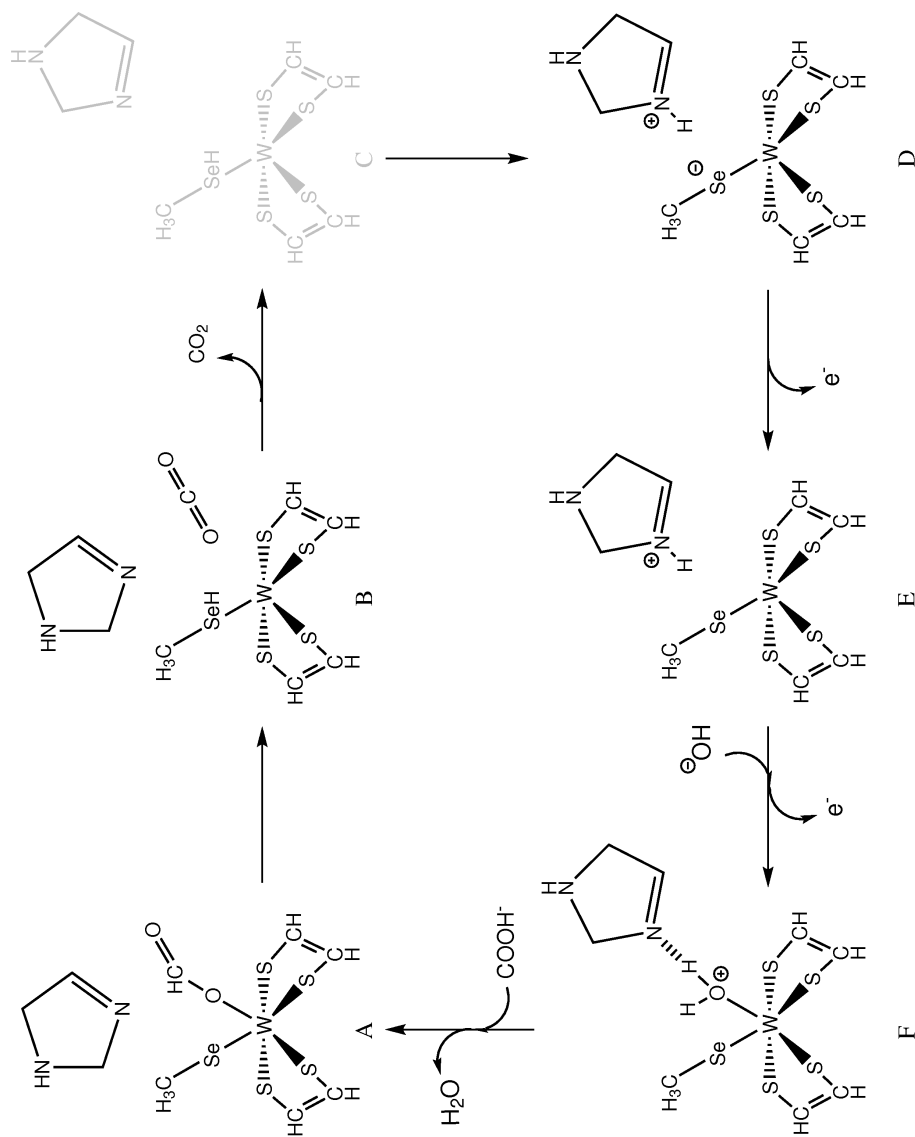
Figure 5.14: Proposed catalytic cycle for the W-FDH enzyme

5.2. QM study of the catalytic mechanism of W-FDH

5.2.2 Computational details

Minima and transition states of the proposed catalytic cycle were computed at QM level using a model of the active center of the enzyme similar to that used previously for FDHh. This model includes the central tungsten atom coordinated by four sulfur atoms from the two MGD's, the selenium from the methylselenolate representing selenocysteine and an oxygen of the formate substrate. The system also includes imidazole ring which acts as proton acceptor. Figure 5.15 shows this model. The starting point for geometry optimizations was taken from the Protein Data Bank (PDB), accession code 1H0H, published by Raaijmakers and co-workers in 2001.[167]
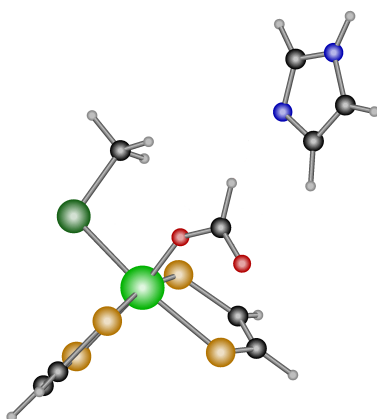
Figure 5.15: QM model for the active center of W-FDH.

All minima and transition states of the proposed catalytic cycle were computed. In each step a geometry optimization was carried out with Becke3LYP density functional[18,19,177] as implemented in the Gaussian03 package[176] using the LANL2DZ basis set[192] for tungsten, selenium and sulfur, and 6-31+G(d) basis set[179,180] for other atoms. The basis set for selenium and sulfur was completed with a d polarization shell. To simulate the effect of the solvent, single point energy calculations were carried out using the PCM method (ϵ=36.64) on

129

each optimized geometry. ϵ close to 40 was used, this solvent was acetonitrile (ϵ=36.64). Again, benzyl viologen was used as electron acceptor.

As for FDHh, to better evaluate the role of selenocysteine, an extra set of calculations was carried out replacing selenium by sulfur. In this way, the role of selenocysteine could be compared with that of cysteine.

5.2.3   Overview of energy profiles

Figure 5.16 presents the energy profile (in kcal/mol) of the computed catalytic cycle (figure 5.14) for W-FDH with selenium. The relative energies of intermediates and transition states are given with respect to intermediate **A**, where the formate anion is coordinated to the tungsten atom. All energies reported are the PCM energies of the gas phase optimized structures.

The only qualitative difference between the computed cycle for W-FDH and that previously reported for FDHh is the absence of intermediate **C** in the system with tungsten. In the FDHh cycle it was possible to compute this intermediate, although it was separated by a transition state with very low barrier from the next species **D**. In the cycle with W-FDHh the suppression of the carbon dioxide molecule (present in the intermediate **B**) leads directly to the transfer of the proton from selenium to imidazole ring representing the histidine aminoacid. The barrier for this step was chemically irrelevant, and the absence of this step has thus a neglectable effect on the overall reactivity of the system.

Apart from this detail, one can see that this profile follows the same pattern of the previously computed system with molybdenum. There are two main barriers in the process, corresponding to the oxidation of the formate (from **A** to **B**) and to the reoxidation of the metal, now tungsten, (from **D** to **F**).

The numeric value of both key barriers changes significantly with respect to the molybdenum system. The barrier for formate oxidation increases to 33.7

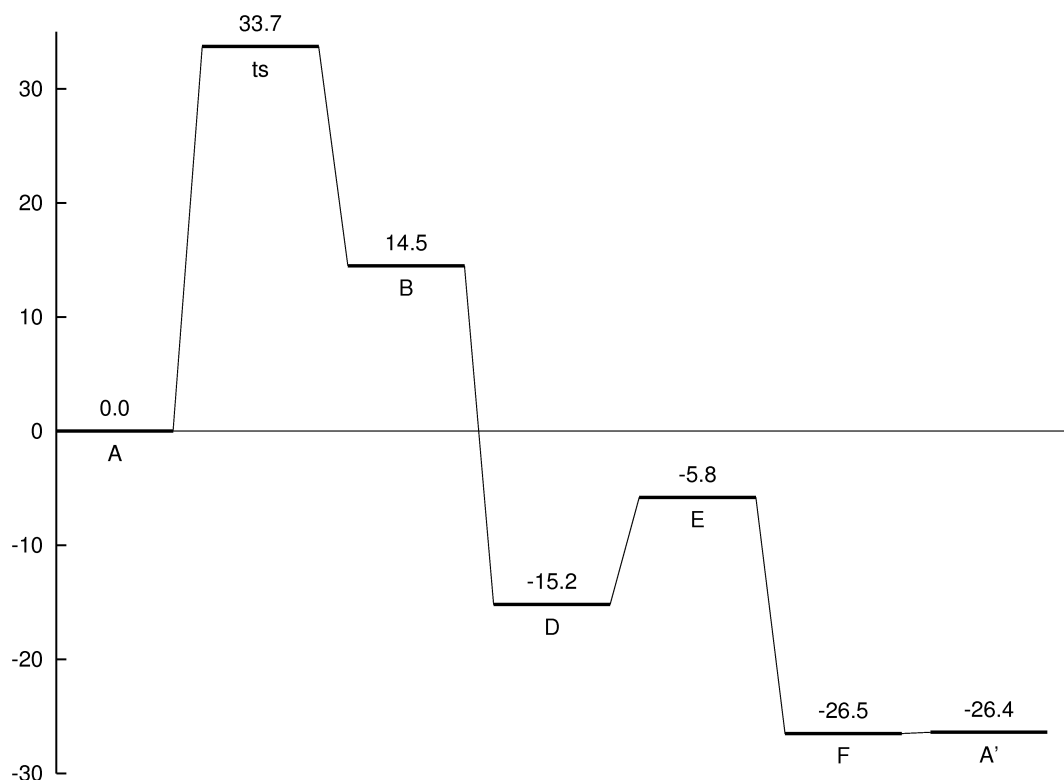## 5.2. QM study of the catalytic mechanism of W-FDH

Figure 5.16: Computed energy profile (in kcal/mol) for the catalytic cycle of W-FDH with the selenium system.

kcal/mol (from 25.3 kcal/mol), and the barrier for reoxidation of the metal (from **D** to **E**) decreases to 9.4 kcal/mol (from 13.7 kcal/mol).

The analogous catalytic cycle for the system where selenium is replaced by sulfur is shown in Figure 5.17. The overall pattern is similar. There are two moderately high barriers for the oxidation of formate (32.0 kcal/mol) and reoxidation of molybdenum (10.2 kcal/mol). The change from selenium to sulfur in W-FDH is the same reported above for FDHh. The barrier for formate oxidation decreases, and that for metal reoxidation increases.

The relative energies of all stationary points presented in this section are collected in table 5.7. The disparity between the barriers for formate oxidation and

131

metal reoxidation is even larger than for the molybdenum systems. The raw data thus point out again to formate oxidation as the rate-determining step. Again, this would suggest cysteine as a more efficient aminoacid than seleno-cysteine. Because of the same reasoning exposed for FDHh on relative accuracy of calculations, we suggest that the rate-determining step is instead the metal reoxidation also for W-FDH.

Table 5.7: Computed relative energies (kcal/mol) for minima and transition states involved in the catalytic cycle of W-FDH with both the selenium and sulfur models.

Molecule	ΔE_{Se}	ΔE_S	Difference (Se-S)
A	0	0.0	0.0
ts	33.7	32.0	1.7
B	14.5	9.0	5.5
D	-15.2	-19.3	4.1
E	-5.8	-9.1	3.3
F	-26.5	-26.4	-0.1
A'	-26.4	-26.4	0.0

Comparison of the energetics for FDHh and W-FDH shows that the barrier for formate oxidation is higher in W-FDH, but the barrier for metal reoxidation is lower. The only explanation we can find therefore for the existence of W-FDH is that it should be more efficient in more reducing environments, where metal reoxidation is intrinsically more difficult. Another aspect of the comparison between both systems concerns their sensitivity to the selenium/sulfur substitution. The difference in barriers is smaller for the tungsten system (1.7 kcal/mol for formate oxidation, 0.8 kcal/mol for metal reoxidation), thus suggesting a

smaller sensitivity to the nature of the aminoacid in the case of tungsten.
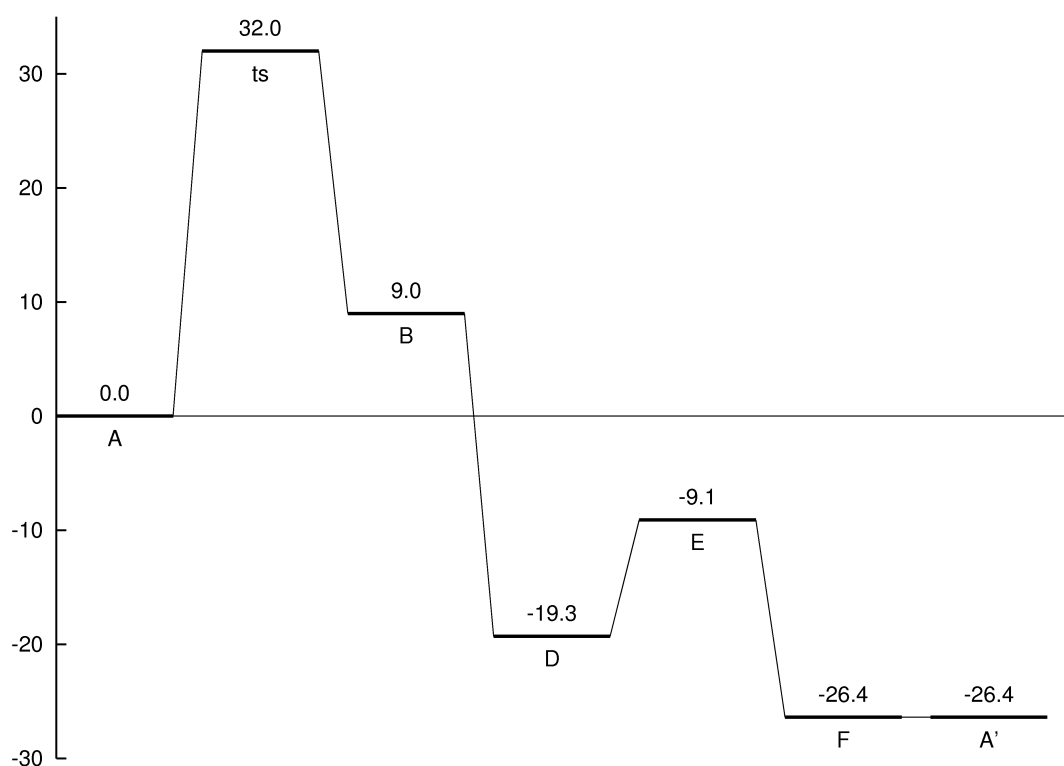
Figure 5.17: Computed energy profile (in kcal/mol) for the catalytic cycle of W-FDH with the sulfur system.

A detailed study of each step will be shown in what follows, and the differential effect of selenium and sulfur in each of them will be analyzed, as well as the difference between molybdenum and tungsten.

5.2.4 Oxidation of formate

The reaction begins in intermediate **A** where the central atom, tungsten, has an oxidation state of +6. Besides the four *cis*-dithiolene sulfurs of MGD cofactors, two more anionic ligands are coordinated to tungsten. They are Me-Se$^-$ and the

133

formate anion (OCHO$^-$). The product of this step, **B**, is the complex between the enzyme and the product of the oxidation of the formate anion, a molecule of carbon dioxide. In the transition state the hydrogen atom the formate ligand is transferred to the selenium center. Carbon dioxide is produced and the tungsten center of the enzyme is reduced to oxidation state +4.

Figure 5.18: Minima and transition states for oxidation of formate in the W-FDH model with selenium. The structures are: minimum **A** (bottom left corner), transition state (top) and minimum **B** (bottom right corner). Distances are in Å.

Figure 5.18 collects the structures involved in this step for the selenium model. The reaction implies an increase of the distance W-Se, from 2.514 Å in **A** to 2.667 Å in **B**. Me-Se-H is obviously a worse ligand than Me-Se$^-$. The coordi-

nation sphere around tungsten changes from an octahedral shape in **A**, typical of hexacordinated transition metals systems, to a square pyramidal shape in **B**. This square pyramid shape is already present in the transition state from **A** to **B**.
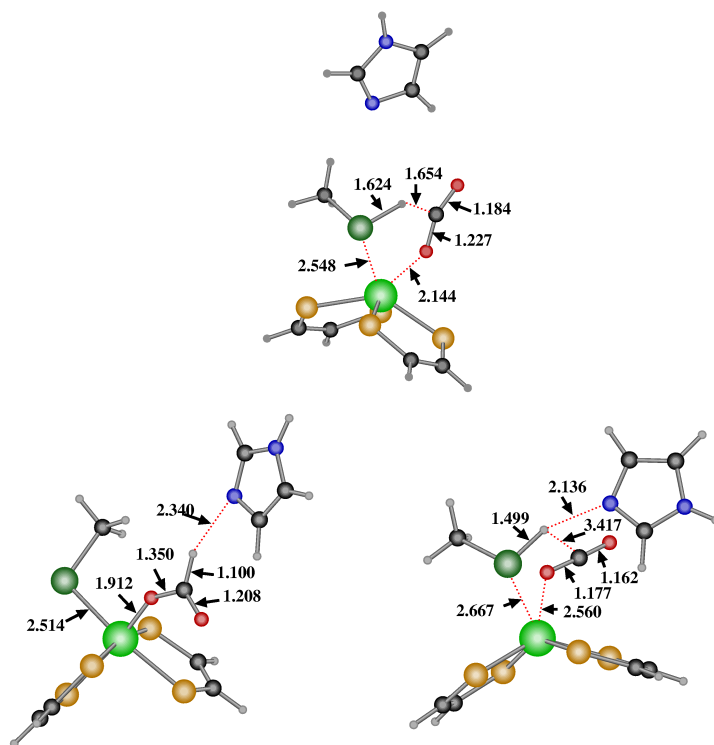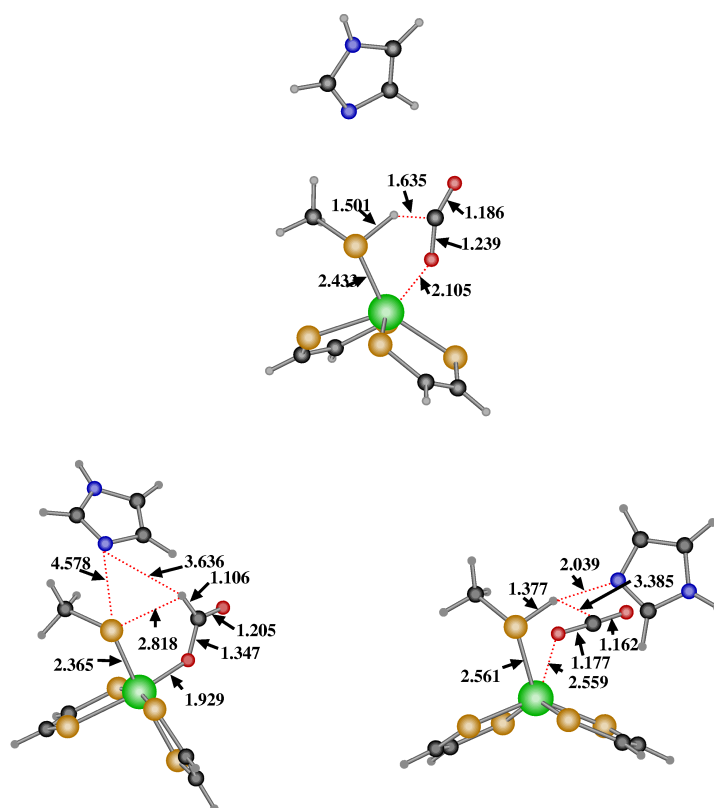

Figure 5.19: Minima and transition states for oxidation of formate in the W-FDH model with sulfur. The structures are: minimum **A** (bottom left corner), transition state (top) and minimum **B** (bottom right corner). Distances are in Å.

The formate anion being oxidized experiences obviously substantial geometrical changes. In **A**, one of its oxygens is coordinated to tungsten, with a W-O distance of 1.912 Å; the carbonyl oxygen is oriented towards one of the

MGD sulfurs, and the hydrogen points towards the basic nitrogen atom of imidazole, with a N-H distance of 2.340 Å. In the transition state the hydrogen is between carbon and selenium with C-H and Se-H distance of 1.654 and 1.624 Å, respectively. In **B** the formate has already been oxidized to carbon dioxide which remains in the active center of the enzyme. The protonated selenium is stabilized by the imidazole ring which has been reoriented to maximize the N-H interaction. In **B**, the Se-H distance is 1.499 Å and the N-H distance is 2.136 Å.

Structures with sulfur follow a similar pattern and one can appreciate the same changes from selenium to sulfur than for FDHh. The increase of the W-Se(S) distance in this reaction step is larger with sulfur than with selenium, being 0.196 Å with sulfur (from 2.365 Å in **A** to 2.561 Å in **B**). A seemingly remarkable feature of the **A** intermediate in this system is in the position of the imidazole ring. The basic nitrogen is at similar distances from the formate hydrogen and the selenium center. Alas, these distances are quite long, 3.636 and 4.578 Å, respectively. Therefore, this is just one of the possible conformations, of likely similar energy for the arrangement between imidazole and the ligands organized around the metal. In the transition state, the hydrogen being transferred is logically closer to sulfur (1.501 Å) than it was to selenium, but the H-C distances are very similar. (1.654 Å for Se; 1.635 Å for S). In the enzyme-product complex, intermediate **B**, the S-H unit is also stabilized by a hydrogen bond with the most basic nitrogen of imidazole, which is placed at 2.039 Å from H (the S-H distance is 1.377 Å).

From an energy point of view, the oxidation of formate in W-FDH is clearly an endothermic reaction. **B** is 14.5 kcal/mol above **A** in the model with selenium; and the difference is 9.0 kcal/mol with sulfur. The barriers follow the same trend as the exothermicities, both in the comparison between selenium

136

and sulfur for W-FDH and in the comparison between FDHh and W-FDH. The difference between sulfur and selenium in W-FDH can be explained by the same reasonings exposed above for FDHh. As for the comparison between FDHh and W-FDH, that is, between molybdenum and tungsten, we can appeal to the different redox potentials, related to electronegativities of the metals. Tungsten has a lower electronegativity, attracts its electrons with less strength, and has thus a lower reduction potential. Oxidation of formate requires reduction of the metal, and tungsten is simply less likely to reduce than molybdenum.

Table 5.8: Computed relative energies (kcal/mol) for the species involved in the step of formate oxidation in W-FDH.

Molecule	ΔE_{Se}	ΔE_S	Difference (Se-S)
A	0.0	0.0	0.0
ts	33.7	32.0	1.7
B	14.5	9.0	5.5

5.2.5   Deprotonation of selenium/sulfur

This step consists in the transfer of the proton from the selenium atom to the imidazole ring representing the histidine aminoacid.

This step has an important qualitative difference with the calculations reported for the molybdenum FDHh system in the previous section. A species as **C**, where the proton stays on selenium after the departure of carbon dioxide does not exist as a local minimum. As a result, only the two structures shown in Figure 5.20 are significant for the W-FDH system. **B** is the product-enzyme complex, with selenium protonated and a W-Se distance of 2.667 Å. There is already

137

a hydrogen bond between the protonated selenium and the basic nitrogen of imidazole, with a N-H distance of 2.136 Å.
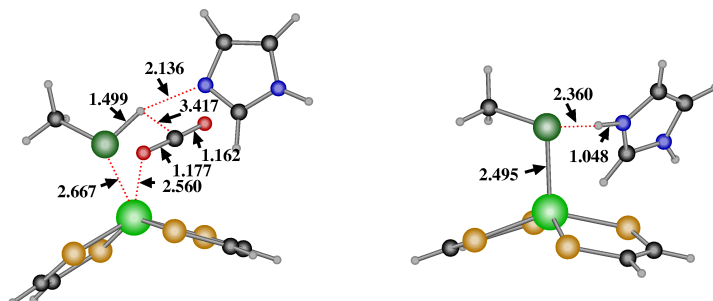

Figure 5.20: Minima involved in deprotonation of selenium in the FDHh model with selenium. The structures are: minimum **B** (left) and minimum **D** (right). Distances are in Å.

When the carbon dioxide is removed from the system, the proton moves spontaneously from selenium to the imidazole nitrogen, resulting in intermediate **D**. This intermediate has a significant charge separation between the two subunits, the positively charged protonated imidazole and the negatively charged metal unit. The W-Se distance is 2.495 Å in this structure, significantly shorter than that in **B**. The loss of the proton increases the affinity of selenium for the metal. Of course, there is also a strong hydrogen bond in **D**. The involved distances are: Se-H, 2.360 Å; N-H, 1.048 Å.

Once again, the structures with the sulfur model (Figure 5.21) follow the same pattern, with all distances involving sulfur being shorter than the corresponding ones with selenium. For instance, in **D** the W-S distance is 2.362 Å and the S-H distance is 2.219 Å.

The step is clearly exothermic: 29.7 kcal/mol for the selenium system, 28.3 kcal/mol for the sulfur system. Following the general trend the sensitivity to the selenium/sulfur replacement is smaller for W-FDH than for FDHh. It

138

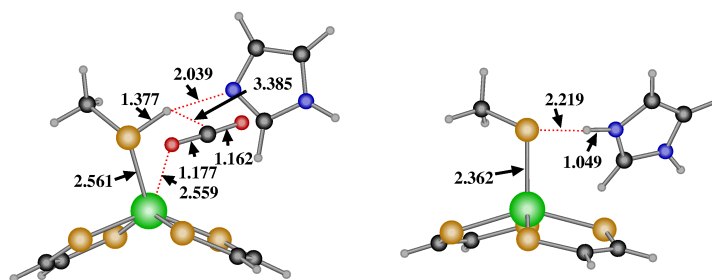5.2. QM study of the catalytic mechanism of W-FDH

Figure 5.21: Minima involved in deprotonation of sulfur in the FDHh model with sulfur. The structures are: minimum **B** (left) and minimum **D** (right). Distances are in Å.

is worth mentioning the substantial effect of the environment in the exothermicity of the step, the gas phase values for selenium and sulfur are 14.2 and 14.3 kcal/mol, respectively. Due to the charge separation in **D**, this structure is strongly favored with respect to **B** by the introduction of a dielectric constant in the environment.

5.2.6 Reoxidation of tungsten

Once the selenium has been deprotonated and this proton transferred to the acceptor, selenium becomes negatively charged while tungsten remains with oxidation state +4. This is the situation in intermediate **D**. Tungsten has to return to the initial oxidation state +6 in order to restart the catalytic cycle. This process of reoxidation of the molybdenum has two steps and in each one an electron leaves the molybdenum. There is an intermediate, **E**, where the metal has the oxidation state +5, and the product of this step, **F**, contains W(VI). The optimized structure for **D**, **E**, **F** are shown in Figure 5.22.

The first part of this tungsten reoxidation step consists of the transfer of the first electron to the external electron receptor, resulting in intermediate **E**. When going from **D** to **E** the W-Se distance decreases by 0.05 Å, indicating that coor-
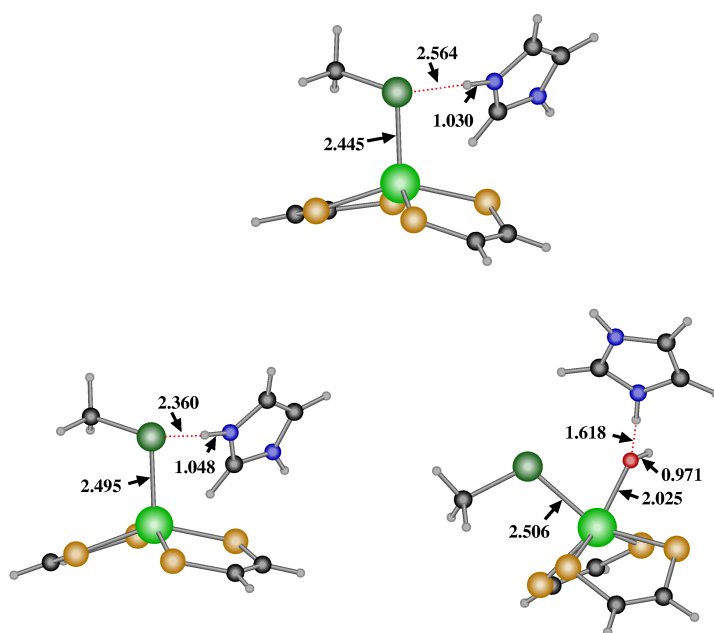
139

Figure 5.22: Minima for reoxidation of tungsten in the W-FDH model with selenium. The structures are: minimum **D** (bottom left corner), minimum **E** (top) and minimum **F** (bottom right corner). Distances are in Å.

dination of the formally negative Me-Se$^-$ ligand is stronger to W^V than to W^{IV}. The distance Se-H increases from 2.360 Å in **D** to 2.564 Å in **E**, indicating a weakening of the hydrogen bond. In the three structures considered, the protonated imidazole system is always involved in hydrogen bonds.

In the step from **E** to **F** a second electron leaves the system, and the oxidation state of tungsten increases from +5 to +6. As happened for FDHh, the departure of the second electron is coupled with the entry of a hydroxyl group.

The most remarkable thing in this step is the entry of a new molecule into the active center of the enzyme. This molecule is an anionic hydroxil which is coordinated to tungsten by its oxygen being the W-O distance, 2.025 Å. This new negative ligand causes a rearrangement of the position of the imidazole ring. The distance H-O is 1.618 Å.

140

5.2. QM study of the catalytic mechanism of W-FDH
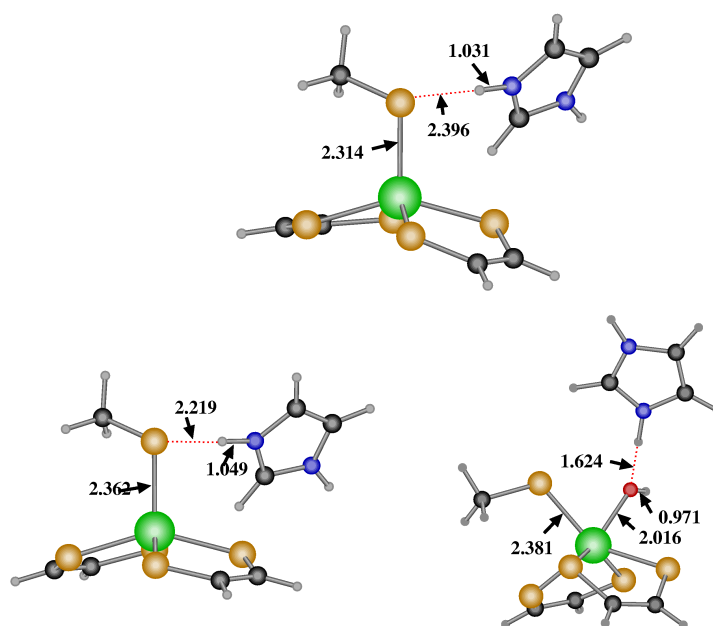
Figure 5.23: Minima for reoxidation of tungsten in the W-FDH model with sulfur. Minimum **D** (bottom left corner), minimum **E** (top) and minimum **F** (bottom right corner). Distances in Å.

Once again, one cannot see important changes in the structure of the intermediates with sulfur (Figure 5.23), apart of the general decrease of the distances involving the sulfur atom.

The tungsten oxidation process from **D** to **F** is an exothermic reaction, by 11.3 kcal/mol with the selenium model, and by 7.1 kcal/mol with the sulfur model. The barrier for the reoxidation process can be estimated as the energy difference from **D** to **E**, as was done in the case of the FDHh. **E** is not a transition state, but we can assume that the electron transfer transition states separating it from other minima will have very low barriers. The barrier for tungsten reoxidation computed in this way is 9.4 kcal/mol for the selenium system and 10.2 kcal/mol for the sulfur system.

As was the case for FDHh, we think that the difference of 0.8 kcal/mol for

141

Table 5.9: Computed relative energies (kcal/mol) for minima and transition states involved in the tungsten reoxidation step of catalytic cycle of W-FDH with both the selenium and sulfur models.

Molecule	ΔE_{Se}	ΔE_S	Difference (Se-S)
D	0.0	0.0	0.0
E	9.4	10.2	0.8
F	-11.3	-7.1	4.2

the system with selenocysteine suggests that this reoxidation is also the rate-determining step for W-FDH. The explanation is again based on the low electronegativity of selenium compared to sulfur. It is also worth mentioning that the barrier for this step is remarkably lower for the tungsten system than for the molybdenum counterpart. This can be explained again by the reduction potential of the metals. Tungsten has a lower reduction potential and it is thus easier to oxidate than molybdenum.

5.2.7 Incorporation of a new formate anion

It is the last step of the catalytic cycle that allows to restart the process. The previous step discussed above, the reoxidation of the tungsten, ended in **F**, where tungsten had recovered the starting oxidation state +6 and a molecule of hydroxyl was coordinated to the enzyme active center. The last step of the catalytic cycle is the first one of a new process. This step completes the enzyme regeneration and a new cycle is started from intermediate **A'** In the transformation of **F** into **A'**, a water molecule leaves the system and a new anionic ligand is coordinated to the metal center.

In intermediate **F** (Figure 5.24, the oxidation state of tungsten is +6. The high affinity of the proton (from the protonated histidine) to the hydroxyl anion

5.2. QM study of the catalytic mechanism of W-FDH

Figure 5.24: Minima for incorporation of a new formate anion in the W-FDH model with selenium. The structures are: minimum **F** (left), minimum **A** (right). Distances are in Å.

generates a water molecule, and from there the water can be replaced by a formate anion. There are minor changes in the position of imidazole, but they have little relevance on the overall energetics. The structures where sulfur replaces selenium (Figure 5.25) show a similar pattern.
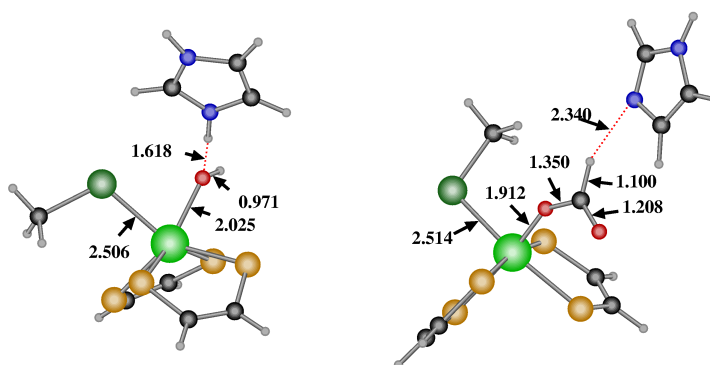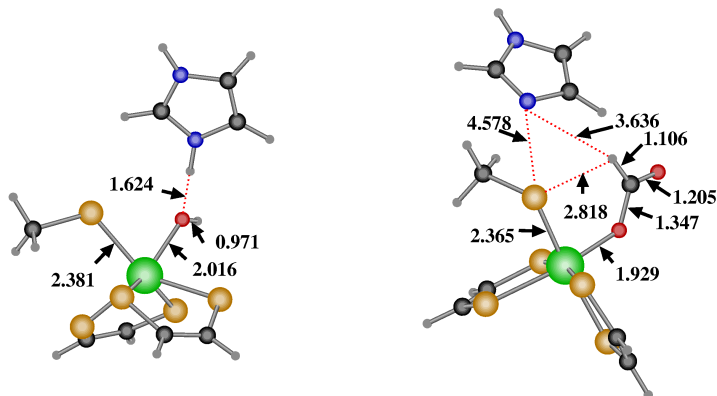

Figure 5.25: Minima for incorporation of a new formate anion in the W-FDH model with sulfur. The structures are: minimum **F** (left), minimum **A** (right). Distances are in Å.

This step has no important energy changes associated, neither for sulfur nor

143

for selenium; neither for molybdenum not for tungsten.

5.2.8 Summary

The main difference between the computed energy profiles for FDHh and W-FDH is that with tungsten, the barrier corresponding to the oxidation of formate is higher, and the barrier corresponding to metal reoxidation is lower. On one hand, this is hardly surprising because high oxidation states are more stable in the case of tungsten, therefore making more difficult for this metal the transition from +6 to +4, and easier the transition from +4 to +6. However, this result can be also related to the characteristics of the bacteria where the two enzymes are present, and ultimately to the biological features of the two metals.

FDHh is present in *Escherichia coli* and W-FDH is present in *Desulfovibrio gigas*. Both *Escherichia coli* and *Desulfovibrio gigas* are mesophilic organisms. This means that they grow better in a medium with moderate temperatures, between $15°C$ and $40°C$. The main difference between them is that the second one is an anaerobic organism, it grows better in absence of oxygen. The absence of oxygen is associated to a more reductive environment, where reductions are easier, and oxidations are more difficult. The presence of tungsten in the enzyme would be thus an advantage only in the case of reducing environments.

In fact, this behavior can be related with the general biological role of molybdenum and tungsten. Both metals are relatively scarce in nature, being in the position 54th and 53rd, respectively, in the ranking of relative abundance of the elements in earth.[195,196] However, their abundance in live organisms is quite different. Molybdenum is present in a wide range of metalloenzymes[197] in bacteria, fungi, algae, animals and plants. In contrast, the presence of tungsten in metalloenzymes is restricted to a limited range of bacteria,[198] many of them anaerobic and hyperthermophilic (living at high temperature) bacteria.[168,199,200]

144

5.2. QM study of the catalytic mechanism of W-FDH

The presence of tungsten-based W-FDH in (*Desulfovibrio gigas*) seems therefore better suited for an anaerobic environment, which was majoritary in earlier stages of earth evolution. The current stage of the planet, with high atmospheric content of oxygen is better suited for molybdenum-based formated dehydrogenases, as W-FDHh.

145

Chapter 6

Conclusions

After the QM and QM/MM studies of the different organic and bioinorganic systems, we have reached the following conclusions.

- The thermal stability of [4]helicenes is directly correlated with the computed interconversion barrier between the M and P forms. This barrier depends on the nature of the substituents. The main effect of the R_1 and R_3 functional groups is steric, while the R_2 substituent has an electronic effect.

- There are two possible paths for the synthesis of [5]helicenequinones from phenantrenes. These two paths, I and II, differ in the conformation of a dihydro[5]helicene intermediate. II is the most stable conformer, and leads to product M. This is the preferred path in absence of steric effects. If there are steric constraints hinder the attack of the oxidating agent, the reaction goes through conformer I, leading to the P enantiomer of the product. This steric effect can come from either the substituents on helicene.

- The enantioselectivity in the ring opening of azlactones by esters catalyzed by guanidine derivatives can be correlated with the computed stabilities

147

of the adducts between guanidinium and the deprotonated azlactone. The structure of these adducts is ruled by the combination of hydrogen bonds and steric effects, and is completely different from those that can be computed for the simplest model systems.

- There are two critical energy barriers in the catalytic cycle of formate dehydrogenase H. The first of them is in the the formate oxidation by the molybdenum-containing catalytic center, and the second one is the reoxidation of the metal. Replacement of cysteine by selenocysteine increases the barrier for the first step, and reduces that of the second one. Metal reoxidation is thus the critical step that requires the presence of selenocysteine instead of cysteine.

- The energy profiles for the catalytic cycles of formate dehydrogenase enzymes containing molybdenum and tungsten follow the same pattern. Tungsten reduces the differences between the behaviors of selenocysteine and cysteine, and seems to be better suited for more reducing environments.

Bibliography

[1] Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003,** *32,* 425–443.

[2] Pons, T.; Montero, L. A.; Febles, J. P. *PLoS. Comput. Biol.* **2007,** *3,* e227.

[3] Szabo, A.; Ostlund, N. *Modern Quantum Chemistry;* MCGraw-Hill: New York, revised 1st edition ed.; 1989.

[4] Jensen, F. *Introduction to Computational Chemistry;* Wiley: New York, 1999.

[5] Koch, W.; Holthausen, M. *A Chemist's Guide to Density Functional Theory;* Wiley-VCH: Weinheim, 2000.

[6] Burkert, U.; Allinger, N. *Molecular Mechanics;* ACS: Washington, 1982.

[7] Rappé, A.; Casewit, C. *Molecular Mechanisc Across Chemistry;* University Science Books: Sausalito, California, 1997.

[8] Karplus, M.; Warshel, A. *J. Am. Chem. Soc.* **1972,** *94,* 5612.

[9] Field, M. J.; Bash, A.; Karplus, M. *J. Comput. Chem* **1990,** *11,* 700.

[10] Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995,** *16,* 1170.

[11] Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996,** *100,* 10580.

[12] Gao, J. *Acc. Chem. Res.* **1996,** *29,* 298.

[13] Monard, G.; Merz, K. M. *Acc. Chem. Res.* **1999,** *32,* 904.

[14] Hartree, D. R. *Proc. Cambridge Phil. Soc.* **1928,** *24,* 89-110.

[15] Fock, V. Z. *Phys.* **1930,** *61,* 126-148.

[16] Levine, I. *Quantum Chemistry;* Prentice Hall: 1999.

[17] Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory;* Wiley-VCH: Weinheim, 2001.

[18] Becke, A. D. *J. Chem. Phys.* **1993,** *98,* 5648.

[19] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev.* **1988,** *37,* B785.

[20] Niu, S.; Hall, M. B. *Chem. Rev.* **2000,** *100,* 353–406.

[21] Torrent, M.; Solà, M.; Frenking, G. *Chem. Rev.* **2000,** *100,* 439–494.

[22] Siegbahn, P. E. M. *Q Rev Biophys* **2003,** *36,* 91–145.

[23] Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976,** *103,* 227–249.

[24] Eichler, U.; Kolmel, C. M.; Sauer, J. *J. Comput. Chem.* **1997,** *18,* 463–477.

[25] Sierka, M.; Sauer, J. *J. Chem. Phys.* **2000,** *112,* 6983–6996.

[26] Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996,** *100,* 19357–19363.

[27] Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003,** *24,* 760–769.

[28] Morokuma, K.; Wang, Q. F.; Vreven, T. *J. Chem. Theory. Computat.* **2006,** *2,* 1317–1324.

150

BIBLIOGRAPHY

[29] Lin, H.; Truhlar, D. G. *J. Phys. Chem. A* **2005,** *109,* 3991–4004.

[30] Antes, I.; Thiel, W. *J. Phys. Chem. A* **1999,** *103,* 9290–9295.

[31] Prabhakar, R.; Musaev, D. G.; Khavrutskii, I. V.; Morokuma, K. *J. Phys. Chem. B.* **2004,** *108,* 12643.

[32] Christmann, U.; Pantazis, D. A.; Benet-Buchholz, J.; McGrady, J. E.; Maseras, F.; Vilar, R. *J. Am. Chem. Soc.* **2006,** *128,* 6376–6390.

[33] Adams, R. D.; Captain, B.; Beddie, C.; Hall, M. B. *J. Am. Chem. Soc.* **2007,** *129,* 986–1000.

[34] Maresca, O.; Maseras, F.; Lledos, A. *New J. Chem.* **2004,** *28,* 625–630.

[35] Solans-Monfort, X.; Sodupe, M.; Branchadell, V.; Sauer, J.; Orlando, R.; Ugliengo, P. *J.Phys. Chem. B* **2005,** *109,* 3539–3545.

[36] Segarra, A. M.; Daura-Oller, E.; Claver, C.; Poblet, J. M.; Bo, C.; Fernández, E. *Chem. Eur. J.* **2004,** *10,* 6456–6467.

[37] Daura-Oller, E.; Segarra, A. M.; Poblet, J. M.; Claver, C.; Fernńdez, E.; Bo, C. *J. Org. Chem.* **2004,** *69,* 2669–2680.

[38] Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *Journal Of Molecular Structure-Theochem* **1999,** *461,* 1–21.

[39] Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Computat.* **2006,** *2,* 815–826.

[40] Horton, E. W. *Chem. Soc. Rev.* **1975,** *4,* 589–600.

[41] Christopherson, R. I.; Lyons, S. D.; Wilson, P. K. *Acc. Chem. Res.* **2002,** *35,* 961–971.

[42] Gallou, I.; Senanayake, C. H. *Chem. Rev.* **2006,** *106,* 2843–2874.

[43] Butler, A. R.; Wu, Y. L. *Chem. Soc. Rev.* **1992,** *21,* 85–90.

[44] Tomaszewski, J.; Rumore, M. M. *Drug. Dev. Ind. Pharm.* **1994,** *20,* 119–139.

[45] Thayer, A. M. *Chem. & Engin. News* **2006,** *84,* 29–31.

[46] Peri, R.; Padmanabhan, S.; Rutledge, A.; Singh, S.; Triggle, D. J. *J. Med. Chem.* **2000,** *43,* 2906–2914.

[47] Kurpiewski, M. R.; Koziolkiewicz, M.; Wilk, A.; Stec, W. J.; JenJacobson, L. *Biochemistry* **1996,** *35,* 8846–8854.

[48] Kumar, C. V.; Buranaprapuk, A.; Sze, H. C. *Chem. Commun.* **2001,** 297–298.

[49] Nakagawa, H.; Kobori, Y.; Yoshida, M.; Yamada, K. *Chem. Commun.* **2001,** 2692–2693.

[50] Cornelissen, J. J. L. M.; Rowan, A. E.; Nolte, R. J. M.; Sommerdijk, N. A. J. M. *Chem. Rev.* **2001,** *101,* 4039–4070.

[51] Ruslim, C.; Ichimura, K. *J. Mat. Chem.* **2002,** *12,* 3377–3379.

[52] Green, M. M.; Cheon, K. S.; Yang, S. Y.; Park, J. W.; Swansburg, S.; Liu, W. *Acc. Chem. Res.* **2001,** *34,* 672–680.

[53] Koeckelberghs, G.; Vangheluwe, M.; Samyn, C.; Persoons, A.; Verbiest, T. *Macromolecules* **2005,** *38,* 5554–5559.

[54] Zech, G.; Kunz, H. *Chem. Eur. J.* **2004,** *10,* 4136–4149.

[55] Xiong, R. G.; Zuo, J. L.; You, X. Z.; Abrahams, B. F.; Bai, Z. P.; Che, C. M.; Fun, H. K. *Chem. Commun.* **2000,** 2061–2062.

BIBLIOGRAPHY

[56] Davis, F. A.; Chen, B. C. *Chem. Rev.* **1992,** *92,* 919–934.

[57] Adam, W.; Korb, M. N.; Roschmann, K. J.; Saha-Moller, C. R. *Journal Of Organic Chemistry* **1998,** *63,* 3423–3428.

[58] Masuno, M. N.; Young, D. M.; Hoepker, A. C.; Skepper, C. K.; Molinski, T. F. *Journal Of Organic Chemistry* **2005,** *70,* 4162–4165.

[59] Hata, S.; Iguchi, M.; Iwasawa, T.; Yamada, K.; Tomioka, K. *Org. Lett.* **2004,** *6,* 1721–1723.

[60] Jacobsen, E.; Pfaltz, A.; Yamamoto, H. *Comprehensive Asymmetric Catalysis;* Springer: 1999.

[61] Knowles, W. S. *Angew. Chem. Int. Ed. Engl.* **2002,** *41,* 1999–2007.

[62] Sharpless, K. B. *Angew. Chem. Int. Ed. Engl.* **2002,** *41,* 2024–2032.

[63] Noyori, R. *Angew. Chem. Int. Ed. Engl.* **2002,** *41,* 2008–2022.

[64] Britni, I.; Sigel, A.; Sigel, H. *Handbook on Metalloproteins;* Marcel Decker: 2001.

[65] Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000,** *100,* 421.

[66] Himo, F.; Siegbahn, P. E. . M. *Chem. Rev.* **2003,** *103,* 2421–2456.

[67] Stepens, P. J.; Jollie, D. R.; Warshel, A. *Chem. Rev.* **1996,** *96,* 2491.

[68] Frenking, G.; Holthausen, M. C. *J. Comput. Chem.* **2006,** *27,* 1221.

[69] Shaik, S.; Kumar, D.; deVisser, S.; Altun, A.; Thiel, W. *Chem. Rev.* **2005,** *105,* 2279.

[70] Bathelt, C. M.; Zurek, J.; Mulholland, A.; Harvey, J. N. *J. Am. Chem. Soc.* **2005,** *127,* 12900–12908.

[71] Blomberg, M.; Siegbahn, P. E. M. *J. Comput. Chem.* **2006,** *27,* 1373.

[72] Blomberg, M.; Siegbahn, P.; Babckock, G.; Wikstrom, M. *J. Am. Chem. Soc.* **2000,** *122,* 12848.

[73] Prabhakar, R.; Morokuma, K.; Musaev, D. G. *J. Comput. Chem.* **2006,** *27,* 1438.

[74] Baik, M.; Newcomb, M.; Friesner, R.; Lippard, S. *Chem. Rev.* **2003,** *103,* 2385.

[75] Pardo, A.; DeLacey, A.; Fernández, V.; Fan, H.; Fan, Y. *J. Biol. Inorg. Chem.* **2006,** *11,* 286.

[76] Tye, J. W.; Darensbourg, M. Y.; Hall, M. B. *J. Comput. Chem.* **2006,** *27,* 1454.

[77] Szilagyi, R. K.; Musaev, D. G.; Morokuma, K. *Inorg. Chem.* **2001,** *40,* 766.

[78] Himo, F.; Eriksson, L.; Maseras, F.; Siegbahn, P. *J. Am. Chem. Soc.* **2000,** *122,* 831.

[79] Nakashima, H.; Hasegawa, J.; Nakatsuji, H. *J. Comput. Chem.* **2006,** *27,* 1363.

[80] Marechal, J. D.; Maseras, F.; Lledos, A.; Mouawad, L.; Perahia, D. *J. Comput. Chem.* **2006,** *27,* 1446.

[81] Comba, P.; A., L.; Maseras, F. *Inorg. Chim. Acta* **2001,** *324,* 21.

[82] Prabhakar, R.; Vreven, T.; Morokuma, K.; Musaev, D. G. *Biochemistry* **2005,** *44,* 11864.

154

BIBLIOGRAPHY

[83] Wynberg, H. *Acc. Chem. Res.* **1971,** *4,* 65-73.

[84] Martin, R. H. *Angew. Chem. Int. Ed. Engl.* **1974,** *86,* 727-738.

[85] Katz, T. J. *Angew. Chem. Int. Ed. Engl.* **2000,** *112,* 1997-1999.

[86] Collins, S. K.; Vachon, M. *Org. Biomol. Chem.* **2006,** *4,* 2518–2524.

[87] Trotter, J. *Acta Crystallogr.* **1963,** *16,* 605-609.

[88] Munday, R.; Sutherland, I. O. *J. Chem. Soc. B* **1968,** 80-84.

[89] Cosmo, R.; Hambley, T. W.; Sternhell, S. *J. Org. Chem.* **1987,** *52,* 3119-3123.

[90] Kiefl, C.; Zinner, H.; Cuyegkeng, M. A.; Eigspelger, A. *Tetrahedron: Assymetry* **2000,** *11,* 3503-3513.

[91] Janke, R. H.; Haufe, G.; Würthwein, E. U.; Borkent, J. H. *J. Am. Chem. Soc.* **1996,** *118,* 6031-6035.

[92] Katzenelson, O.; Edelstein, J.; Avnir, D. *Tetrahedron: Assymetry* **2000,** *11,* 3503-3513.

[93] Cahn, R. S.; Ingold, C.; V., P. *Angew. Chem. Int. Ed. Engl.* **1966,** *78,* 413-447.

[94] Newman, M. S.; Wheatley, W. B. *J. Am. Chem. Soc.* **1948,** *70,* 1913-1916.

[95] Newman, M. S.; Wise, R. M. *J. Am. Chem. Soc.* **1956,** *78,* 450-454.

[96] Okubo, H.; Yamaguchi, M.; Kabuto, C. *J. Org. Chem.* **1998,** *63,* 9500-9509.

[97] Feng, F.; Miyashita, T.; Okubo, H.; Yamaguchi, M. *J. Am. Chem. Soc.* **1998,** *120,* 10166-10170.

[98] Okubo, H.; Feng, F.; Nakato, D.; Hirata, T.; Yamaguchi, M.; Miyashita, T. *Tetrahedron* **1999,** *55,* 14855-14864.

[99] Nakato, D.; Hirano, R.; Yamaguchi, M.; Kabuto, C. *Tetrahedron Lett.* **2003,** *44,* 3683-3686.

[100] Carreño, M. C.; García-Cerrada, S.; Sanz-Cuesta, M. J.; Urbano, A. *Chem. Commun.* **2001,** 1452-1453.

[101] Carreño, M. C.; García-Ruano, J. L.; Urbano, A. *Synthesis* **1992,** 651-653.

[102] Carreño, M. C.; Hernández-Sánchez, R.; Mahugo, J.; Urbano, A. *J. Org. Chem.* **1999,** *64,* 1387-1390.

[103] Carreño, M. C.; García-Cerrada, S.; Urbano, A. *Chem. Eur. J.* **2003,** *9,* 4118-4131.

[104] Terada, M.; Ube, H.; Yaguchi, Y. *J. Am. Chem. Soc.* **2006,** *128,* 1454–1455.

[105] Terada, M.; Nakano, M.; Ube, H. *J. Am. Chem. Soc.* **2006,** *128,* 16044–16045.

[106] Schug, K. A.; Lindner, W. *Chem. Rev.* **2005,** *105,* 67–114.

[107] Luo, R.; David, L.; Hung, H.; Devaney, J.; Gilson, M. K. *J. Phys. Chem. B* **1999,** *103,* 727-736.

[108] Echavarren, A.; Galán, A.; Lehyn, J. M.; de Mendoza, J. *J. Am. Chem. Soc* **1989,** *111,* 4994-5995.

[109] de Mendoza, J. "Personal communication", .

[110] Arai, S.; Takahashi, F.; Tsuji, R.; Nishida, A. *Heterocycles* **2006,** *67,* 495-501.

[111] Chowdari, N. S.; Barbas, C. F. *Org. Lett.* **2005,** *7,* 867-870.

[112] Wood, M. E.; Penny, M. J.; Dteere, J. S.; Horton, P. N.; Light, M. E.; Hursthouse, M. B. *Chem. Commun.* **2006,** 2983-2985.

156

BIBLIOGRAPHY

[113] Karle, I. L.; Kaul, R.; Rao, R. B.; Raghothama, S.; Balaram, P. *J. Am. Chem. Soc.* **1997,** *119,* 12048-12054.

[114] Berzelius, J. J. *Afhandl. Fys. Kemi. Mineralog.* **1818,** *6,* 42.

[115] Schwarz, K.; Foltz, C. M. *J. Am. Chem. Soc.* **1957,** *79,* 3292.

[116] Andreesen, J. R.; Ljungdahl, L. G. *J. Bacteriol.* **1973,** *116,* 867–873.

[117] Turner, D. C.; Stadtman, T. C. *Arch. Biochem. Biophys.* **1973,** *154,* 366–381.

[118] Andreesen, J. R.; Ljungdahl, L. G. *J. Bacteriol.* **1974,** *120,* 6–14.

[119] Ljungdahl, L. G.; Andreesen, J. R. *FEBS Lett.* **1975,** *54,* 279–282.

[120] Ljungdahl, L. G.; Andreesen, J. R. *Methods Enzymol.* **1978,** *53,* 360–372.

[121] Flohé, L.; Günzler, W. A.; Schock, H. H. *FEBS Lett.* **1973,** *32,* 132–134.

[122] Rotruck, J. T.; Pope, A. L.; Ganther, H. E.; Swanson, A. B.; Hafeman, D. G.; Hoekstra, W. G. *Science* **1973,** *179,* 588–590.

[123] Thomson, C. D. *Eur. J. Clin. Nutr.* **2004,** *58,* 391.

[124] Goldhaber, S. B. *Regul. Toxicol. Pharmacol.* **2003,** *38,* 232.

[125] Rayman, M. P. *Lancet* **2000,** *356,* 233–241.

[126] Bardia, A.; Tleyjeh, I. M.; Cerhan, J. R.; Sood, A. K.; Limburg, P.; Erwin, P.; Montori, V. M. *Mayo Clin. Proc.* **2008,** *83,* 23–34.

[127] Brigelius-Flohé, R. *Chem. Biodivers.* **2008,** *5,* 389–395.

[128] Marshall, J. R. *Gastroenterol. Clin. North. Am.* **2008,** *37,* 73–82.

[129] Zeng, H.; Combs, G. F. *J. Nutr. Biochem.* **2008,** *19,* 1–7.

[130] Diwadkar-Navsariwala, V.; Prins, G. S.; Swanson, S. M.; Birch, L. A.; Ray, V. H.; Hedayat, S.; Lantvit, D. L.; Diamond, A. M. *Proc. Natl. Acad. Sci.* **2006,** *103,* 8179.

[131] Ursini, F.; Heim, S.; Kiess, M.; Maiorino, M.; Roveri, A.; Wissing, J.; Flohé, L. *Science* **1999,** *285,* 1393–1396.

[132] Foresta, C.; FlohÃ©, L.; Garolla, A.; Roveri, A.; Ursini, F.; Maiorino, M. *Biol. Reprod.* **2002,** *67,* 967–971.

[133] Arnér, E. S. J.; Holmgren, A. *Semin. Cancer Biol.* **2006,** *16,* 420–426.

[134] Xu, J.; Yang, F.; An, X.; Hu, Q. *J. Agric. Food Chem.* **2007,** *55,* 5349–5353.

[135] Lei, X. G.; Cheng, W. H.; McClung, J. P. *Annu. Rev. Nutr.* **2007,** *27,* 41–61.

[136] Valko, M.; Rhodes, C. J.; Moncol, J.; Izakovic, M.; Mazur, M. *Chem. Biol. Interact.* **2006,** *160,* 1–40.

[137] Battin, E. E.; Perron, N. R.; Brumaghim, J. L. *Inorg. Chem.* **2006,** *45,* 499–501.

[138] Sun, Q. A.; Kirnarsky, L.; Sherman, S.; Gladyshev, V. N. *Proc. Natl. Acad. Sci.* **2001,** *98,* 3673–3678.

[139] Kryukov, G. V.; Castellano, S.; Novoselov, S. V.; Lobanov, A. V.; Zehtab, O.; Guigó, R.; Gladyshev, V. N. *Science* **2003,** *300,* 1439–1443.

[140] Rudolf, E.; Rudolf, K.; Cervinka, M. *Cell. Biol. Toxicol.* **2008,** *24,* 123–141.

[141] Zhao, R.; Domann, F. E.; Zhong, W. *Mol. Cancer. Ther.* **2006,** *5,* 3275–3284.

[142] Chung, H. J.; Yoon, S. I.; Shin, S. H.; Koh, Y. A.; Lee, S.; Lee, Y.; Bae, S. *J. Cell Physiol.* **2006,** *209,* 131–141.

BIBLIOGRAPHY

[143] Fischer, J. L.; Lancia, J. K.; Mathur, A.; Smith, M. L. *Anticancer Res.* **2006,** *26,* 899–904.

[144] Smith, M. L.; Lancia, J. K.; Mercer, T. I.; Ip, C. *Anticancer Res.* **2004,** *24,* 1401–1408.

[145] Jiang, C.; Hu, H.; Malewicz, B.; Wang, Z.; Lü, J. *Mol. Cancer Ther.* **2004,** *3,* 877–884.

[146] Lanfear, J.; Fleming, J.; Wu, L.; Webster, G.; Harrison, P. R. *Carcinogenesis* **1994,** *15,* 1387–1392.

[147] of Medicine, I.; Food,; (USA), N. B. *Dietary Reference Intakes: Vitamin C, Vitamin E, Selenium, and Carotenoids;* 2000.

[148] of Agriculture, U. D.; Service., A. R. *National Nutrient Database for Standard Reference (http://www.nal.usda.gov/fnic/foodcomp);* volume 16 2003.

[149] Kryukov, G. V.; Kryukov, V. M.; Gladyshev, V. N. *J. Biol. Chem.* **1999,** *274,* 33888–33897.

[150] Lescure, A.; Gautheret, D.; Carbon, P.; Krol, A. *J. Biol. Chem.* **1999,** *274,* 38147.

[151] Castellano, S.; Morozova, N.; Morey, M.; Berry, M. J.; Serras, F.; Corominas, M.; Guigó, R. *EMBO Rep.* **2001,** *2,* 697–702.

[152] Boyington, J. C.; Gladyshev, V. N.; Khangulov, S. V.; Stadtman, T. C.; Sun, P. D. *Science* **1997,** *275,* 1305–1308.

[153] Garcin, E.; Vernede, X.; Hatchikian, E. C.; Volbeda, A.; Frey, M.; Fontecilla-Camps, J. C. *Structure* **1999,** *7,* 557–566.

[154] Andreesen, J. R.; Wagner, M.; Sonntag, D.; Kohlstock, M.; Harms, C.; Gursinsky, T.; Jäger, J.; Parther, T.; Kabisch, U.; Gräntzdörffer, A.; Pich, A.; Söhling, B. *Biofactors* **1999,** *10,* 263–270.

[155] Behne, D.; Kyriakopoulos, A.; Meinhold, H.; KÃ¶hrle, J. *Biochem. Biophys. Res. Commun.* **1990,** *173,* 1143–1149.

[156] Tamura, T.; Stadtman, T. C. *Proc. Natl. Acad. Sci.* **1996,** *93,* 1006–1011.

[157] Lee, S. R.; Kim, J. R.; Kwon, K. S.; Yoon, H. W.; Levine, R. L.; Ginsburg, A.; Rhee, S. G. *J. Biol. Chem.* **1999,** *274,* 4722–4734.

[158] Arteel, G. E.; Mostert, V.; Oubrahim, H.; Briviba, K.; Abel, J.; Sies, H. *Biol. Chem.* **1998,** *379,* 1201–1205.

[159] Motsenbocker, M. A.; Tappel, A. L. *J. Nutr.* **1984,** *114,* 279–285.

[160] Kim, H. Y.; Gladyshev, V. N. *Mol. Biol. Cell.* **2004,** *15,* 1055–1064.

[161] Kim, H. Y.; Fomenko, D. E.; Yoon, Y. E.; Gladyshev, V. N. *Biochemistry* **2006,** *45,* 13697–13704.

[162] Unden, G.; Bongaerts, J. *Biochim. Biophys. Acta* **1997,** *1320,* 217–234.

[163] Richardson, D. J. *Microbiology* **2000,** *146 (Pt 3),* 551–571.

[164] Lamzin, V. S.; Aleshin, A. E.; Strokopytov, B. V.; Yukhnevich, M. G.; Popov, V. O.; Harutyunyan, E. H.; Wilson, K. S. *Eur. J. Biochem.* **1992,** *206,* 441–452.

[165] Lamzin, V. S.; Dauter, Z.; Popov, V. O.; Harutyunyan, E. H.; Wilson, K. S. *J. Mol. Biol.* **1994,** *236,* 759–785.

BIBLIOGRAPHY

[166] Jormakka, M.; Törnroth, S.; Abramson, J.; Byrne, B.; Iwata, S. *Acta. Crystallogr. D. Biol. Crystallogr.* **2002,** *58,* 160–162.

[167] Raaijmakers, H.; Teixeira, S.; Dias, J. M.; Almendra, M. J.; Brondino, C. D.; Moura, I.; Moura, J. J.; ao, M. J. R. *J Biol. Inorg. Chem.* **2001,** *6,* 398–404.

[168] Raaijmakers, H.; Macieira, S.; Dias, J. M.; Teixeira, S.; Bursakov, S.; Huber, R.; Moura, J. J.; Moura, I.; ao, M. J. R. *Structure* **2002,** *10,* 1261–1272.

[169] Raaijmakers, H. C. A.; ao, M. J. R. *J. Biol. Inorg. Chem.* **2006,** *11,* 849–854.

[170] Schirwitz, K.; Schmidt, A.; Lamzin, V. S. *Protein Sci.* **2007,** *16,* 1146–1156.

[171] Axley, M. J.; Grahame, D. A. *J. Biol. Chem.* **1991,** *266,* 13731.

[172] Axley, M. J.; Gahame, D. A.; Stadtman, T. C. *J. Biol. Chem.* **1990,** *265,* 18213.

[173] Axley, M. J.; Bock, A.; Stastman, T. C. *Proc. Natl. Acad. Sci.* **1991,** *88,* 8450.

[174] Carreno, M. C.; Enríquez, A.; García-Cerrada, S.; Sanz-Cuesta, M. J.; Urbano, A.; Maseras, F.; Nonell-Canals, A. *Chem. Eur. J.* **2008,** *14,* 603–620.

[175] Dewar, M.; Thiel, W. *J. Am. Chem. Soc.* **1977,** *99,* 4499.

[176] Frisch, M. J. *et al.* "Gaussian 03, Revision C.02", Gaussian, Inc., Wallingford, CT, 2004.

[177] Stephens, P. J.; Devlin, F. J.; Chabalowski, M. J.; Frisch, J. *J. Phys. Chem.* **1994,** *98,* 11623.

[178] "Jaguar 5.5", Schrödinger, Inc., Portland, OR, 2003.

[179] Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys* **1972,** *56,* 2257.

[180] Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973,** *28,* 213.

[181] Wurche, F.; Sicking, W.; Sustmann, R.; Klärner, F.-G.; Rüchardt, C. *Chem. Eur. J.* **2004,** *10,* 2707–2721.

[182] Chan, B.; Radom, L. *J Phys Chem A* **2007,** *111,* 6456–6467.

[183] Noodleman, L. *J. Chem. Phys.* **1981,** *74,* 5737.

[184] Bacskay, G. B. *Chem. Phys.* **1981,** *61,.*

[185] Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard III, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992,** *114,* 10024.

[186] Curtin, D. Y. *Record Chem. Progr.* **1954,** *15,* 111.

[187] "Macromodel 8.5", Schrödinger, Inc., Portland, OR, 2003.

[188] Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989,** *111,* 4379.

[189] Saunders, M.; Houk, K. N.; Wu, Y.; Still, C.; Lipton, M.; Chang, G.; Guida, W. C. *J. Am. Chem. Soc.* **1990,** *112,* 1419.

[190] Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am Chem. Soc.* **1996,** *118,* 11225-11236.

[191] Villà, J.; Warshel, A. *J. Phys. Chem. B* **2001,** *105,* 7887.

[192] Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985,** *82,* 299.

[193] Höllwarth, A.; Böhme, M.; Dapprich, S.; Ehlers, A. W.; Gobbi, A.; Jonas, V.; Köhler, K. F.; Stegmann, R.; Veldkamp, A.; Frenking, G. *Chem. Phys. Lett.* **1993,** *208,* 237.

162

BIBLIOGRAPHY

[194] Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997,** *106,* 5151–5158.

[195] Krauskopf, K. B. *Tungsten. In: handbook of geochemistry;* Springer-Verlag, New York: 1972.

[196] Evans, H. T.; Manheim, F. T.; Landergen, S. *Molybdenum. In: handbook of geochemistry;* Springer Verlag, New York: 1974.

[197] Mendel, R. R. *Dalton Trans* **2005,** 3404–3409.

[198] Kletzin, A.; Adams, M. W. *FEMS Microbiol. Rev.* **1996,** *18,* 5–63.

[199] Yamamoto, I.; Saiki, T.; Liu, S. M.; Ljungdahl, L. G. *J Biol Chem* **1983,** *258,* 1826–1832.

[200] Bertram, P. A.; Karrasch, M.; Schmitz, R. A.; Böcher, R.; Albracht, S. P.; Thauer, R. K. *Eur J Biochem* **1994,** *220,* 477–484.