Chemical activity of anticancer compounds : computational studies on the mechanism of bleomycin and the recognition of flavonoids

Karawajczyk, A.

Chemical Activity of
Anticancer Compounds

Computational studies on the mechanism of bleomycin and the recognition of flavonoids

Proefschrift

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus prof. mr. P.F. van der Heijden,

volgens besluit van het College voor Promoties

te verdedigen op woensdag 31 oktober 2007

klokke 15:00 uur

door

Anna Karawajczyk
geboren te Legnica, Polen
in 1977

Promotiecommissie

Promotor:

Prof. dr. H. J. M. de Groot

Copromotor:

Dr. F. Buda

Referent:

Prof. dr. E. J. Baerends

Overige leden:

Prof. dr. J. Brouwer

Prof. dr. M. C. van Hemert

Prof. dr. G. W. Canters

Contents

List of Abbreviations

aALA	Aminoalanine
ABLM	Activated bleomycin
ADF	Amsterdam density functional
BLM	Bleomycin
BLYP	Becke, Lee, Yang, Parr
BP	Becke, Perdew
CP-ANN	Counterpropagation artificial neural network
CPMD	Car-Parrinello molecular dynamics
DFT	Density functional theory
DNA	Deoxyribonucleic acid
EPR	Electron paramagnetic resonance
Eq.	Equation
FPMD	First principles molecular dynamic
GA	Genetic algorithm
GGA	Generalized gradient approximation
H-bond	Hydrogen bond
HIS	Histidine
HOMO	Highest occupied molecular orbital
K_I	Inhibition constant
LDA	Local density approximation
MD	Molecular dynamics
MM	Molecular mechanics
NMR	Nuclear magnetic resonance
PW	Plane waves
PYR	Pyrimidine
QM	Quantum mechanics
QM/MM	Quantum mechanics/Molecular mechanics
RMSD	Root mean square deviation
RMSE	Root mean square error
Ry	Rydbergs
2D	Two-dimensional
3D	Three-dimensional

Chapter one

INTRODUCTION

1

ABSTRACT

In this introductory chapter the work on the bleomycin anticancer drug is put in the broader context of the drug design process. We discuss the contribution of different computational methods into this field emphasizing the growing role played by quantum mechanical methods.

1.1 Computational chemistry in drug design

The design of pharmaceuticals is an extremely complex process as it is shown on the scheme 1.1 [6]. Computational chemists combine their knowledge of molecular interactions and drug activity, together with visualization techniques, detailed energy calculations, geometric considerations, and data filtered out of huge databases, in an effort to narrow down the search for effective drugs.

A fundamental assumption for rational drug design is that drug activity is obtained through the molecular binding of one molecule, the ligand, to the pocket of another and usually larger molecule, the receptor. In their active or binding conformations, the molecules exhibit geometric and chemical complementarity, both of which are essential for successful drug activity [6, 7]. By binding to macromolecules, drugs may modulate signal pathways, for example, by altering sensitivity to hormonal action, or by altering metabolism, or by interfering with the catalytic activity of the enzyme. Most commonly, this is achieved by binding in a specific cavity of the enzyme, the active site, which catalyzes the reaction, thus preventing access of the natural substrate.

Computer-aided drug design will be a significant component of future rational drug design strategies, and is becoming more relevant as the understanding of molecular activity improves and the amount of available experimental data that requires processing increases [8]. The role of quantum mechanical methods has been until now very limited in the drug design process, mostly because of the high computational demands involved that allowed to deal only with small molecules [9]. The recent progress in first-principles electronic structure calculations together with the steady increase in computational power have considerably broadened the range and scope of application of these theoretical methods. Of particular interest are density functional theory calculations that have been proven to be a powerful tool for studying a large variety of problems in chemistry and more recently in highly complex systems of biophysics and biochemistry.
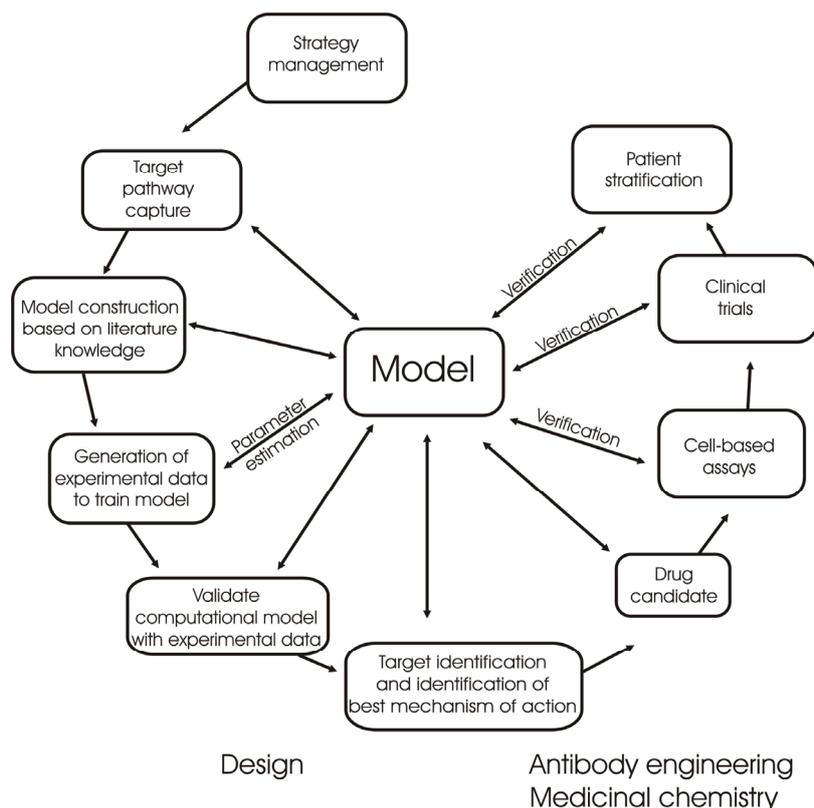
Scheme 1.1 Model-based design in drug discovery. Between each step in the design there is an iterative exchange of information resulting in an improvement of the model.

While computational modeling techniques are increasingly used in all the major steps in the drug design process, they differ in accuracy and amount of molecules they can deal with. In order to define the lead compounds, *i.e.* the compounds that have some activity against a disease, combinatorial chemistry techniques are used. Combinatorial chemistry serves to generate vast libraries of molecules that can be screened for biologically active compounds. However, combinatorial chemistry gives rise to an enormous number and range of compounds that are not necessarily synthetically accessible drug-like molecules. Useful new compounds may not emerge unless intelligently designed in the library production. Sets of compounds (libraries) may be well targeted or diversified, depending on the degree of available information. These libraries can also be compared to database of existing compounds.

An important step to narrow down the numbers of potentially useful compounds is to use molecular modelling to find out the best position, orientation and the

most favorable conformation of a compound based on energy considerations. In addition, the active-analog approach assists the design of a ligand based on similarities to a set of compounds known to possess the desired activity. The affinity score is calculated to select candidate compounds with strong binding to the target site. In addition, when the target site is not known for a ligand, various programs allow to characterize likely sites through computational approaches for functional site mapping. This involves repeatedly placing small functional groups into the possible site to approximate the shape of the binding region. Likely sites may also be inferred from similarity to known site structures. In this step several different computational techniques are used like docking, structure based design or molecular modeling where steric and electrostatic interactions are taken into account.

Another essential step in the process of drug design is to refine the drug activity. For instance, statistical techniques such as QSAR (quantitative structure activity relationship) analysis may be used in order to choose targeted compounds with required features. In QSAR, or QSPR (quantitative structure property relationship), statistical correlation is explored between an activity or a property and geometric or chemical characteristics (pharmacophores) of the molecule. It is often used to analyse the effect of a particular substructure on the activities or properties of compounds. The attributes of the compound being analysed such as the activity, property, or structure are referred to as a descriptor.

Ideally there is a continuous exchange of information between the researchers doing QSAR studies, synthesis and testing. These techniques are frequently used and often very successful since they do not rely on knowing the biological basis of the disease which can be very difficult to determine. However, they are not able to investigate directly the chemical activity of a lead compound or drug interacting with a specific target. Here quantum mechanics based computational tools may become essential for validation of a small number of potential drugs before going into the expensive and time consuming clinical stage.

1.2 DNA-interactive agents

Deoxyribonucleic acid or DNA, the polynucleotide that carries the genetic information in cells, is also one of the receptors with which drugs can interact. Because this receptor is so vital to human functioning, and since from the perspective of a medicinal chemist the overall shape and chemical structure of DNA found in normal and abnormal cells is nearly indistinguishable, DNA-interactive drugs that interact with this receptor are generally very toxic to normal cells. Therefore, these drugs are reserved only for life-threatening diseases such as cancers. There is little information that can guide the design of selective agents against abnormal DNA. One feature differentiating cancer cells from most normal cells is that the cancer cells undergo a rapid, abnormal, and uncontrolled cell division. Genes coding for differentiation in cancer cells appear to be shut off or inadequately expressed, while genes coding for cell proliferation are expressed when they should not be. Because the cells are continuously undergoing mitosis, there is a constant need for rapid production of DNA. Because of the similarity of normal and abnormal DNA, a compound that reacts with a cancer cell will react with a normal cell as well. However, because of the rapid cell division, cancer cell mitosis can be halted more easily than in normal cells where there is sufficient time for repair mechanisms to act. Hence, anticancer drugs are most effective against malignant tumors with a large proportion of rapidly dividing cells, such as leukemias and lymphomas. In addition, DNA damage in a cell is sensed by several as yet poorly defined mechanisms involving a number of proteins. Tumor cells, however, are defective in their ability to undergo cell cycle arrest or apoptosis in response to DNA damage [10]. Cancer cells that cannot undergo cell cycle arrest are sensitive to DNA damaging agents.

There are three major classes of clinically important DNA-interactive drugs: (1) reversible binders, which interact with DNA through the reversible formation of noncovalent interaction; (2) alkylators, which react covalently with DNA bases; (3) DNA strand breakers, which generate reactive radicals that produce cleavage of the polynucleotide strands.

The reversible binders interfere with the interaction of nucleic acids with a variety of small molecules, including water, metal cations, small organic molecules, and proteins, all of which are essential for stabilization of the nucleic acid structure inside the cell [11]. This interference can disrupt the DNA structure. There are three important ways small molecules can reversibly bind to duplex DNA and interfere with DNA function: (1a) by electrostatic binding along the exterior of the helix, (1b) by interaction with the edges of the base pairs in either the major or minor groove and (1c) by intercalation between the base pairs. The *electrostatic interactions* are generally not dependent on the DNA sequence and they are possible due to the negatively charged sugar phosphate backbone. *Groove binders* can be elongated to extend the interaction within the groove, which leads to highly sequence-specific recognition by these molecules. The major and minor grooves have significant differences in their electrostatic potential, hydrogen bonding characteristics, steric effects, and degree of hydration. Proteins exhibit binding specificity primarily through major groove interactions, but small molecules prefer minor groove binding. Finally, flat, generally aromatic or heteroaromatic molecules bind to DNA by *intercalating* between the base pairs of the double helix. The principal driving forces for intercalation are stacking and charge-transfer interactions, while hydrogen bonding and electrostatic forces also play a role in stabilization [12]. Intercalation, first described in 1961 by Lerman [13], is a noncovalent interaction in which the drug is held rigidly perpendicular to the helix axis. This causes the base pairs to separate vertically, thereby distorting the sugar-phosphate backbone and decreasing the pitch of the helix. Intercalation is an energetically favorable process. Presumably, the van der Waals forces that hold the intercalated molecules to the pairs are stronger than the forces stabilizing the stacked pairs.

The last, third, class of DNA-interactive drugs is *DNA strand breakers*. They initially intercalate into DNA, and can react in such a way as to generate radicals depending on the local environmental and cellular metabolism. These radicals typically abstract hydrogen atoms from the DNA sugar-phosphate backbone or from the DNA bases, leading to DNA strand scission. Therefore, these DNA-

interactive compounds are metabolically activated radical generators. Examples of drugs that operate with this mechanism are the anthracycline antitumor antibiotics, tirapazamine and the enediyne antitumor antibiotics, as well as the glycopeptyde antibiotic bleomycin, which is the topic of this thesis.

1.3 Bleomycin

The anticancer drug bleomycin is actually a mixture of several glycopeptyde antibiotics isolated from a strain of the fungus *Streptomyces verticellus*. The major component is bleomycin A_2 ($R = NH(CH_2)_3S^+(CH_3)_2$) (Fig. 1.1). Bleomycin cleaves double-stranded DNA selectively at 5′-GC and 5′-GT sites in the minor groove by a process that is both metal ion and oxygen dependent [14-16]. There are three principal domains in bleomycin [17] (see Fig. 1.1 and Fig. 1.2). The pyrimidine, the β-aminoalanine, and the β-hydroxyimidazole moieties make up the first domain, which is involved in the formation of a stable complex with iron (II). This complex interacts with O_2 to give a ternary complex, which is believed to be responsible for the DNA cleaving activity [18]. The second domain is comprised of the bithiazole moiety (the five-membered N and S heterocycles) and the attached sulfonium ion-containing side-chain. The bithiazole is important for sequence selectivity, presumably because of its intercalation properties with DNA [19]. Possibly the sulfonium ion is attracted electrostatically to a phosphate group [20]. The third domain, consisting of the gulose and carbamoylated mannose disaccharide moiety, may be responsible for selective accumulation of bleomycin in some cancer cells, but it does not appear to be involved in DNA cleavage.

The primary mechanism of action of bleomycin is the generation of single- and double-strand breaks in DNA. This results from the production of radicals by a 1:1:1 ternary complex of bleomycin, Fe(II), and O_2.
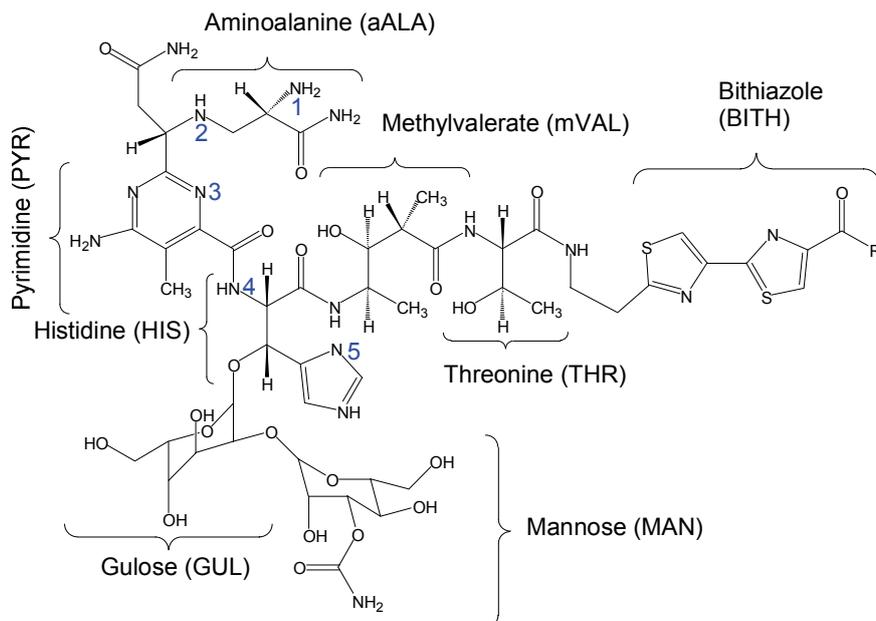
Figure 1.1 The chemical structure of bleomycin A_2 (R = $NH(CH_2)_3S^+(CH_3)_2$), B_2 (R = $NH(CH_2)_4NHC(NH)(NH_2)$), and pepleomycin (R = $NH(CH_2)_3NHCH(CH_3)Ph$). The numbers from 1 – 5 indicate the coordination sites to the metal.

This ternary complex may be self-activated by the transfer of an electron from a second unit of the ternary complex or activation may be initiated by a microsomal NAD(P)H-cytochrome P450 reductase-catalyzed reduction [21, 22]. The activated bleomycin, the peroxide iron (III) bleomycin complex (BLM-Fe(III)-OOH) binds tightly to guanine bases in DNA, principally via the amino-terminal tripeptide containing the bithiazole unit [23]. The two major monomeric products formed when activated bleomycin reacts with DNA are nucleic base propenals (**4** in Scheme 1.2) and nucleic acid bases. Base propenal formation consumes an equivalent of O_2 in addition to that required for bleomycin activation and is accompanied by DNA strand scission with the production of 3′-phosphoglycolate (**5** in Scheme 1.2) and 5′-phosphate-modified DNA fragments (**3** in Scheme 1.2). DNA base formation does not require additional O_2 and results in destabilization of the DNA sugar-phosphate backbone. Evidence for the 4′C radical (**1** in Scheme 1.2) and the peroxy radical (**2** in Scheme 1.2) come from the model studies of Giese and coworkers who used chemical methods to generate a 4′C radical in a single-stranded oligonucleotide [24, 25]. They detected the 4′C radical
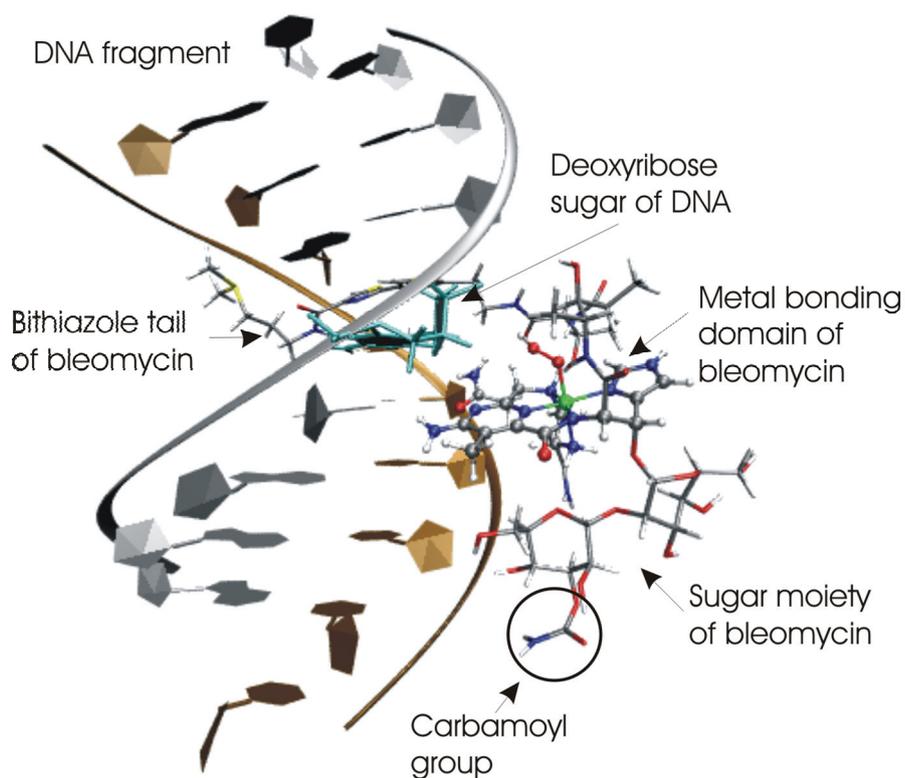
Figure 1.2 The NMR structure of BLM-Co(III)-OOH bound to the fragment of DNA (D(GGAAGCTTCC)) [4]. The structure of the bleomycin complex is represented in color following standard atomic color definition. The three parts of bleomycin are indicated: (i) the metal bonding domain represented in sticks and balls; (ii) the bithiazole tail that is inserted between two pairs of DNA; (iii) the sugar moiety with the carbamoyl group. The deoxyribose sugar of DNA, indicated in cyan in the figure, is the one attacked by activated bleomycin in the first step of DNA degradation.

and the peroxy radical in line with the products that are detected in the reaction of activated bleomycin with DNA.

DNA strand scission is sequence selective, occurring most frequently at 5'-GC-3' and 5'-GT-3' sequence [26]. The specificity for DNA cleavage at a residue located at the 3' side of G appears to be absolute. Preferences for cleavage at 5'-GC and 5'-GT instead of corresponding 5'-AC or 5'-AT sites can be attributed to reduced binding affinity of bleomycin, since guanine can engage in an additional hydrogen bond compared with adenine [27, 28].
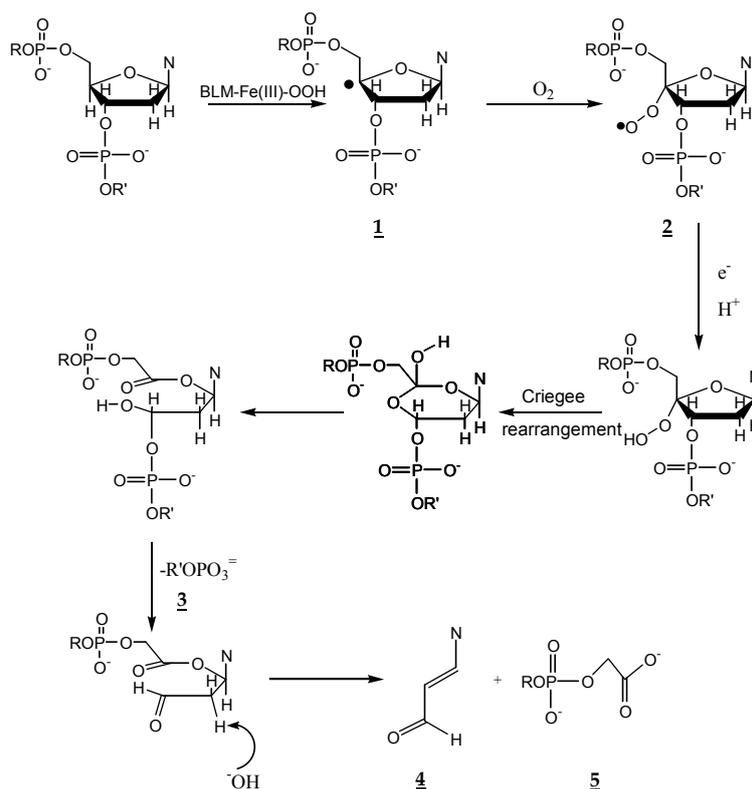
Scheme 1.2 Base propenal formation and DNA strand scission by activated BLM

1.4 Aim and structure of the thesis

In this work the anticancer drug bleomycin is investigated using DFT based computational tools. Although this drug is in use clinically since the early eighties, there are a number of fundamental open issues that are addressed in this thesis work. The investigations are focused on the structure of the Fe(II)BLM complex and the activation mechanism of the O-O bond, which is crucial for the DNA degradation process and for the mechanism of self-inactivation of bleomycin. The work underlines the importance of applying quantum mechanics based methods to the scientific questions in biochemistry. The work shows that the quantum mechanical calculations cannot be avoided at some stage of the drug-activity investigation and that this type of methods is complementary to other methods of lower accuracy. In some cases, like the study of the reactivity, where bond breaking and forming takes place, electronic-structure methods are the only reasonable choice.

In traditional quantum chemistry, highly accurate calculations are carried out for small molecules in vacuum at zero temperature, and molecular properties are deduced from these models. Multidimensional potential energy surfaces can be constructed for very small molecules or for larger systems by considering only a limited set of degrees of freedom, which can be chosen *a priori* by chemical intuition. However, the recent progress in first-principles electronic structure calculations together with a steady increase in computational power have considerably broadened the range and scope of application of quantum chemistry methods. In particular, DFT provides a versatile tool for the study of medium sized to large molecules with a good accuracy (see Section 2.3). On the other hand, the most appropriate method to study reaction pathways is first principles molecular dynamics. Within this approach the system is allowed to evolve at a finite temperature and can possibly cross barriers between minima on the potential energy surface without any *a priori* assumption on the reaction path. An elegant way of carrying out first principles molecular dynamics based on DFT is the Car-Parrinello approach, where the dynamics of the nuclei as well as the adiabatically evolving electronic wave function are described by Newtonian equations of motion (see Section 2.3). When dealing with very large biomolecules it is necessary to use a hybrid quantum mechanics – molecular mechanics approach. Here, I use a recently developed hybrid QM/MM Car-Parrinello scheme [29, 30]. This approach enables efficient and robust hybrid Car-Parrinello simulations of extended systems with the chemically relevant part treated on the quantum mechanical level while the remainder of the system is described with less accuracy in order to simulate the effects due to the environment at a satisfactory level.

The thesis is organized as follows: In Chapter 2, the theoretical foundations of the applied computational methods are introduced. Chapter 3 presents a classical approach to the search for potential drug-like molecules. This chapter provides an example of currently used computational methods in the drug design process and underlines the need of cooperation between experimentalists and theoreticians. In

Chapter 4, I describe results obtained by static DFT calculations applied to the structure of the Fe(II) bleomycin complex. The Car-Parrinello molecular dynamics is used in Chapter 5 for modeling the activation of the O-O bond in the activated bleomycin. In this chapter the QM/MM approach is introduced to study the bleomycin case. Chapter 6 is dedicated to the investigation of the reaction mechanism of activated bleomycin with the deoxyribose sugar since it is known from experiment that the degradation of DNA starts by forming a radical at the 4'C position of a deoxyribose sugar.

Chapter two

COMPUTATIONAL METHODS
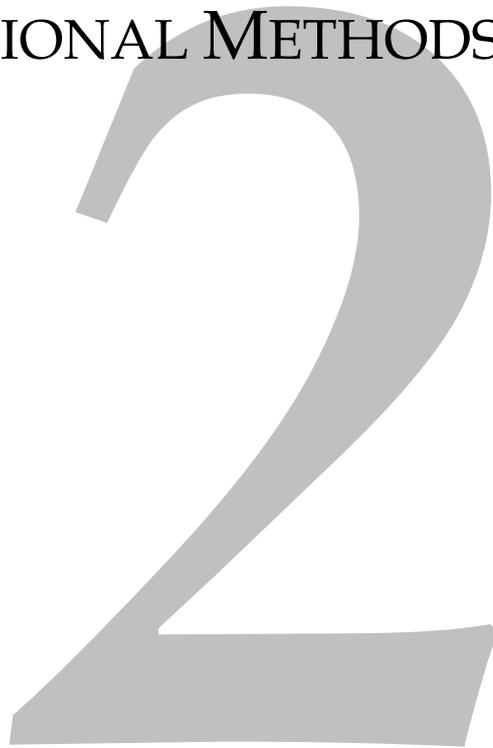
ABSTRACT

First principles molecular dynamics (Car-Parrinello) simulations based on density functional theory have emerged as a powerful tool for studying physical, chemical and biological systems. Its implementation into a QM/MM approach is especially attractive for the *in situ* investigation of chemical reactions that occur in a complex and heterogeneous environment. Hereby, the theoretical backgrounds of all the computational techniques applied in the investigation are presented.

2.1 Classical Molecular Dynamics

Molecular dynamics simulations yield an atomistic time-dependent description of particles in a system and hence provide an insight into its dynamic and thermodynamic properties. From the trajectory, a number of properties such as free-energy differences, reaction rates, and different space and time correlation functions can be calculated.

To describe the time evolution of a molecular system by MD simulations, Newton's equation of motion

$$\vec{F}_i = -\frac{\partial V}{\partial \vec{R}_i} = M_i \frac{d^2 \vec{R}_i}{dt^2} \tag{2.1}$$

has to be integrated. $\vec{F}_i$ is the force acting on atom i with position $\vec{R}_i$ and mass M_i, and V is the potential energy of the system.

In general, there is no analytical solution for the integration of Eq. 2.1, and numerical algorithms based on time discretization have to be used. The size of the time step depends on the characteristic dynamical time scale of the system and for classical MD it is typically between 1–2 fs. A commonly used integration algorithm is the velocity-Verlet algorithm [31], which employs a Taylor expansion truncated beyond the quadratic term for the coordinates

$$\vec{R}(t + \Delta t) = \vec{R}(t) + \vec{v}(t)\Delta t + \frac{\vec{F}(t)}{2M}\Delta t^2. \tag{2.2}$$

The update for the velocities is given by

$$\vec{v}(t + \Delta t) = \vec{v}(t) + \frac{\vec{F}(t + \Delta t) + \vec{F}(t)}{2M}\Delta t. \tag{2.3}$$

The thermodynamic state of a system is defined in terms of macroscopic parameters that are constant during a MD simulation. If the number of particles N, the volume V, and the energy E are fixed, then a constant energy ensemble is sampled. Under the ergodic hypothesis, *i.e.*, the assumption that a system will sample the whole phase space given an infinite amount of time, the time averages over an infinite trajectory correspond to averages over a microcanonical (NVE)

ensemble [32]. However, most chemical and biological processes take place at constant temperature and constant pressure. A canonical ensemble (NVT) or an isothermal-isobaric ensemble (NPT) can be sampled in MD simulations by applying a thermostat and/or a barostat algorithm [33-35].

In classical MD the potential energy is determined by an empirical force field parametrized in order to reproduce experimental or *ab initio* data. Force fields consist of an interaction function and interaction parameters. For biomolecular applications several force fields have been developed [36-38]. Here we use the AMBER8/parm99 [39, 40] force field. Its pair-wise additive potential is of the form

$$V = \sum_{\text{bonds}} k_R \left(\vec{R} - \vec{R}_{\text{eq}} \right)^2 + \sum_{\text{angles}} k_\theta \left(\theta - \theta_{\text{eq}} \right)^2 + \sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi)]$$
$$+ \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\varepsilon R_{ij}} \right].$$

(2.4)

The interactions are divided into bonding interactions, which only act within a molecule, and non-bonding interactions, which act between all atoms with the exception of bonded neighbors. Eq. 2.4 contains three terms for the bonded interactions, namely one for the chemical bonds between two neighboring atoms, one for the bond angles between three atoms, and one for the dihedral angles between four atoms. In addition, improper dihedral-angle terms can be applied to maintain planar or tetrahedral conformations. The functional form of the bond and angle terms is quadratic, while the dihedral term uses a trigonometric function.

The last term in Eq. 2.4 contains the non-bonded interactions, which are composed of a Lennard-Jones term for the van der Waals interactions and a Coulomb term for the electrostatic interactions between atoms *i* and *j*. Van der Waals and Coulomb interactions between atoms that are involved in direct (1-2) or indirect (1-3, 1-4) bonded interactions are rescaled or excluded from the potential. The van der Waals term

$$V_{\text{vdW}} = \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} \right]$$

(2.5)

describes the repulsive forces at small interatomic distances due to the Pauli repulsion between electrons (decaying exponentially, modeled with R_{ij}^{-12}), and the attractive forces at intermediate distances due to instantaneous dipole-induced dipole interactions (decaying with R_{ij}^{-6}).

Finally, the Coulomb term

$$V_{el} = \sum_{i<j} \left[\frac{q_i q_j}{\varepsilon R_{ij}} \right] \tag{2.6}$$

takes into account the electrostatic interactions between charged particles. In the AMBER8/parm99 force field, the charge distribution in a molecule is reproduced by atom-centered point charges derived from the electrostatic potential [41-43]. The Coulomb interaction is a long-range interaction and the sum in Eq. 2.6 converges very slowly. Therefore different algorithms have been developed for a fast and accurate treatment of electrostatic interactions, based on Ewald summations. Particle-mesh Ewald [44] and particle-particle/particle-mesh Ewald [45] algorithms are widely used in classical MD programs.

In classical MD, the size of the system can be of the order of hundred thousand atoms and the total time of simulation is in the nanoseconds range. Since nevertheless one cannot reach macroscopic system sizes for simulations in the condensed phase, usually periodic boundary conditions are used to prevent surface artifacts. The central simulation box is periodically surrounded by images of itself. Care has to be taken to avoid the effects of artificial periodicity, such as the interaction of a molecule with its image in the neighboring box.

The most obvious limitation of empirical force fields is their inability to describe reactive events. Since they make use of fixed parameters, *e.g.*, for bond distances, they cannot adapt to different electronic situations. Additionally, most of the standard force fields employ fixed point charges and do not contain terms that allow for an explicit polarization of the atoms. To be able to describe chemical reactions one has to go beyond an empirical force field description by including explicitly the quantum mechanical description of the electronic structure.

2.2 Density Functional Theory

Among all quantum chemistry techniques, Density Functional Theory is often the method of choice because of its good compromise between accuracy and computational cost. In a few words, DFT provides a way to obtain the electron density and the ground state energy of a polyatomic system given its atomic coordinates [46]. Most programs based on DFT are capable to search for energy minima and compute several molecular properties such as atomic charges, multipole moments, vibrational frequencies, and spectroscopic constants. DFT is also the basis of first principles molecular dynamics techniques such as the Car–Parrinello method, in which the molecules evolve in real time and finite temperature under the forces derived from the instantaneous ground state of the electron cloud [47]. Since the electronic density changes during the simulation, polarization effects are described in a natural way, as well as changes in the bonding pattern of the atoms (*e.g.*, bond breaking and bond-forming processes).

The development of DFT in the area of computational chemistry dates from the mid 1960s when Hohenberg and Kohn [48] demonstrated that the ground-state energy of a system of interacting electrons subject to an external potential $V(\vec{r})$ is a unique functional of the electron density $E = E[\rho(\vec{r})]$, and it can be obtained by minimizing the energy functional with respect to the density,

$$E^{\mathrm{DFT}} = \min_{\rho(\vec{r})} E[\rho(\vec{r})]. \tag{2.7}$$

Later, Kohn and Sham [49] demonstrated that there is equivalence between the electronic density of the real system and a model comprising noninteracting electrons that are subject to an effective potential, V_{eff}. This provides a way to solve the problem of finding the density of the many-electron interacting system, via obtaining the electron density of the noninteracting system. This density can be expressed in terms of single-electron orbitals $\psi_i(\vec{r})$, known as Kohn–Sham (KS) orbitals,

$$\rho(\vec{r}) = 2\sum_{i}^{\mathrm{occ}} \left| \psi_i(\vec{r}) \right|^2, \tag{2.8}$$

where the sum extends over the occupied single-particle orbitals. Eq. 2.8 describes the simplest situation in which all orbitals are doubly occupied and is easily generalized to spin polarized systems [50].

Because of the relation in Eq. 2.8, the energy functional can be either expressed in terms of the density (Eq. 2.7) or using the single-electron orbitals,

$$E^{\mathrm{DFT}} = \min_{\{\psi_i\}} E^{\mathrm{KS}}\left[\{\psi_i(\vec{r})\}, \{\vec{R}_N\}\right], \tag{2.9}$$

where $\{\vec{R}_N\}$ are the nuclear coordinates fixed within the Born-Oppenheimer approximation. The energy functional (Eq. 2.7) can be written in atomic units as:

$$E^{\mathrm{KS}} = 2\sum\int\psi_i^*(\vec{r})\left(-\frac{\nabla^2}{2}\right)\psi_i(\vec{r})d\vec{r} + \int V(\vec{r})\rho(\vec{r})d\vec{r}$$
$$+ \frac{1}{2}\int\frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r}-\vec{r}'|}d\vec{r}d\vec{r}' + E_{\mathrm{XC}}[\rho(\vec{r})] + \sum_{I>J}\frac{Z_I Z_J e^2}{|\vec{R}_I - \vec{R}_J|}. \tag{2.10}$$

The first term on the right-hand side of this expression is the kinetic energy of the noninteracting electrons. The second term corresponds to the interaction of the electrons with the nuclear charges and $V(\vec{r})$ is the potential resulting from the nuclei. In case only valence electrons are explicitly considered in the calculation, $V(\vec{r})$ would be a pseudopotential. The third term corresponds to the classical Coulomb interaction of a density distribution $\rho(\vec{r})$. The fourth term, $E_{\mathrm{XC}}[\rho(\vec{r})]$, is a functional of the density that accounts for the remaining contributions to the electron–electron interaction. The last term accounts for the nucleus-nucleus electrostatic repulsion, which is a constant in the Born-Oppenheimer approximation since the nuclear coordinates are fixed.

The only unknown quantity in Eq. 2.10 is E_{xc}, which contains the exchange and correlation energies. In principle DFT in the Kohn-Sham formulation is an exact theory, but in practice approximations have to be made for the unknown exchange and correlation functional.

In the local density approximation [49], which is based on the homogeneous electron gas approximation, E_{xc} is defined as

$$E_{\mathrm{XC}}^{\mathrm{LDA}}[\rho] = \int\rho(\vec{r})\varepsilon_{\mathrm{XC}}(\rho)d\vec{r}, \tag{2.11}$$

where $\varepsilon_{xc}(\rho)$ is the sum of the exchange and correlation energy per electron of a homogeneous electron gas of density ρ:

$$\varepsilon_{XC}(\rho) = \varepsilon_X(\rho) + \varepsilon_C(\rho), \tag{2.12}$$

with

$$\varepsilon_X(\rho) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \times (\rho)^{1/3} \tag{2.13}$$

and

$$\varepsilon_C(\rho) = \varepsilon_C^{VWN}(\rho). \tag{2.14}$$

For the correlation term ε_c, no analytical expression is available, but accurate values have been obtained from quantum Monte Carlo [51] calculations and analytic forms of ε_c have been parametrized with these results by Vosko, Wilk, and Nusair [52]. LDA yields good results for solid state systems, but the accuracy provided by this approximation is not enough for most applications in chemistry and biology. One of the main drawbacks of LDA is that van der Waals interactions, which originate from correlated motions of electrons caused by Coulomb interactions between distant atoms, cannot be properly described and also bond distances and binding energies can have large errors that appear in a nonsystematic way.

Better results are obtained with functionals that do not only depend on the local density $\rho(\vec{r})$, but also on the local gradient $\nabla\rho(\vec{r})$ of the density. Many functionals have been developed in the framework of the generalized gradient approximation. Among the most popular GGA exchange and correlation functionals used in biological applications are the ones denoted as BP with the exchange part by Becke [53] and the correlation according to Perdew [54], BLYP, a combination of Becke exchange and correlation developed by Lee, Yang and Parr [55], PBE developed by Perdew, Burke and Ernzerhof [56]. The use of the GGA approximation improves considerably the description of bonding, in particularly hydrogen bonding, with respect to pure LDA with a very low additional computational cost. The description of weak van der Waals interactions, however, remains problematic. Therefore, special care should be taken when addressing problems in which van der Waals interactions might play a relevant role, such as

stacking interactions between π-systems and the diffusion of ligands in purely hydrophobic cavities [57]. The introduction of hybrid functionals, which contain a certain amount of exact non-local exchange from Hartree-Fock theory, has been an important step towards a higher accuracy of DFT calculations for molecules. The most popular hybrid functional is the three-parameter functional from Becke (B3LYP) [58]. However, the high computational cost of calculating the two-electron integral of non-local exchange within the plane-wave basis set employed in the Car-Parrinello MD code hampers the use of hybrid functionals for FPMD.

2.3 Car-Parrinello Molecular Dynamics

The basic idea of the Car-Parrinello approach [47] is to treat the electronic degrees of freedom $\{\psi_i\}$ as fictitious classical dynamical variables, exploiting the timescale separation between the fast electronic and the slow ionic motion to avoid energy exchange between the two subsystems. The Car-Parrinello Lagrangian

$$L_{CP} = T_n + T_e - U,$$

or explicitly

$$L_{CP} = \sum_I \frac{1}{2} M_I \, \dot{\vec{R}}_I^2 + \mu \sum_i \int d\vec{r} \left| \dot{\psi}_i \right|^2 - E^{KS}\left[\{\psi_i\}, \{\vec{R}_I\}\right] + \sum_{i,j} \lambda_{ij} \left(\langle \psi_i | \psi_j \rangle - \delta_{ij} \right) \quad (2.15)$$

contains the kinetic energy T_n of the nuclei, the fictitious kinetic energy T_e of the electrons, the potential energy U, and the last term is the constraint ensuring the orthonormality of orbitals. The potential energy term U is given by the Kohn-Sham energy density functional $E^{KS}\left[\{\psi_i\}, \{\vec{R}_I\}\right]$ (Eq. 2.10). The electronic Hamiltonian contains the coulomb interaction between the nuclei. A fictitious mass or inertia parameter μ is assigned to the orbital degrees of freedom and can be tuned to ensure adiabaticity. The Newtonian equations of motion are obtained from the associated Euler-Lagrange equations

$$\frac{d}{dt} \frac{\partial L_{CP}}{\partial \dot{\vec{R}}_I} = \frac{\partial L_{CP}}{\partial \vec{R}_I} \quad (2.16)$$

and

$$\frac{d}{dt}\frac{\delta \mathcal{L}_{CP}}{\delta \langle \dot{\psi_i}|} = \frac{\delta \mathcal{L}_{CP}}{\delta \langle \psi_i|} \qquad (2.17)$$

and the corresponding Car-Parrinello equations of motions are found to be of the form

$$M_I \ddot{\vec{R}}_I = -\nabla_I E^{KS}[\vec{R},\psi] \qquad (2.18)$$

and

$$\mu \ddot{\psi}_i = -\frac{\delta E^{KS}}{\delta \langle \psi_i|} + \frac{\partial}{\partial \psi_i} \sum_{i,j} (\langle \psi_i | \psi_j \rangle - \delta_{ij}). \qquad (2.19)$$

They can be solved numerically using, for example, the Verlet algorithm [59]. The constant of motion is

$$E_{cons} = \sum_I \frac{1}{2} M_I \dot{\vec{R}}_I^2 + \sum \mu \langle \dot{\psi}_i | \dot{\psi}_i \rangle + E^{KS}[\{\psi_i, \vec{R}^N\}]. \qquad (2.20)$$

The parameter μ has to be chosen in a way that ensures (i) that the lowest electronic frequency ω_e^{min} is larger than the highest frequency ω_I^{max} of the nuclei $(\omega_e \gg \omega_I)$, in order to avoid energy transfer, and (ii) that the highest electronic frequency ω_e^{max} is compatible with the chosen time step Δt. Typical values are $\mu =$ 400 − 800 a.u. in combination with a time step $\Delta t = 4 - 6$ a.u. (0.096 – 0.144 fs). The time step in CPMD is smaller than for classical MD because of the fast electronic motions. Since the electronic degrees of freedom are explicitly included, the size of a system that can be treated with FPMD is of the order of 100 – 1000 atoms with a total simulation time in the range of 1 – 10 ps.

The original Car-Parrinello method imposes periodic boundary conditions and expands the wave function in plane waves. Plane waves are defined as

$$f_G^{PW}(\vec{r}) = \frac{1}{\sqrt{\Omega_{cell}}} \exp[i\vec{G}\cdot\vec{r}], \qquad (2.21)$$

with the reciprocal space vector G and the cell volume Ω_{cell}. Plane waves form a complete and orthonormal basis, and the Kohn-Sham orbitals can be written in the form

$$\psi_i(\vec{r}) = \frac{1}{\sqrt{\Omega_{cell}}} \sum_i c_i(\vec{G}) \exp[i\vec{G}\cdot\vec{r}]. \qquad (2.22)$$

The expansion has to be truncated at an energy cutoff $E_{\text{cut}} = \frac{1}{2} G_{\text{max}}^2$, which determines the number of plane waves

$$N_{\text{PW}} = \frac{1}{2\pi^2} \Omega_{\text{cell}} E_{\text{cut}}^{3/2} \qquad (2.23)$$

and therefore the accuracy of the calculation. The advantage of a plane-wave basis set is that all terms of the Car-Parrinello equations can conveniently be solved either in real or in reciprocal space, making use of fast Fourier transformation algorithms. In addition, the Pulay forces are zero, because the basis set does not depend on the atomic positions, which makes evaluation of nuclear forces easier. However, a large number of plane waves would be needed for the description of the highly localized core electrons that are chemically inactive. This problem has prompted the development of pseudopotentials for the description of core electrons. Norm-conserving pseudopotentials [60] depend on the angular-momentum and correctly represent the long-range interactions of the core electrons as well as the full wave function outside the core radius.

2.4 Hybrid QM/MM

The computer simulations of chemical reactions in a realistic environment is a particularly challenging task. Chemical bonds are broken and formed during this process, which implies the use of quantum mechanical methods that can take into account instantaneous changes in the electronic structure explicitly. On the other hand, the systems of interest in computational biology are quite complex with many thousands of atoms. In spite of the considerable progress that was achieved in the development of DFT approaches, it is clear that, in order to treat complex biological systems, we still need to be able to combine various computational chemistry methodologies with different accuracies and cost of calculations. One possible solution for modeling such systems is the choice of a hierarchical hybrid approach in which the whole system is partitioned into a localized chemically active region treated with a quantum mechanical method and the environment,

treated with empirical potentials. In this quantum mechanical/molecular mechanics method the computational effort can be concentrated on the part of the system where it is most needed while the effects of the surroundings, such as mechanical constraints, electrostatic perturbations and dielectric screening, are taken into account with a more expedient model [61]. The idea of a QM/MM scheme is not new and the first published example appeared already thirty years ago [62]. However, in the last few years this subject has developed very rapidly and QM/MM approaches have been implemented in the most commonly used computational packages.

The particular QM/MM Car-Parrinello method [30] that has been used in this work is based on a mixed Hamiltonian of the form

$$H = H_{QM} + H_{MM} + H_{QM/MM}, \tag{2.24}$$

in which the quantum part H_{QM} is described with the extended Car-Parrinello Lagrangian (Eq. 2.15). Since this QM/MM Car-Parrinello implementation establishes an interface between the Car-Parrinello code CPMD [63] and the classical force fields GROMOS96 [38] and AMBER [39], the classical part H_{MM} follows the formalism used in these packages according to Eq. 2.4.

The intricacies of QM/MM methods are in the challenge of finding an appropriate treatment for the coupling between the QM and MM regions as described by the interaction Hamiltonian $H_{QM/MM}$. Special care has to be taken that the QM/MM interface is treated in an accurate and consistent way, in particular in combination with a plane-wave-based Car-Parrinello scheme. If the QM/MM boundary cuts through a covalent bond, care has to be taken to saturate the valence orbitals of the QM system. In the present implementation [30], this can be done by "capping" the QM site with a hydrogen atom or an empirically parametrized pseudopotential ("dummy atom") and such a bond is treated on the QM level. The remaining bonding interactions of the interface region, *i.e.* angle bending and dihedral distortions, are described within the classical force field. The same holds for the van der Waals interaction between QM and MM parts of the system.

The electrostatic effects of the classical environment, on the other hand, are taken into account in the quantum mechanical description as an additional contribution to the external field of the quantum system,

$$H_{\mathrm{QM/MM}}^{\mathrm{el}} = \sum_{i \in MM} q_i \int d\vec{r}\, \rho(\vec{r})\, v_i\!\left(\left|\vec{r}-\vec{r}_i\right|\right),$$

(2.25)

where q_i is the classical point charge located at r_i and $v_i\!\left(\left|\vec{r}-\vec{r}_i\right|\right)$ is a Coulombic interaction potential modified at short range in such a way as to avoid spill-out of the electron density to nearby positively charged classical point charges [29]. In the context of a plane-wave-based Car-Parrinello scheme, a direct evaluation of Eq. 2.25 is prohibitive, because it involves on the order of $N_r N_{MM}$ operations, where N_r is the number of real space grid points, typically $\sim 100^3$, and N_{MM} is the number of classical atoms, usually of the order 10 000 or more in system of biochemical relevance. Therefore, the term in Eq. 2.25 is included exactly only for a set of MM atoms in the vicinity of the QM system. The electrostatic interaction between the classical point charges of the more distant MM atoms and the QM system is calculated by a multipolar expansion of the full interaction given in Eq. 2.25. In this way, efficient and consistent QM/MM Car-Parrinello simulations of complex extended systems can be performed. The steric and electrostatic effects of the surroundings can be taken into account explicitly while the total energy of the coupled QM/MM system is conserved during the dynamics.

Chapter three

Investigation of Properties of Flavonoids Influencing the Binding to Bilitranslocase: a Neural Network Modeling

ABSTRACT

Bilitranslocase is a plasma membrane carrier. This work is aimed at characterizing the interaction of bilitranslocase with flavonols, a flavonoid sub-class. The results obtained show that, contrary to anthocyanins, flavonol glycosides do not interact with the carrier, whereas just some of the corresponding aglycones act as relatively poor ligands to bilitranslocase. These data point to a clear-cut discrimination between anthocyanins and flavonols occurring at the level of the bilitranslocase transport site. A quantitative structure-activity relationship based on counter propagation artificial neural network modelling was undertaken in order to shed light on the nature of flavonoid interaction with bilitranslocase. It was found that binding relies on the ability to establish hydrogen bonds, ruling out the involvement of charge interactions. This requisite might be at the basis of the discrimination between anthocyanins and flavonols by bilitranslocase and could lie behind some aspects of the distinct pharmacokinetic properties of anthocyanins and flavonols in mammals.

3.1 Introduction

Flavonoids are heterocyclic, polyphenolic compounds characterized by a common basic structure consisting of two aromatic rings (A and B), bound to an oxygenated heterocycle (ring C) (Fig. 3.1). The chemical repertoire of flavonoids is large, due to different patterns of hydroxylation, methoxylation and glycosylation of their common structure. They are plant secondary metabolites occurring at relatively high concentrations in several kinds of fruits, grains and vegetables harvested for human consumption [64]. The prevalence of such food in the human diet has recently been associated with significant reductions of the risk factors in chronic human pathologies, such as diabetes, cancer, neuro-degenerative and cardiovascular diseases [65-67]. At the cellular level, flavonoids have been found to exert a variety of biological effects [68], presumably mediated by specific interactions with molecular targets. Indeed flavonoids have been shown to interact with biological macromolecules, such as nucleic acids [69-71], polysaccharides [72, 73] and proteins [72-77]. The critical step determining the ability of any compound to reach an intracellular target is its translocation through the cell plasma membrane, for which the activity of specific transport proteins is mandatory in the case of hydrophilic and sterically complex compounds [78]. Among the carriers possibly involved in flavonoid membrane transport is bilitranslocase, a membrane transporter firstly identified in the liver,
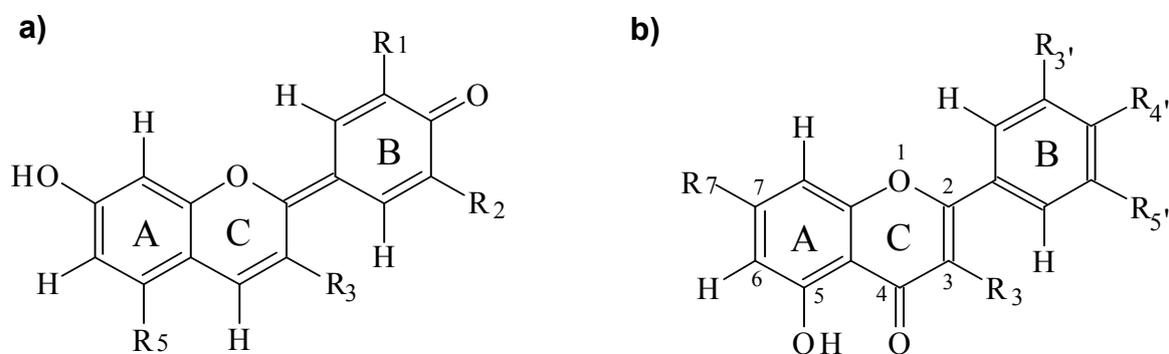
Figure 3.1 Chemical structures of anthocyanins (a) and flavonols (b). The substitutions occur at positions denoted by R (R is specified in Table A.1 in Appendix).

where it is expressed on the sinusoidal domain of the plasma membrane [79, 80]. At this level, its physiological function is to mediate the diffusion of organic anions from the blood into the liver, thus playing a role in the hepatic detoxification pathway(s) of endo- and xenobiotics. Its established substrates are bilirubin [81, 82] and nicotinic acid [81] with dissociation constants K_d = 2 nM and K_d = 11 nM, respectively, sulfobromophtalein with K_d = 5 μM) [83, 84] and anthocyanins with K_d = 1.5-22 μM [85]. The interaction mechanism of bilitranslocase binders has not yet been established, since the secondary structure of bilitranslocase is not known.

This investigation focuses on the structural properties of the ligands in order to infer the mechanisms of their interaction with the transporter. Thus, we study the nature of the interactions between bilitranslocase and flavonoid ligands, both anthocyanins and flavonols, by the counter propagation artificial neural network method, one of the computational approaches already validated as a proper tool in the investigation of ligand activity [86]. We have used this approach first to classify the tested molecules into three categories, according to their effect on bilitranslocase transport activity: i) competitive inhibitors (C), ii) noncompetitive inhibitors (N), iii) inactive molecules (I). With CP-ANN modeling the inhibition constant K_I of competitive inhibitors can be predicted. Special attention is dedicated to the examination of the kind of molecular descriptors needed to create the model. The results of this work show that, contrary to dietary anthocyanins, most of dietary flavonols do not interact with bilitranslocase, while some flavonol aglycones act as poor ligands of that carrier [87]. A quantitative analysis of the structure-activity relationship leads to the identification of parts of ligands potentially involved in the binding to bilitranslocase, along with an inference on the kind of interaction between the ligand and the target.

3.2 Methods and models

A detailed description of the counterpropagation artificial neural network architecture and its learning strategy are given in many articles and text books [88, 89]. In Appendix A.1 a short description of the method is given in reference to the specific application that is presented here.

The experimental data for the molecules listed in Table A.1 were obtained by studying in vitro bilitranslocase transport activity [87]. The models for 22 anthocyanins and 21 flavonols were built up and a structural optimisation was performed for each of them. The semiempirical AM1 method within MOPAC packages [90] was used to obtain the equilibrium structures. Next, the CODESSA program was used to calculate the descriptors on the basis of optimised geometrical parameters [91]. We obtained 353 descriptors for each molecule as CODESSA output. Structural descriptors are illustrated in Appendix A.3.

The crucial point in chemometrics is to obtain a proper set of descriptors, which very often means a reduction in the number of the originally calculated descriptors. They need to be carefully chosen in order to obtain the best distribution of molecules in the top-map. Visual inspection of the top-map gives us information on clusters as structural similarity relationships between molecules. Descriptors can provide the knowledge about molecular features, which make the selected molecule, for instance a good competitive inhibitor. Based on such a study we are able to draw some hypothesis of the potential nature of the ligand – target interaction. The selection of relevant descriptors is done for two independent purposes. First we wanted to classify the molecules according to their effect on bilitranslocase transport activity (I, C, N). To achieve this aim the average of the absolute deviation was calculated and all descriptors with values smaller than a threshold of 0.8 were removed. In this way the non-discriminative variables that are similar for all molecules are effectively eliminated. The models were tested with the leave-one-out method [89]. The correlation coefficient of a leave-one-out test was used as a criterion to estimate

the quality of the classification models. After following these procedures the initial number of descriptors is reduced to 207 descriptors.

Second, we wanted to predict the K_I value of competitive inhibitors. For this purpose we used only 18 molecules from the data set, those that show the competitive inhibition of bilitranslocase. With these molecules we started the selection procedure once again with the whole set of descriptors. The initial selection was done using the transposed data matrix, in which one row is an object representing one descriptor and contains the values of that particular descriptor for all compounds. The dimension of the input vectors is now equal to the number of molecules in the data set. After inputting the transposed data matrix into the input layer of CP-ANN, we obtained the Kohonen map of descriptors. Some of them either occupied the same neuron or formed clusters. The descriptors, which occupied one single neuron were selected as specific and chosen to be submitted to the next selection method. In the next step we applied the genetic algorithm [92] (see also Appendix A.2) that generates the lowest possible number of descriptors, thus yielding a satisfactory prediction for K_I for each individual inhibitor in the data set. Once again the leave-one-out cross-validation method for the training set data was used for verifying the models with different numbers of descriptors. After the iterative procedure the best results were obtained for a set of 155 variables and these results are presented in the following section. The set of these 155 descriptors is listed in the appendix in Table A.2.

3.3 Results

3.3.1 CP-ANN model for classification of the bilitranslocase ligands

The set of 43 molecules was studied for the classification purpose. In this study the model adaptive parameters are: number of training epochs, dimension of network (n_x, n_y) and maximal correction factor. We divided the molecules into the training and test sets on the basis of clusters formed on the Kohonen map.
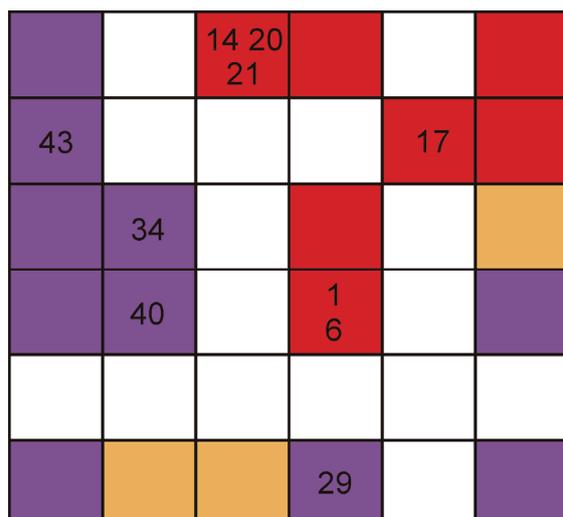
Figure 3.2 The Kohonen map of the classification CP-ANN model, 6 × 6 neurons. The neurons occupied by the competitive, non-competitive or inactive ligands are colored red, orange or purple, respectively. The test compounds are denoted by their ID numbers.

We took care to include into the test set all representatives of groups of structurally similar molecules. The CP-ANN was trained with the molecules of the training set with the following network parameters: $n_x = n_y = 6$, $a_{max} = 0.3$, $a_{min} = 0.01$, epochs = 200.

The training molecules were well separated into clusters irrespective of their inhibition potency. The testing set of molecules (ID 1, 6, 14, 17, 34, 40, 43; see Table A.1) served for an evaluation of the classification model. All molecules were correctly classified, so we went one step further and tested the classification CP-ANN model with three other molecules, which had not been tested experimentally (ID 20, 21, 29; see Table A.1). Two of them were classified as competitive inhibitors and one as an inactive molecule. The Kohonen map is illustrated in Figure 3.2. After a detailed study of structures and descriptors of the compounds not experimentally tested and in the view of their similarity to experimentally tested molecules, we concluded that the classification was correct.

The most problematic group for classification contained the non-competitive inhibitors. The data were available for only four molecules, *i.e.* ID 18, 19, 23, and 24. Molecule No. 25 (quercetin) was treated as competitive because of higher competitive than non-competitive K_I (see Table A.1). We even had to remove one of the non-competitive molecules (ID 18) from the training set, because it caused a conflicting situation with corresponding descriptors too similar to be distinguished from other competitive molecules. The general problem of the

investigated molecules is that slight changes in the structure cause big changes in the activity. This was particularly evident with compounds ID 17 (Cyanidin 3-O-α-L-arabinopyranoside, competitive inhibitor) and ID 18 (Cyanidin 3-O-β-D-galactopyranoside, non-competitive inhibitor). In the final classification model only three non-competitive inhibitors were included (ID 19, 23, 24). A set of three compounds is obviously not enough for extracting the proper statistical evaluation of the predictions obtained by an experience-based model as CP-ANN is. We can only say that the three non-competitive inhibitors, with the exception of ID 18 that was discarded, were separated from the other compounds (see Fig. 3.2, orange squares). In the absence of test compounds, no prediction was possible.

3.3.2 CP-ANN model for prediction of K_I

The prediction of K_I is an essential step to validate the CP-ANN analysis of experimental data. To study the prediction of K_I value, a set of 18 molecules was used, which experimentally show competitive activity. The molecules with ID numbers from 1 to 17 and quercetin (ID 25, see Table A.1)
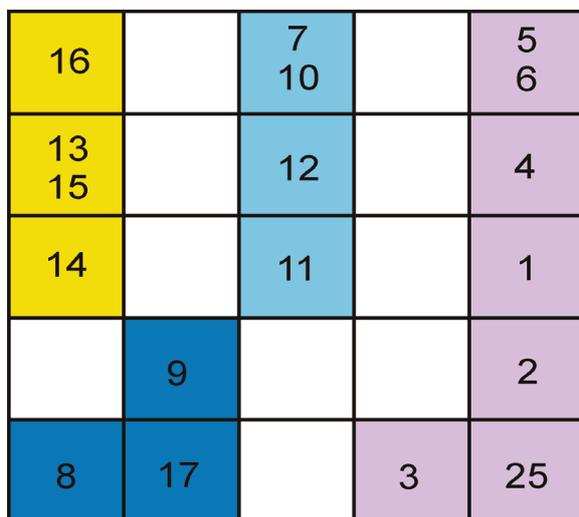
Figure 3.3 The Kohonen map of the predictive CP-ANN model, 5 × 5 neurons. The ID numbers of the 18 competitive inhibitors are shown as they appear in the trained network. The colors indicate structural similarity of groups of compounds: yellow for diglucosides, pink for aglycones, dark and light blue for monoglucosides with lower and higher affinity, respectively.

a)

6		**4**		**1**	**2** **25**
	5				
12					**3**
10 **7**					
				9	**17**
13 **14**	**15** **16**		**11**		**8**

b)

c)
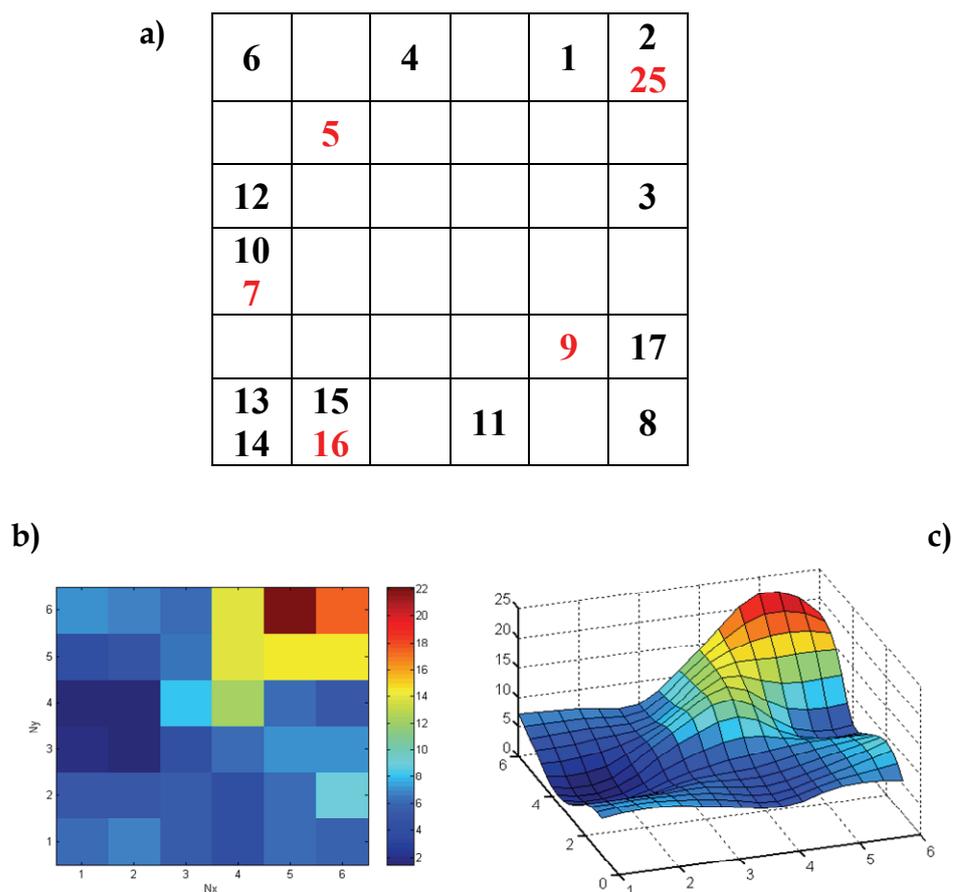
Figure 3.4 The top-map with 6 × 6 neurons **(a)**, the response surface **(b)**, and 3D view of the response surface **(c)** of the final predictive CP-ANN model trained with 15 training compounds.

belong to this set. By using the 155 descriptors set and following computational parameters $n_x = n_y = 5$, $a_{max} = 0.3$, $a_{min} = 0.01$, epochs = 250, we obtained the Kohonen map reported in Figure 3.3. The molecules are well clustered with respect to their K_I values. However, perhaps even more significant, such clusters grouped molecules with various degree of glycosylation: the yellow cluster gathers the di-glucosides, the pink cluster gathers the aglycones, and the blue cluster gathers the mono-glucosides with either higher (light blue) or lower (dark blue) affinities for bilitranslocase. The next modelling step was to split the molecules into the training and testing set. For the testing set we chose 5 molecules (ID 5, 7, 9, 16, 25). We trained the model until a suitable distribution of training molecules was obtained (Fig. 3.4 a, black numbers). The network parameters were the following: $n_x = n_y = 6$, $a_{max} = 0.35$, $a_{min} = 0.01$, epochs = 200.

We tested the model with the molecules from the test set and, as a result, all the molecules (Fig. 3.4 a, red numbers) were placed in the correct clusters of trained molecules.

The predictions of the trained CP-ANN model are stored in the output layer shown in Figures 3.4 b and c. From Figure 3.4 b it can be seen that we obtain 6×6 points of the K_I response surface for 6×6 neurons in particular network architecture. The colour bar is given to associate the individual colour of the square with the K_I value; the range 2 to 22 corresponds to the range of K_I values of the training compounds. The prediction of a test compound is obtained by

ID	Experimental K_I^e	Predicted K_I^p
Training set		
1	22.21	22.16
2	17.55	17.59
3	5.27	5.27
4	6.23	6.23
6	7.20	7.20
8	5.78	5.81
10	1.83	1.83
11	4.03	4.03
12	1.42	1.42
13	6.42	6.09
14	5.77	6.09
15	6.81	6.80
17	9.16	9.13
Testing set		
5	7.57	4.85
7	2.79	1.83
9	8.57	6.23
16	6.36	6.79
25	21.10	17.59
External test compound		
BSP	5.32	4.03

Table 3.1 Experimental and predicted values (µM) of competitive inhibition constant for 18 compounds (13 training and 5 testing compounds, ID numbers given in Table A.1) and for external test compound BSP (sulfobromophthalein)

pinpointing a co-ordinate on the output surface. In Figure 3.4 c, a 3D view of the output surface with larger resolution (3 × 3 squares per neuron) is given. The colour bar is the same as in Figure 3.4 b.

The RMSE of the predicted K_I values is equal to 2.2. For a small dataset such as that investigated in the current study this error is expected. The K_I values of the new molecules predicted by this model could be a good approximation of their actual values. We are especially satisfied with the results, because the tested molecules can be correctly placed in clusters of the training model. That result proves that the set of descriptors is properly chosen for prediction of the K_I value. Table 3.1 reports the experimental and predicted K_I values.

3.3.3 Structural descriptors influencing the binding of anthocyanins to bilitranslocase

The detailed investigations of the 155 descriptors used to classify the compounds lead us to conclude that the ligand ability to act as H-bond donors and/or acceptors is very important for establishing an interaction with the host molecule. Descriptors directly defining the ability for creating this kind of bonds were indicated by an automated genetic-algorithm-based variable selection procedure. Only the selected descriptors enabled the CP-ANN training procedure to cluster the molecules according to their affinity to the target.

3.3.4 Final validation of the CP-ANN model

The final validation of the CP-ANN model was performed by testing the inhibition activity of BSP, which is an established substrate of bilitranslocase with the K_I = 5.32 ± 0.63 μM [83, 84]. The 3D structure was prepared in the same way as for all other molecules in the study. Once structural descriptors were obtained, only those selected in the model building and variable reduction procedure were accepted, *i.e.* 155 descriptors. The predicted K_I = 4.03 μM (see Table 3.1). The deviation $\left| K_I^e - K_I^p \right| = 1.29\,\mu M$ is twice as large as the established experimental error of ± 0.63. However, it is within range of the obtained model error (RMSE = 2.2) for the testing set, which validate our CP-ANN model.

3.4 Discussion

The main result of this work is in the identification of the nature of molecular interactions between the flavonoids and bilitranslocase. This enables a structure-based classification of molecules according to their activity. Taking into account the structural features of the molecules belonging to a given cluster and focusing on the nature of molecular interactions between the carrier and the ligands, a system of deductions about the structural requirements for the activity of bilitranslocase ligands is proposed below.

3.4.1 The structural basis of classification of molecules

To understand the structural basis for the partition of molecules in well separated clusters of inactive and active (competitive and non-competitive) molecules, a comparative analysis of structural and functional features of anthocyanins *vs* flavonols was particularly helpful.

3.4.1.1 Active *vs* inactive molecules: the importance of the 3D structure of the molecule

The role played by the 3D structure of the ligands was noticed for the compounds 2 (cyanidin, an anthocyanidin) and 25 (quercetin, a flavonol) and their analogues 8 and 32, substituted with the sugar at position R3 (2→8 and 25→32). Substitution on cyanidin (ID 2→8) made it a better inhibitor, since its K_I decreased from 17.55 to 5.78 μM. On the contrary, substitution on quercetin (ID 25→32) made it an inactive molecule. We compared the optimised 3D structures of these molecules, which are shown in Figure 3.5, and we noticed that they differ mostly in the orientation of the sugar moiety. The sugar in ID 8 lies in the same plane as the planar aromatic rings. In this way, the hydrogen atoms of the

a)

b)

c)
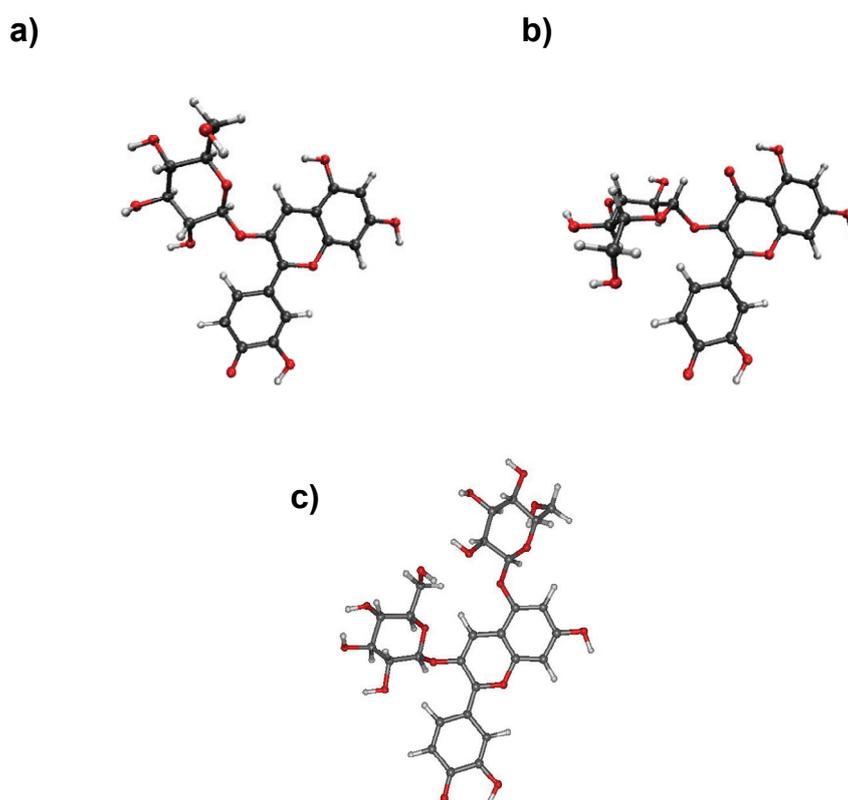
Figure 3.5 3D structures of **(a)** cyanidin 3-glucoside (ID = 8), **(b)** quercetin 3-glucoside (ID = 32) and **(c)** cyanidin 3,5-diglucoside (ID = 14).

sugar OH groups point to the same direction as the substitutes of B ring, ready to create hydrogen bond(s) with the target. On the contrary, the same sugar moiety bound to an equivalent position in quercetin is located perpendicularly to the planar condensed rings of the molecule. This position is forced by the presence of the oxygen atom in ring C. That may be a good reason for the molecule ID 32 to be inactive even in a protein environment – it simply does not fit the binding pocket of the host molecule. The energy needed to flip down the sugar moiety could be too high, especially if the carbonyl group of ring C acted as an H-bond acceptor from the carrier active site. We looked at the structure of ID 14 (cyanidin 3,5-di-O-β-Dglucopyranoside), in order to evaluate the 3D structure of the two sugar substitutions. We found that both are almost co-planar with rings A and C. From

these observations it can be deduced that the ligands of bilitranslocase, whether acting as competitive or non competitive inhibitors, have to be planar. This conclusion is in agreement with the original experimental observation that the only phthalein tautomer transported by bilitranslocase is the planar, quinoidal one [93].

3.4.1.2 Competitive *vs* non-competitive inhibition by aglycones: the importance of hydroxylation of B ring

As a step further, we tried to understand the structural basis underlying the property of bilitranslocase inhibitors to act either competitively or non-competitively. Some hints were provided by the behaviour of the three active flavonols (ID 23-25). Galangin (ID 23) is the only aglycone displaying pure non-competitive activity. Unique among all flavonoid aglycones tested (both anthocyanins and flavonols), its ring B has no oxygen atom at position R4'. We deduced that its interaction with the target must involve the condensed rings A and C, the only moieties suitable for H-bond formation. By contrast, the introduction of an OH group in the 4' position of the B ring of flavonols (as in ID 24 and 25) confers them the further property to interact at the level of the competitive site, like most of the anthocyanins (ID 1-19).

3.4.1.3 Pure competitive *vs* mixed-type inhibition: the possible role of tautomerism of ring B

To understand the nature of mixed-type inhibition, typical of two flavonols (ID 24-25), we started from the evidence that the CP-ANN model clustered competitive molecules separately from non-competitive ones, reflecting two well distinct modes of interaction of the same molecule with the carrier. Given the importance of hydroxylation of the B ring in determining the type of inhibition (and its strength as well, as discussed below), we speculated that mixed-type inhibition is compatible with the binding of two alternative structures of ID 24-25 to the carrier, arising from pH-dependent phenolic-quinoidal tautomerism of the
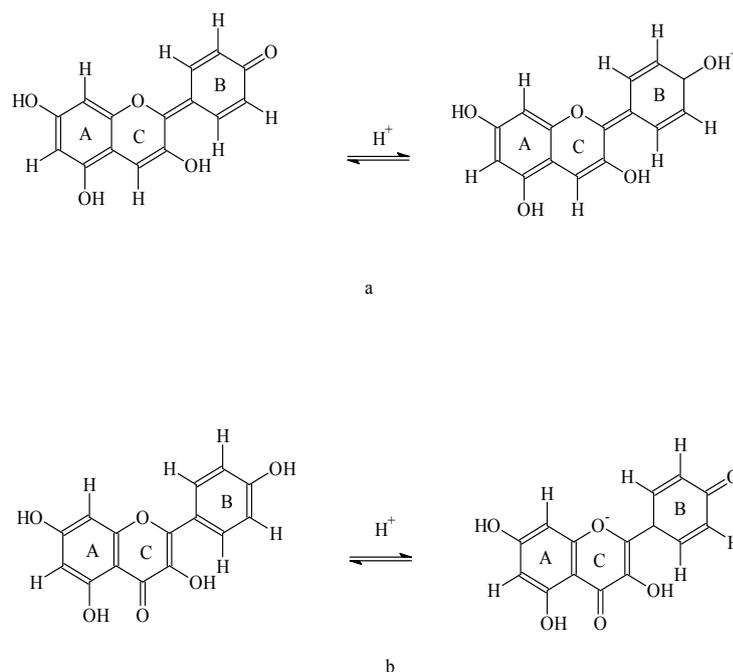
Figure 3.6 pH-dependent acidic equilibrium of anthocianins **(a)** and flavonols **(b)**.

B ring (see Fig. 3.6). Given the fact that the pK_a values of ID 24 and 25 are 8.1-8.2 and 7.0-7.03, respectively [94], both structures of either ID 24 or 25 occur in the assay medium enough to generate mixed-type inhibition.

In the evaluation of this hypothesis, it should be observed that only flavonols have such dual activity, while anthocyanins are purely competitive inhibitors. Phenolic and quinoidal species of both classes of compounds occur in the assay medium. However, while the phenolic species of anthocyanins are cationic, those of flavonols are neutral.

The structural descriptors, selected by the GA, that enabled both the classification and the prediction of the activity of competitive inhibitors (Fig. 3.3 and 3.4), referred to the capability of establishing H-bonds. Interestingly, they did not include, and likely excluded, charge interactions with the target. If so, anthocyanin aglycones might be all pure competitive inhibitors simply because their phenolic species are cationic and, as such, could establish no interaction with the target. It should therefore be concluded that quinoidal tautomers are likely to be the species competitively interacting with the carrier. By contrast, both the phenolic and the quinoidal species of flavonols are neutral and both could

therefore bind the carrier. However, while quinoidal species of flavonols would act as competitive inhibitors, in analogy with anthocyanins, phenolic species would act as non-competitive inhibitors, in analogy with galangin (ID 23) whose B ring cannot originate a quinoidal tautomer.

The concept that *para*-quinoidal species of flavonoids competitively interact with bilitranslocase is supported by the experimental evidence that the pH-indicator phthaleins are transported only as *para*-quinoidal species [93]; it is quite realistic that competitive inhibition of phthalein transport should be exerted by the structurally closest tautomer of flavonoids. Another point in favour of this concept is that only the *para*-quinoidal species of flavonols definitely set the B ring in the same plane of the rest of the molecule, a structural requirement for phthaleins to be transported and for glycosylated flavonoids to be active.

Yet, what would prevent phenolic species, freely rotating around C1, from binding to the competitive site? With respect to the formation of a network of H-bonds, we have to deduce that the main, obvious difference between *para*-phenolic and *para*-quinoidal species is that the former are H-bond donors whereas the latter are H-bond acceptors. This would suggest that a decisive structural feature for a molecule to act as a competitive inhibitor of bilitranslocase transport activity is the availability of atoms acting as H-bond acceptors from corresponding H-bond donors of the active site of the carrier. Conversely, the H-bond formation by donating a hydrogen by a *para*-phenolic tautomer to the active site of the carrier might involve a configuration of the flavonol B ring compatible with non-competitive inhibition.

3.4.2 The structural basis of the activity of competitive inhibitors

The visual inspection of the Kohonen map shown in Figure 3.3 highlights the relationship between the structure and the activity of investigated compounds. Clusters are formed according to the degree of glycosylation. The best competitive inhibitors are anthocyanin mono-glucosides, separated in two sub-classes grouping molecules with $K_I < 4$ μM (ID 7, 10-12; light blue) and $K_I > 5.8$ μM (ID 8, 9, 17, dark blue) respectively. The di-glucosides form the next group, if ranged by

affinity and the last are the aglycones, with the lowest affinity (in Fig. 3.3 shown as yellow and pink squares, respectively). This map indicates that the carrier active site somehow "reads" the glycosylation state of anthocyanins, thus implying not only that it is a complex and multi-compliant structure, but also that it provides a polar environment for the interaction with polar moieties of the substrates.

3.4.2.1 Flavonoid aglycones: the importance of hydroxylation of the B ring

Among competitive inhibitors, the least polar anthocyanidin (pelargonidin, ID 1) is the least active one (K_I = 22.21 μM). Similarly, the least polar flavonol (kaempferol, ID 24) has a very poor competitive inhibition activity (K_I = 131.6 μM). The presence of a second hydroxyl group on ring B, as in either cyanindin (ID 2, an anthocyanin) or in quercetin (ID 25, a flavonol), enhances the competitive inhibition of both compounds. While among anthocyanins the gain in activity is just above the statistical significance: K_I = 22.2±1.65 (ID 1) *vs* K_I = 17.5±1.68 μM (ID 2), the effect is marked among the flavonols, as the K_I drops from 131.6±3.8 (ID 24) to 21.1±1.7 μM (ID 25). Interestingly, the difference in K_I of either cyanidin (ID 2) or quercetin (ID 25) is not significant, suggesting that both compounds interact with the target through the moiety they have in common, *i.e.* the ring B. Thus, the number of H-bonds established at the level of the B ring is indeed very important in shaping the strength of interaction of the aglycones. In more detail, the B ring interactions are based on the presence of two vicinal oxygen atoms, possibly acting as a couple of H bond acceptor (the *para*-quinoid carbonyl) and H-bond donor (the R3' hydroxyl group). At the same time, the steric hindrance exerted by the carbonyl group of the C ring, a specific feature of the flavonols, appears to be successfully overcome if the B ring interactions prevail over the others.

3.4.2.2 Flavonoid aglycones: the importance of steric hindrance and charge distribution caused by the carbonyl group of the C ring

The carbonyl group of ring C in flavonols appears to heavily influence the magnitudes of the competitive K_I values. That was realised by comparing

pelargonidin (ID 1 an anthocyanin) with kaempferol (ID 24, a flavonol). The anthocyanin is nearly 6-fold more active than the flavonol (22.2 vs 131.6 µM). The flavonol appears to be sterically hindered by the carbonyl group of the C ring. It can be also considered that the carbonyl group on the C ring favours the establishment of intramolecular H bonds with the OH groups in R3 or in R5 [35], thus possibly limiting the overall tendency of flavonols to interact with the carrier and therefore increasing their apparent competitive K_I.

In flavonols, the presence of a hydroxyl group in R5′ (ID 26 vs ID 23, 24, 25) or, as predicted by the CP-ANN (Fig. 3.2), O-methylation of the B ring (ID 29 vs ID 25) results in the total loss of activity. In none of the anthocyanins were these all-or-none effects observed on activity (ID 3 vs ID 1, 2 and ID 4 vs ID 2). On the contrary, triple hydroxylation of the B ring of anthocyanins is better than double hydroxylation (ID 3 vs ID 2) and O-methylation is also well accepted, apparently because the anthocyanin aglycone can interact with the target also through the A ring, as previously proposed [85]. In flavonols, the functional groups introduced in the B ring, as in ID 26 and ID 29, appear to be equally effective in preventing their interaction with the carrier, presumably by steric hindrance. It results therefore that the carbonyl group of the C ring in flavonols acts as constraint for their ability to take different orientations at the level of the transport site(s) of the carrier. Thus, in interacting with the carrier, flavonols might be seen as rather stiff molecules, with a limited repertoire of interactions, whereas anthocyanins appear to be more flexible and thus better at fitting the chemical environment of bilitranslocase active site.

3.4.2.3 Anthocyanidin mono- and diglucosides

The distinctive feature for clustering mono-glucosides in two groups according to their activity (Fig. 3.3, light and dark blue) is the OH group at position R1 of the B ring, which occurs in the low-affinity molecules (K_I > 5.8 µM, ID 8, 9, 17) and is missing in the high-affinity ones(K_I < 4 µM ID 7, 10-12). In the case of mono-glucosides, an increase in the polarity of B ring (see Table A.1) lowers the affinity

for the target (ID 7-9, 10-12). While this finding emphasises one more time the critical role of this moiety in the interaction with the carrier, it also reflects the existence of a specific binding pocket for the glycosyl moiety attached to C3, improving the binding affinity to such an extent that the interaction with the B ring is devalued. Any changes in the sugar moiety result in the decrease of affinity to the target (*e.g.* ID 8 *vs* ID 17) or even in a change of the type of inhibition (*e.g.* ID 8 *vs* ID 18 and ID 19, Table A.1). This may be an indication that the ligand does interact with the target at two sites, *i.e.* at one site through the sugar moiety and at the other through the B ring.

In contrast, for diglucosides, changes at substituents of ring B don't influence the K_I value so much. That can suggest that the diglucoside-target interactions occur either through both sugars or through ring A and its glucosyl moiety, as proposed earlier [85]. It appears therefore that the target has a structurally complex active site, providing specific, hydrogen bond-based interactions for a potentially wide spectrum of structurally unrelated ligands.

3.4.3 A hypothesis for the structural basis of the activity of ID 18

A glycosyl moiety in C3 improves the interaction of anthocyanins with the carrier, because an additional set of H bonds can be established. Cyanidin 3-O-β-D-galactopyranoside (ID 18) clustered with competitive inhibitors, without being one. This would indicate that ID 18 might indeed bind to the carrier at the same level as the other mono-glucosides, although with lower affinity (K_I = 35.2 µM *vs* K_I = 1.42-9.16 µM). This concept is in apparent conflict with the tenet that non-competitive inhibitors bind to a site that is close to, but does not coincide with an enzyme active site. However, binding of the galactosyl moiety to the sugar-specific pocket of the carrier might affect the conformation of the active site of the carrier, thus perturbing the kinetics of BSP transport. Two possibilities are equally applicable, *i.e.* ID 18 either induces an alternative, catalytically less efficient, conformational modification of the transport site or it causes a blockade of a conformational modification essential for transport catalysis.

3.5 Conclusions

It can be concluded that the requirement for a flavonoid molecule to interact with the target is the overall planarity. Then, to interact in a competitive way, and thus to be presumably transported, it has to have a *p*-quinoid group on the ring B, in order to be a hydrogen bond acceptor.

If it has a *p*-hydroxyl group on the ring B, it is a non-competitive inhibitor. The OH group may therefore play a dual role, either as a proton donor, or, in the deprotonated form, as a strong hydrogen bond acceptor. It would remain a challenge to verify the above statements experimentally or using other computational tools, such as quantum mechanics based method.

It cannot be stated yet that inhibitors of electrogenic BSP transport activity are indeed transported, unless they are tested directly as transport substrates. However, the dependence of the interaction of flavonoids with bilitranslocase on H bonds shows that those molecules interact without charge compensation, an essential requirement to be transported electrogenically. By an independent experimental approach, transport of malvidin 3-glucoside in liver cells has been shown to be mediated by bilitranslocase [95], suggesting that, at least for competitive inhibitors, this could be the case.

3.6 Appendix

A.1 Counterpropagation Artificial Neural Network

The CP-ANN modeling is based on a two-step learning procedure, which is unsupervised in the first step and supervised in the second one, where both steps are completely independent, but have to follow each other in the iterative process of learning [89, 92]. As it is shown in the Figure A.1 the architecture of CP-ANN is made of two layers of neurons represented as vectors of weights arranged in a two-dimensional rectangular matrix. The first layer is called input or Kohonen layer. Its structure and learning strategy imitates the structure of the brain, which is why it is named an artificial neural network. The Kohonen
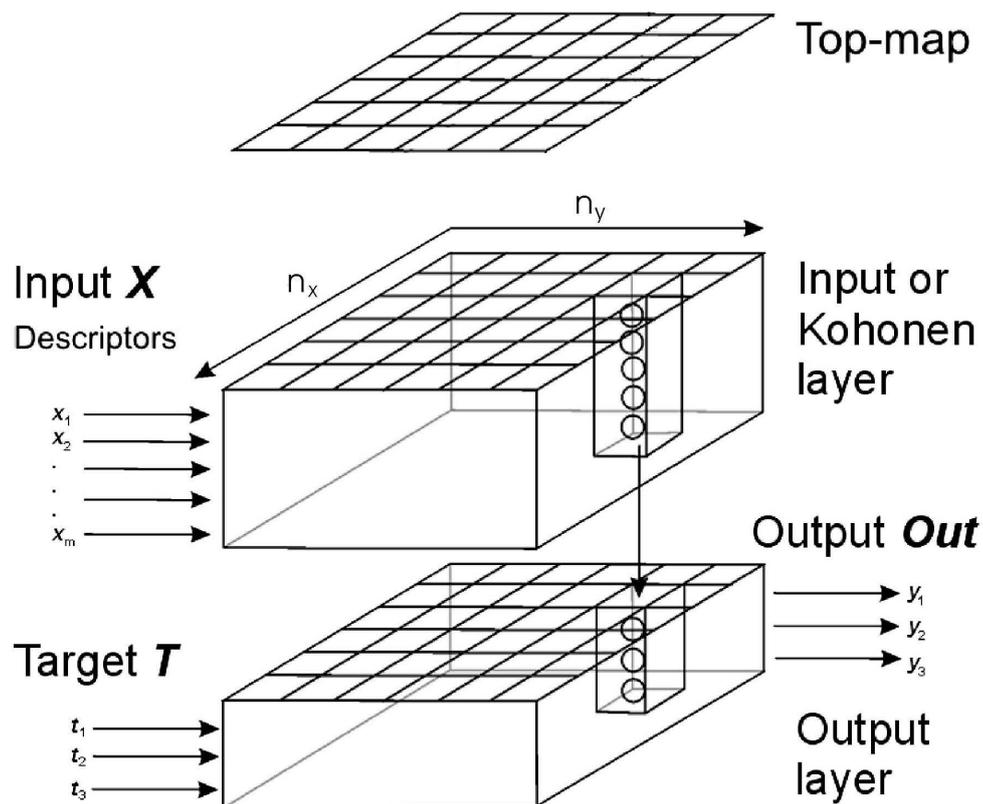
Figure A.1 The scheme of a counterpropagation artificial neural network (CP-ANN).

layer is represented as a tensor in three dimensions $n_x \times n_y \times m$. n_x, n_y are the number of neurons represented in the X and Y dimensions, while m is the number of descriptors that spans the Z axis (Fig. A.1). The number of neurons $N = n_x \times n_y$ is a computational parameter related to the number of molecules in the training set. The Kohonen layer has a well-defined topology of neurons with defined lattice structure. The first step, the unsupervised learning procedure is determined by descriptors $X = x_1, \dots x_k, \dots x_m$. During the training molecules with similar descriptors are placed close to each other and form clusters that can be observed in the top-map or Kohonen-map of the CP-ANN (see Fig. A.1). It is important that only the descriptors guide the determination of the position of the molecules in the top-map. This top-map contains the ID number of molecules that were placed at particular positions in the Kohonen layer after learning, or a blank, if the corresponding neuron in the input layer is empty. When we plot the top-map, as it is done in Figures 3.2, 3.3, 3.4, we can see which positions are occupied by individual molecules at the end of the training. The Kohonen layer contains the "weights" produced in the training procedure. The weights $w_{i,j,k}$ are real numbers, one for each element i,j,k in the three-dimensional tensor ($n_x \times n_y \times m$).

The second, supervised step of the learning procedure involves training the output layer, guided by the input layer (see Fig. A.1). The position of molecules stored in the Kohonen layer is imposed on the output layer in each iteration. As a result of the second step in the learning procedure each neuron in the output layer provides an output vector ***Out*** for each molecule as close as possible to the initial target vector ***T***. In the application of CP-ANN for the classification of flavonoid molecules, the target ***T*** is a three-component binary vector (t_1, t_2, t_3, as shown in Fig. A.1) representing one of three possible classes of bilitranslocase substrates, with $T_C = (1,0,0)$ for the competitive inhibitors, $T_N = (0,1,0)$ for non-competitive inhibitors and $T_I = (0,0,1)$ for the inactive molecules. The artificial neuron network trained in such a way can be used to classify new molecules by first situating them in the Kohonen layer on the neurons with the most similar weights. These positions are then projected to the output layer, which provides the output class C, N or I of substrate.

In the second application of CP-ANN for the prediction of K_I, the target T is the experimental value of K_I and the output layer is a plane, since it has only one value at each position. The neural network again has to be trained in two steps, but now with respect to predicting a single scalar value, the K_I, for the competitive inhibitors.

The computational parameters of the neural network construction can influence the results and should be tested and optimized. In particular the number of training epochs, the dimension of the network (n_x, n_y), the maximal correction factor (a_{max}), the minimal correction factor (a_{min}) [89].

A.2 Genetic Algorithm

A genetic algorithm can be applied to address problems in which optimal values of a set of parameters or an optimal reduced set of parameters, resulting from a combination of original parameters, is sought. In the present application the method was used to reduce the number of descriptors needed for proper clustering of the training set of molecules. The origin of its name comes from the resemblance with the biological process: the GA consists of three basic processes mimicking Darwinian evolution: crossover, mutation and survival of the fittest. In the crossover step, new chromosomes are generated by mixing fractions of old individuals, so called parent chromosomes. Mutation is introduced to randomly change individual bits of the chromosome. In the last step the survivals are chosen for the next generation, *i.e.* the cycle of the GA. The survivals are the chromosomes having the best criterion value determined by the fitness function. This last step should yield the lowest possible number of descriptors which gives a satisfactory prediction for K_I for each individual inhibitor in the data set. The length of the chromosome, *i.e.* the number of genes or bits, is determined by the number of input parameters. A full description of the GA can be found in Ref. [88]

A.3 Structural descriptors

Descriptors are used to characterize mathematically the molecules. The set of descriptors in the CODESSA software is split into five categories: (i) **constitutional descriptors**, depending on the number and type of atoms, bonds, and functional groups; (ii) **geometrical descriptors**, which give molecular surface area, projections, and gravitational indices; (iii) **topological descriptors**, which are molecular connectivity indices, related to the degree of branching in the compounds; (iv) **electrostatic descriptors**, such as partial atomic charges and other parameters describing the possibility for some sites in the molecule to form hydrogen bonds; (v) **quantum-chemical descriptors**, *i.e.* total energy of molecule, the energies of the lowest unoccupied and highest occupied orbital, ionization potential, etc. All the descriptors have to be normalized, maintaining the original distribution, using a range scaling procedure according to:

$$x_k^{norm} = \frac{x_k - \overline{x_k}}{s(x_k)},$$

where x_k^{norm} is the normalized value of the descriptor x_k, the $\overline{x}_k$ is the average value and s is a standard deviation

$$s(x_k) = \sqrt{\frac{\sum (x_k - \overline{x_k})^2}{n-1}},$$

with n the number of molecules.

Table A.1 Anthocyanin and flavonol derivatives with experimental data on inhibition constants (K_I) of bilitranslocase transport activity (C – competitive, N – non-competitive, I – inactive); SE: standard error; R: substitutions at positions defined in Figure 3.1.

ANTOCYANINS							
COMPOUND	ID	R_1	R_2	R_3	R_5	K_I	SE
Pelargonidin	1	H	H	OH	OH	22.21 [C]	1.65
Cyanidin	2	OH	H	OH	OH	17.55 [C]	1.68
Delfinidin	3	OH	OH	OH	OH	5.27 [C]	0.38
Peonidin	4	OCH$_3$	H	OH	OH	6.23 [C]	0.51
Petunidin	5	OCH$_3$	OH	OH	OH	7.57 [C]	0.99
Malvidin	6	OCH$_3$	OCH$_3$	OH	OH	7.20 [C]	0.40
Pelargonidin 3-O-β-D-glucopyranoside	7	H	H	β-D-glucopyranosyl	OH	2.79 [C]	0.18
Cyanidin 3-O-β-D-glucopyranoside	8	OH	H	β-D-glucopyranosyl	OH	5.78 [C]	0.39
Delfinidin 3-O-β-D-glucopyranoside	9	OH	OH	β-D-glucopyranosyl	OH	8.57 [C]	0.2
Peonidin 3-O-β-D-glucopyranoside	10	OCH$_3$	H	β-D-glucopyranosyl	OH	1.83 [C]	0.19
Petunidin 3-O-β-D-glucopyranoside	11	OCH$_3$	OH	β-D-glucopyranosyl	OH	4.03 [C]	0.19
Malvidin 3-O-β-D-glucopyranoside	12	OCH$_3$	OCH$_3$	β-D-glucopyranosyl	OH	1.42 [C]	0.13
Pelargonidin 3,5-di-O-β-D-glucopyranoside	13	H	H	β-D-glucopyranosyl	β-D-glucopyranosyl	6.42 [C]	0.29
Cyaniding 3,5-di-O-β-D-glucopyranoside	14	OH	H	β-D-glucopyranosyl	β-D-glucopyranosyl	5.77 [C]	0.39
Peonidin 3,5-di-O-β-D-glucopyranoside	15	OCH$_3$	H	β-D-glucopyranosyl	β-D-glucopyranosyl	6.81 [C]	0.77
Malvidin 3,5-di-O-β-D-glucopyranoside	16	OCH$_3$	OCH$_3$	β-D-glucopyranosyl	β-D-glucopyranosyl	6.36 [C]	0.45
Cyanidin 3-O-α-L-arabinopyranoside	17	OH	H	α-L-arabinopyranosyl	OH	9.16 [C]	0.99
Cyanidin 3-O-β-D-galactopyranoside	18	OH	H	β-D-galactopyranosyl	OH	35.22 [N]	0.58
Malvidin 3-O-(6-O-acetoyl)-β-D-glucopyranoside	19	OCH$_3$	OCH$_3$	(6-O-acetoyl)-β-D-glucopyranosyl	OH	58.33 [N]	0.09
Delfinidin 3,5-di-O-β-D-glucopyranoside	20	OH	OH	β-D-glucopyranosyl	β-D-glucopyranosyl	Not tested	---
Petunidin 3,5-di-O-β-D-glucopyranoside	21	OCH$_3$	OH	β-D-glucopyranosyl	β-D-glucopyranosyl	Not tested	---
Malvidin 3-O-(6-O-p-coumaroyl)-β-D-glucopyranoside	22	OCH$_3$	OCH$_3$	-(6-O-p-coumaroyl)-β-D-glucopyranosyl	OH	I	I

FLAVONOLS								
COMPOUND	ID	R_3	$R_{3'}$	$R_{4'}$	$R_{5'}$	R_7	K_I	SE
Kampferol	23	OH	H	H	H	OH	60.60[N]	1.00
Kaempferol	24	OH	H	OH	H	OH	131.6[C] 63.90[N]	3.80 3.40
Quercetin	25	OH	OH	OH	H	OH	21.10[C] 79.60[N]	1.70 3.60
Myricetin	26	OH	OH	OH	OH	OH	I	
Syringetin	27	OH	O-CH_3	OH	O-CH_3	OH	I	
Rhamnetin	28	OH	OH	OH	H	O-CH_3	I	
Isorhamnetin	29	OH	O-CH_3	OH	H	OH	Not tested	---
Quercetin 4'-glucopyranoside	30	OH	OH	Glu	H	OH	I	
Quercetin 3,4'-diglucopyranoside	31	Glu	OH	Glu	H	OH	I	
Quercetin 3-glucopyranoside	32	Glu	OH	OH	H	OH	I	
Quercetin 3-xyloside	33	Xyl	OH	OH	H	OH	I	
Quercetin 3-rhamnoside	34	Rham.	OH	OH	H	OH	I	
Quercetin 3-galactoside	35	Gal	OH	OH	H	OH	I	
Quercetin 3-O-gluco-pyranosyl-6''-acetate	36	Glu-Ac	OH	OH	H	OH	I	
Quercetin 3-O-sulfate	37	SO_4	OH	OH	H	OH	I	
Isorhamnetin 3-glucoside	38	Glu	O-CH_3	OH	H	OH	I	
Isorhamnetin 3-O-rutinoside	39	Rut	O-CH_3	OH	H	OH	I	
Kaempferol-glucoside	40	Glu	H	OH	H	OH	I	
Kaempferol 3-O-rutinoside	41	Rut.	H	OH	H	OH	I	
Syringetin 3-galactoside	42	Gal	O-CH_3	OH	O-CH_3	OH	I	
Syringetin 3-glucoside	43	Glu	O-CH_3	OH	O-CH_3	OH	I	

Table A.2 List of 155 descriptors selected for a purpose of K_I prediction through CP-ANN. This set of descriptors corresponds to the input X in Figure A.1

Nr desc.	Descriptors Name
1	Relative number of H atoms
2	Relative number of O atoms
3	Relative number of aromatic bonds
4	Relative molecular weight
5	Gravitation index (all pairs)
6	Average Information content (order 2)
7	Information content (order 2)
8	Average Bonding Information content (order 2)
9	XY Shadow / XY Rectangle
10	YZ Shadow / YZ Rectangle
11	ZX Shadow / ZX Rectangle
12	Molecular volume
13	Max partial charge for a C atom [Zefirov's PC]
14	Min partial charge for a C atom [Zefirov's PC]
15	Max partial charge for a O atom [Zefirov's PC]
16	Min partial charge for a O atom [Zefirov's PC]
17	Max partial charge for a H atom [Zefirov's PC]
18	Min partial charge for a H atom [Zefirov's PC]
19	Max partial charge (Qmax)
20	Min partial charge (Qmin)
21	Polarity parameter / square distance
22	DPSA-1 Difference in CPSAs (PPSA1-PNSA1) [Zefirov's PC]
23	FPSA-1 Fractional PPSA (PPSA-1/TMSA) [Zefirov's PC]
24	FNSA-1 Fractional PNSA (PNSA-1/TMSA) [Zefirov's PC]
25	PNSA-2 Total charge weighted PNSA [Zefirov's PC]
26	FNSA-2 Fractional PNSA (PNSA-2/TMSA) [Zefirov's PC]
27	WNSA-2 Weighted PNSA (PNSA2*TMSA/1000) [Zefirov's PC]
28	PNSA-3 Atomic charge weighted PNSA [Zefirov's PC]
29	WNSA-3 Weighted PNSA (PNSA3*TMSA/1000) [Zefirov's PC]
30	RPCG Relative positive charge (QMPOS/QTPLUS) [Zefirov's PC]
31	RPCS Relative positive charged SA (SAMPOS*RPCG) [Zefirov's PC]
32	RNCG Relative negative charge (QMNEG/QTMINUS) [Zefirov's PC]
33	RNCS Relative negative charged SA (SAMNEG*RNCG) [Zefirov's PC]
34	min(#HA, #HD) [Zefirov's PC]
35	count of H-acceptor sites [Zefirov's PC]
36	count of H-donors sites [Zefirov's PC]
37	HA dependent HDSA-1 [Zefirov's PC]
38	HA dependent HDSA-1/TMSA [Zefirov's PC]
39	HA dependent HDSA-2 [Zefirov's PC]
40	HA dependent HDSA-2/TMSA [Zefirov's PC]
41	HA dependent HDSA-2/SQRT(TMSA) [Zefirov's PC]
42	HA dependent HDCA-1 [Zefirov's PC]
43	HA dependent HDCA-1/TMSA [Zefirov's PC]
44	HA dependent HDCA-2 [Zefirov's PC]
45	HA dependent HDCA-2/TMSA [Zefirov's PC]
46	HA dependent HDCA-2/SQRT(TMSA) [Zefirov's PC]
47	HASA-1 [Zefirov's PC]
48	HASA-1/TMSA [Zefirov's PC]
49	HASA-2 [Zefirov's PC]

50	HASA-2/TMSA [Zefirov's PC]
51	HASA-2/SQRT(TMSA) [Zefirov's PC]
52	HACA-1 [Zefirov's PC]
53	HACA-1/TMSA [Zefirov's PC]
54	HACA-2 [Zefirov's PC]
55	HACA-2/TMSA [Zefirov's PC]
56	HACA-2/SQRT(TMSA) [Zefirov's PC]
57	HOMO-1 energy
58	HOMO energy
59	LUMO energy
60	LUMO+1 energy
61	HOMO - LUMO energy gap
62	Tot hybridization comp. of the molecular dipole
63	Tot dipole of the molecule
64	Image of the Onsager-Kirkwood solvation energy
65	PNSA-1 Partial negative surface area [Semi-MO PC]
66	DPSA-1 Difference in CPSAs (PPSA1-PNSA1) [Semi-MO PC]
67	FPSA-1 Fractional PPSA (PPSA-1/TMSA) [Semi-MO PC]
68	FNSA-1 Fractional PNSA (PNSA-1/TMSA) [Semi-MO PC]
69	PPSA-2 Total charge weighted PPSA [Semi-MO PC]
70	PNSA-2 Total charge weighted PNSA [Semi-MO PC]
71	DPSA-2 Difference in CPSAs (PPSA2-PNSA2) [Semi-MO PC]
72	PPSA-3 Atomic charge weighted PPSA [Semi-MO PC]
73	FPSA-3 Fractional PPSA (PPSA-3/TMSA) [Semi-MO PC]
74	FHDSA Fractional HDSA (HDSA/TMSA) [Semi-MO PC]
75	HASA H-acceptors surface area [Semi-MO PC]
76	FHASA Fractional HASA (HASA/TMSA) [Semi-MO PC]
77	HBSA H-bonding surface area [Semi-MO PC]
78	FHBSA Fractional HBSA (HBSA/TMSA) [Semi-MO PC]
79	HDCA H-donors charged surface area [Semi-MO PC]
80	FHDCA Fractional HDCA (HDCA/TMSA) [Semi-MO PC]
81	HACA H-acceptors charged surface area [Semi-MO PC]
82	FHACA Fractional HACA (HACA/TMSA) [Semi-MO PC]
83	HBCA H-bonding charged surface area [Semi-MO PC]
84	FHBCA Fractional HBSA (HBSA/TMSA) [Semi-MO PC]
85	HA dependent HDSA-2/TMSA [Semi-MO PC]
86	HA dependent HDSA-2/SQRT(TMSA) [Semi-MO PC]
87	HA dependent HDCA-1 [Semi-MO PC]
88	HA dependent HDCA-1/TMSA [Semi-MO PC]
89	HA dependent HDCA-2 [Semi-MO PC]
90	HA dependent HDCA-2/TMSA [Semi-MO PC]
91	HA dependent HDCA-2/SQRT(TMSA) [Semi-MO PC]
92	HASA-1 [Semi-MO PC]
93	HASA-1/TMSA [Semi-MO PC]
94	HASA-2 [Semi-MO PC]
95	HASA-2/TMSA [Semi-MO PC]
96	HASA-2/SQRT(TMSA) [Semi-MO PC]
97	HACA-1 [Semi-MO PC]
98	HACA-1/TMSA [Semi-MO PC]
99	HACA-2 [Semi-MO PC]
100	HACA-2/TMSA [Semi-MO PC]
101	HACA-2/SQRT(TMSA) [Semi-MO PC]
102	Max SIGMA-SIGMA bond order
103	Max SIGMA-PI bond order

104	Max PI-PI bond order
105	Max bonding contribution of a MO
106	Max antibonding contribution of a MO
107	Min valency of a O atom
108	Max valency of a O atom
109	Max bond order of a O atom
110	Min e-e repulsion for a O atom
111	Min e-n attraction for a O atom
112	Max atomic state energy for a O atom
113	Min e-e repulsion for a C atom
114	Max e-e repulsion for a C atom
115	Min e-n attraction for a C atom
116	Max e-n attraction for a C atom
117	Max e-n attraction for a H atom
118	Max atomic state energy for a H atom
119	Min resonance energy for a C-O bond
120	Max e-n attraction for a C-O bond
121	Min n-n repulsion for a C-O bond
122	Max n-n repulsion for a C-O bond
123	Min coulombic interaction for a C-O bond
124	Max coulombic interaction for a C-O bond
125	Min total interaction for a C-O bond
126	Max total interaction for a C-O bond
127	Min e-e repulsion for a C-C bond
128	Max e-e repulsion for a C-C bond
129	Min e-n attraction for a C-C bond
130	Max e-n attraction for a C-C bond
131	Min n-n repulsion for a C-C bond
132	Max n-n repulsion for a C-C bond
133	Min coulombic interaction for a C-C bond
134	Max coulombic interaction for a C-C bond
135	Min total interaction for a C-C bond
136	Max total interaction for a C-C bond
137	Min resonance energy for a H-O bond
138	Max resonance energy for a H-O bond
139	Min exchange energy for a H-O bond
140	Max exchange energy for a H-O bond
141	Min e-e repulsion for a H-O bond
142	Max e-e repulsion for a H-O bond
143	Min e-n attraction for a H-O bond
144	Max e-n attraction for a H-O bond
145	Min n-n repulsion for a H-O bond
146	Min e-e repulsion for a C-H bond
147	Max e-n attraction for a C-H bond
148	Min n-n repulsion for a C-H bond
149	Max n-n repulsion for a C-H bond
150	Min coulombic interaction for a C-H bond
151	Tot molecular 2-center resonance energy / # of atoms
152	Tot molecular electrostatic interaction
153	Tot molecular electrostatic interaction / # of atoms
154	1X GAMMA polarizability (DIP)
155	(1/6)X GAMMA polarizability (DIP)

Chapter four

The Metal Bonding Domain of the Antitumor Drug Fe(II)-bleomycin: a DFT Investigation
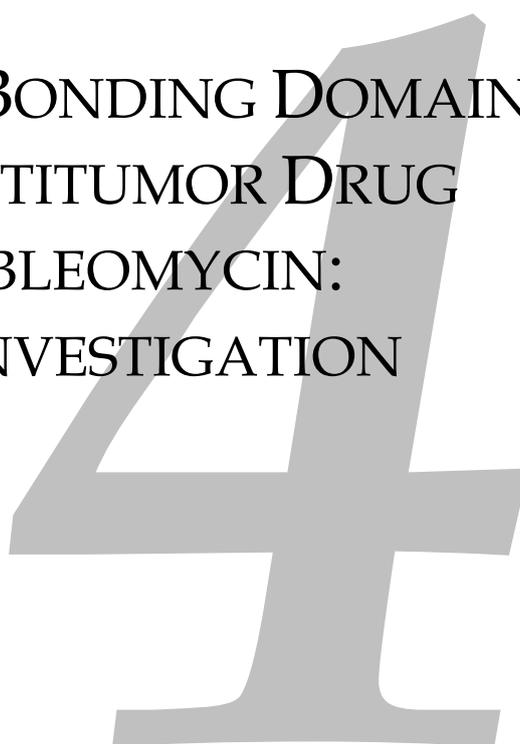
ABSTRACT

The geometric and electronic structure of ferrous complexes of bleomycin (Fe(II)BLM) has been investigated by means of Density Functional Theory calculations. The active site of this antitumor drug is a highly distorted octahedral complex with the coordination sphere completed by the five known endogenous ligands including pyrimidine, imidazole, deprotonated amide, and secondary and primary amine. We have addressed the controversial issue of the nature of the sixth axial ligand that we have identified as the oxygen of the carbamoyl group. Our conclusions are further validated by a comparison with structural data derived from NMR experiments. Moreover, because of the high sensitivity of structural data on the pH of the environment, we have investigated the effect of a different protonation state of the histidine amide on the geometric structure of the Fe(II)BLM complex. The extensive model of the active site of bleomycin considered in this work allows us to check the limitations of previous investigations based on simplified models.

4.1 Introduction

Bleomycin is a DNA-cleaving antibiotic, which is isolated as a copper complex from the culture medium of *Streptomyces verticillis* and administrated to patients in its metal-free form. BLM is important for clinical use in the treatment of head and neck cancer, certain lymphomas and testicular cancer [96]. Interest in the chemical activity of bleomycin began with the discovery that the presence of either reducing agents or hydrogen peroxide permits bleomycin to cleave DNA into simple sets of DNA degradation products [17]. Indications that metals and oxygen species were also required led to the demonstration that the presence of Fe(II) and O_2 was sufficient to allow BLM to degrade DNA. The first step is the formation of a high-spin Fe(II)BLM complex which reacts with O_2 [1]. In the second step, the so-called activated bleomycin (ABLM) – a peroxide-Fe(III)BLM complex – is produced and this then cleaves DNA. The products of DNA degradation are well known [17], but major gaps in our understanding remain where the drug activation and DNA degradation pathways intersect. A precise knowledge of the geometrical and electronic structure of the Fe(II)BLM complex is a prerequisite to understanding the activity of this drug in DNA degradation.

Unfortunately, no crystal structure is yet available for Fe(II)BLM. There are several different synthesized models mimicking BLM as a ligand and crystallized with various metals, for example copper [3, 97], cobalt [98], zinc [99] and even iron[100]. However, all models thus far lack both the bithiazole and the sugar residues. For example, the PMAH ligand has five nitrogen donor centers located in the primary and secondary amines, pyrimidine and imidazole rings, and the amide moiety (see Fig. 4.1). The ligand synthesized for iron is only tridentate, since the β-aminoalanine fragment is also not present. Therefore these crystallographic studies could not give a definitive answer as to whether the bleomycin is a five- or six-coordinated ligand and, if the sixth ligand is indeed present, whether it is endo- or exogenous. The exact coordination of the ligand has thus been a matter of dispute for several years.
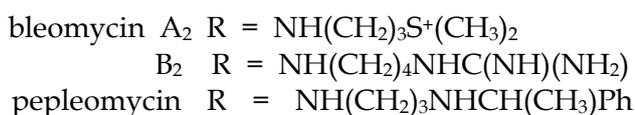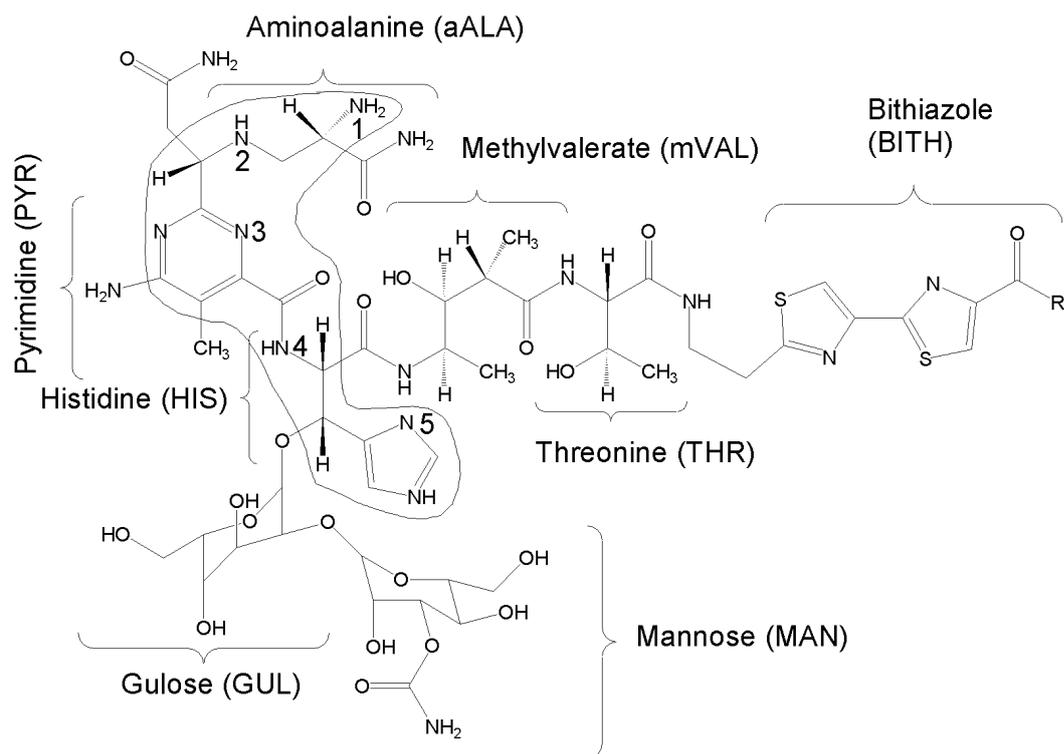
Figure 4.1 The structure of bleomycin A_2, B_2, and pepleomycin. The numbers from 1 – 7 indicate the possible coordination sites to the metal. The solid line shows the PMAH ligand (2-[[N-(aminoethyl)amino]methyl]-4-[N-[2-(4-imidazolyl)ethyl]carbamoyl] – 5 - bromopyrimidine) [3]

Most of the knowledge available on the structure of various metallo–BLM complexes in solution is derived from spectroscopic data and studies combining multinuclear NMR experiments and MD simulations [101-104]. Recently Lehmann [101] proposed a very reasonable structure for Fe(II)BLM based on the results of such studies. The proposed model assumes that the complex has a six-coordinate structure with only endogenous ligands. The equatorial ligands to the metal center are the secondary amine of the β-aminoalanine (A''aALA – N2) segment, the pyrimidine (PYR – N3), the imidazole ring (HIS-I – N5) and the amide nitrogen (HIS-A – N4) of the β-hydroxyhistidine. The primary amine of the β-aminoalanine residue (A'aALA – N1) and the nitrogen atom from the

carbamoyl group ($N_{carbamoyl}$ – N7) are the two axial ligands. This model is in agreement with other experimental work, where spectroscopic methods like optical absorption, circular diachronic and magnetic circular dichroism were applied to determine the structure of Fe(II)BLM, except that the oxygen atom in the carbamoyl group ($O_{carbamoyl}$ – O6) is proposed as the sixth ligand [103]. Therefore it is still unclear whether the nitrogen or the oxygen atom from the carbamoyl group is the second axial ligand.

The main goal of the present study is to resolve the second axial ligand and to determine the structure of the Fe(II)BLM complex using DFT calculations. It should be emphasized that the coordination sphere of the iron atom depends crucially on the total spin of the complex. This important electronic contribution should be taken into account when using a force field method, as *e.g.*, it has been recently done for Cu(II) complexes by adding a ligand field contribution to the potential [105]. The model investigated here by first principles DFT is a realistic and comprehensive model of Fe(II)-BLM, allowing us to check the limitations of previous studies on simplified models. In addition, we investigate the effect of protonation of the histidine amide nitrogen on the geometry of the complex. This aspect is important in the context of the long dispute over the inclusion of the amide nitrogen in the coordination sphere of the Fe(II)-BLM complex. The protonation state appears to be dependent strongly on the pH: The ^{1}H NMR experiment by Oppenheimer at pH 6.4 provided evidence for a protonated state of the histidine amide and thus excluded it as a ligand [106]. However, another NMR experiment by Akkerman at pH 7.0 showed that the histidine amide is deprotonated and bonded to the metal center under these conditions [107].

In the next section we describe in detail the models used for our studies and the computational method. In the third section we present our results and discuss them in comparison with available experimental data. The final section is devoted to conclusions.

4.2 Models and computational details

Figure 4.1 shows a schematic structure of a metal-free BLM ligand. We can distinguish two parts: A main part, which is isostructural for the whole bleomycin family and a long tail, which varies only in the terminal substituent denoted by "R". The main part is the metal bonding domain while the bithiazole tail is responsible for DNA docking [1]. We have investigated a model (Fig. 4.2) where only the bithiazole tail has been excluded since it is most likely irrelevant to the coordination environment of the metal domain of the complex. Thus, our model provides a realistic description which can be used to develop insight into the geometrical and electronic structure of the Fe(II)BLM complex. The starting structure was based on Co(II)BLM, generated from an NMR experiment [102], with the corresponding Co(II)–to–Fe(II) substitution of the metal center. Figure 3 shows the models used for this investigation. In Figure 3a we show the six-coordinated complex with all endogenous ligands. Throughout the chapter we define complex **A** as the model including in the coordination sphere the oxygen atom from the carbamoyl group. We call complex **B** the model
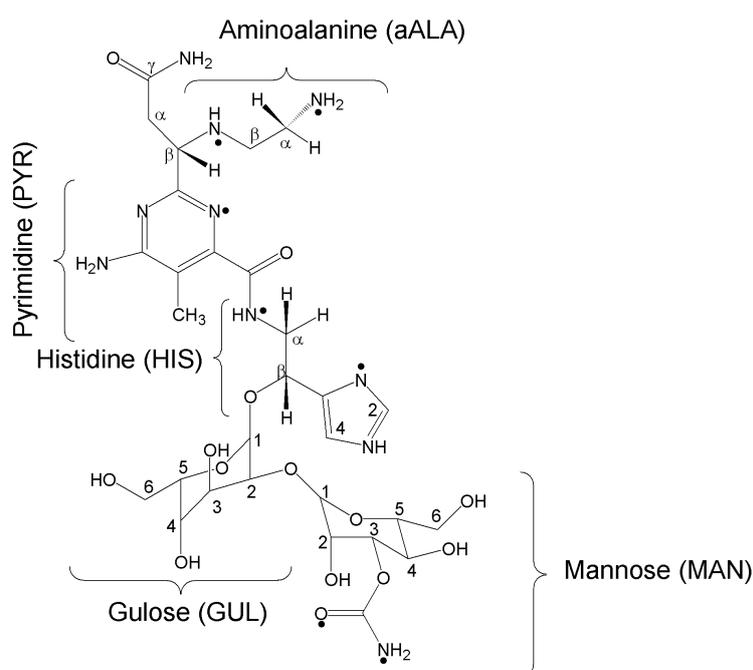

Figure 4.2 Model of the BLM metal bonding domain used in this study. Some of the atoms are indicated with numbers and Greek letters following the same notation as in Ref. [2], for a direct comparison.
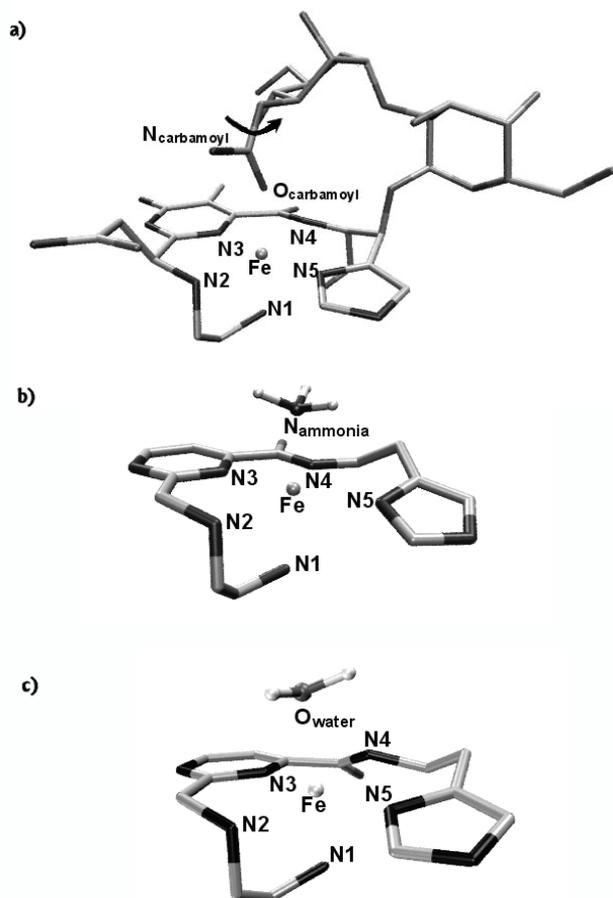
Figure 4.3 Models of the metal coordination center domain of Fe(II)BLM with different axial ligands:

a) the oxygen atom from the carbamoyl group (complex **A** and **AH**) and the nitrogen atom from carbamoyl group (complex **B** and **BH**). Index H in **AH** and **BH** indicates the protonated amide nitrogen (N4H). Complexes **B** and **BH** were created by rotating the carbamoyl group via the torsion angle pointed in the picture.

b) the nitrogen atom from the ammonia molecule replacing the disaccharide moiety (complex **C**, see text)

c) the oxygen atom from water molecule replacing the disaccharide moiety (complex **D**, see text)

The hydrogen atoms have been omitted for simplicity.

obtained from complex **A** by rotation of the carbamoyl group in such a way that the nitrogen atom becomes the sixth ligand. Since we are interested in the effect of the protonation of BLM on the geometry of the complex, we have considered also the complexes **AH** and **BH** generated from complex **A** and **B**, respectively, by protonation of the histidine amide nitrogen (N4). The index H indicates the protonated amide nitrogen (N4H).

In order to check the importance of the explicit inclusion of the original sixth ligand, we have also performed calculations on simpler models where the sugar moiety is replaced by either an ammonia molecule (complex **C**, Fig. 4.3b) or a water molecule (complex **D**, Fig. 4.3c). In this way we were able to check whether it is possible to answer the question about the second axial ligand based only on simplified models.

The DFT calculations were performed using the B3LYP functional, which was shown to predict the correct spin state for a number of different iron complexes

[108, 109]. We used the effective core potential basis set LanL2DZ [110, 111] for the iron atom and the 6-31G basis set for the other atoms as implemented in the program Gaussian03 [112]. This choice of the basis set is supported by the results of test calculations performed by Lehnert *et al.* on similar model systems [5]. Their conclusion is that the effective core potential basis set LanL2DZ is good enough to describe the geometry of ABLM and that a triple-ζ all electron basis set gives a comparable description.

When comparing our results with structural information derived by NMR data in solution, one should take into account that we are neglecting solvent effects in our model. Although we can expect small changes in the geometrical parameters due to the solvent, we think that our conclusions on the coordination shell will not be affected. Indeed, it is now established in the literature that solvent molecules are not participating in the coordination sphere of the Fe(II)-BLM complex [101, 113].

4.3 Results and discussion

For each system under consideration full geometry optimizations assuming low, intermediate or high spin states have been carried out. The final geometric parameters and electronic energies corresponding to singlet, triplet and quintet states are summarized in Table 4.1 and Table 4.2 for protonated and deprotonated complexes, respectively. Complex **AH** was found to have the lowest energy in its quintet electronic state and this energy is taken as the reference energy for other spin states of complexes **AH** and **BH**. The same spin order is observed for deprotonated models and the energy of complex **A** in its quintet electronic state is taken as the reference energy for other spin states of complexes **A** and **B**. Our investigation shows that the geometry of the drug complex is highly dependent on the environment of the iron ion and on its spin state.

4.3.1 Complexes A and AH

For complexes **A** and **AH** the Fe–$O_{carbamoyl}$ distance changes as a function of the spin state, reaching a maximum value for $S = 1$ (see Table 4.1 and 4.2). In the complex **A**, this value is equal to 3.26 Å for $S = 1$ indicating that the bond is broken. For $S = 2$ the Fe–$O_{carbamoyl}$ bond is again present with a length of 2.29 Å. Generally, the bond length between the iron center and its axial ligands ($O_{carbamoyl}$ and N1) in complex **A** are longer than in complex **AH** for each spin state (see Tables 4.1 and 4.2). This can be interpreted as a consequence of the deprotonation of the equatorial nitrogen donor ligand (N4) and the related shortening of the Fe–N4 bond length. It is known from ligand field theory, that strong equatorial ligands cause a tetragonal distortion, which corresponds to an extension along the z axis and compression on the x and y axes with a stabilizing effect on the d_{z^2} orbital. A Natural Bond Orbital analysis shows that the d_{z^2} orbital is already lower in energy than the $d_{x^2-y^2}$ orbital for the protonated complex. Deprotonation of the amide nitrogen (N4) causes further lowering of the energy of the d_{z^2} orbital.

Table 4.1 Relative energies and optimized geometric parameters of complexes **AH**, **BH**.

	Complex AH			Complex BH		
Total spin	0	1	2	0	1	2
ΔE[kcal/mol][a]	11.48	15.12	0.00	36.50	31.41	16.92
Fe-X[b]	2.12	2.25	2.12	2.37	3.70	4.21
Fe-N1	2.04	2.20	2.20	2.05	2.23	2.18
Fe-N2	2.10	2.24	2.26	2.11	2.13	2.28
Fe-N3	1.93	1.94	2.14	1.91	1.92	2.07
Fe-N4	2.10	2.17	2.59	2.11	2.12	2.43
Fe-N5	2.02	2.01	2.22	2.04	2.04	2.16
X-Fe-N1	176.9	171.6	159.4	173.9	173.3	166.1
2-3-4-5	-1.6	-3.7	3.8	0.8	-3.2	-1.8

[a] For complexes **AH** and **BH** the reference energy is the ground quintet electronic state energy of the complex **AH**

[b] X = $O_{carbamoyl}$, $N_{carbamoyl}$

All distances are in Å and angles in degrees

As the spin increases, each metal – equatorial ligand bond length generally increases and, in the particular case of N4 (complex **AH**), the distance increases up to 2.59 Å, while the average for the other equatorial ligands is 2.2 Å. Although this distance seems long, one can find in the Cambridge Structural Database structures of iron (II) complexes, where the distances between iron(II) and its nitrogen donor are 2.50 Å [114], 2.54 Å [115] and even 2.65 Å [116]. In additional, the analysis of the HOMO orbital for β electrons shows that some weak bonding interaction is present between these two atoms as indicated by overlap between the d_{z^2} orbital on iron and the lone pair on the N4 nitrogen atom. The same bond in complex **A** is less sensitive to a change in the spin (see Table 4.1). These results explain why it has been difficult to establish which atoms participate to the iron coordination shell. Hilbers *et al.* [107, 117] have proposed that BLM binds the metal through positions N2, N3, N4, N5, N7 while Openheimer and Hecht [106] proposed N1, N2, N3, N5, N7 thus excluding the N4 atom. These two studies used different experimental conditions including a different pH = 7.0 and pH = 6.4, respectively. Our study clearly demonstrates that the problem of the ligand assignment can be related to the very weak bonding interaction between the N4 atom and the metal

Table 4.2 Relative energies and optimized geometry parameters of complexes **A**, **B**, **C**, **D**.

	Complex A			Complex B			Complex C			Complex D		
Total spin	0	1	2	0	1	2	0	1	2	0	1	2
ΔE [kcal/mol][a]	12.29	6.96	0.0	32.83	26.10	23.78	5.34	8.00	00.0	4.05	7.17	0.00
Fe-X[b]	2.17	3.26	2.29	2.29	3.93	4.02	2.07	2.28	2.26	2.07	2.32	2.22
Fe-N1	2.03	2.26	2.22	2.04	2.27	2.19	2.06	2.29	2.26	2.03	2.31	2.25
Fe-N2	2.16	2.13	2.40	2.14	2.15	2.30	2.17	2.35	2.48	2.15	2.25	2.42
Fe-N3	1.92	1.92	2.13	1.90	1.91	2.07	1.91	1.93	2.14	1.91	1.91	2.13
Fe-N4	1.97	1.94	2.09	1.98	1.94	2.03	1.99	1.99	2.07	1.99	1.97	2.06
Fe-N5	2.01	2.03	2.22	2.01	2.03	2.18	2.03	2.01	2.22	2.03	2.02	2.20
X-Fe-N1	170.3	126.2	140.8	176.8	168.4	166.2	179.3	165.6	161.9	177.9	157.8	155.3
2-3-4-5	7.0	9.4	14.7	5.9	7.0	9.4	5.5	5.3	4.5	-5.1	-5.6	-8.8

[a] For complexes **A** and **B** the reference energy is the ground quintet electronic state energy of the complex **A**, which has the value E_{ref} = -2968.6627 a.u. For complexes **C** and **D** the reference energies are the quintet electronic ground state energies, which have the values E_{ref} = -1144.3408 a.u. and E_{ref} = -1164.1943 a.u., respectively.
[b] X = $O_{carbamoyl}$, $N_{carbamoyl}$, $N_{ammonia}$, O_{water} for complex **A**, **B**, **C**, **D**, respectively.
All distances are in Å and angles in degrees

center when the pH is too low for the histidine amide to be deprotonated.

In the original model of the Co(II)BLM complex arising from NMR experiments combined with force field molecular dynamics simulation, the equatorial ligands are lying almost in the same plane [102]. We observe similar small distortions from planarity only for complex **AH** (see 2-3-4-5 dihedral angle in Table 4.1). The largest distortion is observed for the deprotonated complex **A** in its quintet ground state, with a dihedral angle N2–N3–N4–N5 of 14.7°. The smallest distortion occurs for $S = 0$ (see Tables 4.1 and 4.2). Thus we conclude that the degree of distortion is highly dependent on the spin state of the complexes.

We also observe a large change in the valence angle of $O_{carbamoyl}$–Fe–N1 for different spin states (Tables 4.1 and 4.2). For complex **A**, the angle changes from 170.3° for $S = 0$ to 140.8° for $S = 2$. This angle distortion together with the distortion from planarity discussed above brings the complex into a highly distorted octahedral geometry. Because of these modifications, the Fe–$O_{carbamoyl}$ bond becomes weaker and more susceptible to exchange with an oxygen molecule to produce the $Fe(II)O_2BLM$ complex. This complex then accepts an additional electron to form the activated BLM species. In previously proposed models for activated bleomycin based on force field methods, the Fe–N1 bond was not present and the oxygen molecule entered the coordination sphere from the β-aminoalanine site [102, 118]. However the crystal structure for copper - BLM [119] and some NMR studies for other metallo–BLM [4], suggest a different structure in agreement with our results. Our results strongly support the picture in which the atoms N1, N2, N3, N4, N5 remain in the coordination sphere of the metal center and the flexible sugar moiety drifts away making space for an exogenous ligand.

4.3.2 Complexes B and BH

Complex **BH** also has the lowest energy for total spin $S = 2$, but this energy is still higher than the energy of complex **AH** for all spin states and particularly it is about 17 kcal/mol higher than the reference energy, which is the ground state of complex **AH** for $S = 2$. For complex **BH** the length of the Fe–$N_{carbamoyl}$ bond also increases with increasing total spin (see Table 4.1). Specifically, for $S = 2$ the

distance becomes 4.21 Å. This leads us to conclude that the $N_{carbamoyl}$ atom moves out of the coordination sphere of the iron atom. We notice that the breaking of the Fe–$N_{carbamoyl}$ bond also occurs in the **B** complex with the anionic ligand. Therefore, the behavior of the Fe–$N_{carbamoyl}$ bond is not related to the specific protonation state of the ligand, but rather to steric interaction. For $S = 0$ the distance between the $O_{carbamoyl}$ atom and the $C\gamma$ atom from the pyrimidine residue ($C\gamma$PYR) is equal to 2.89 Å, while the sum of the van der Waals radii of these atoms is 3.10 Å. When the Fe–$N_{carbamoyl}$ distance increases, the $O_{carbamoyl}$–$C\gamma$PYR distance increases as well, minimizing in this way the steric interaction. However, the presence of the proton affects the Fe–N4 bond. In particular for $S = 2$ the Fe–N4 distance changes from 2.43 Å in **BH** to 2.03 Å in the complex **B**. Interestingly, the geometries of complexes **A** and **B** are very similar for the lowest spin state $S = 0$, both having a six-fold coordination shell. It is only when we go to the high spin state $S = 2$ that is the experimentally observed we see major changes in the coordination shell. This strong dependence of the final geometry on the spin state of the complexes is an important finding that should be taken into account when describing these systems in terms of force field modeling [105].

4.3.3 Complexes C and D

We have also investigated simplified models using an ammonia molecule (complex **C**) and a water molecule (complex **D**) to represent the nitrogen and the oxygen atoms of the carbamoyl group, respectively, to validate the importance of the inclusion of the big sugar moiety explicitly in the model. We find that the quintet electronic spin state is the ground state for both complexes **C** and **D**, which is in agreement with the experimentally observed spin state for the Fe(II)BLM complex. However, the differences between the energies of the different spin states and the reference energies are only a few kcal/mol (see Table 4.2) and we cannot directly compare energies of complex **C** and **D**. Both complexes form a distorted octahedral geometry, providing no grounds to distinguish between them. The Fe–$N_{ammonia}$ bond length increases slightly with increasing spin, but we have not observed the bond breaking that occurs when the

actual sugar moiety is included, in the model complex **B** (Table 4.2). By comparing the ground state structures of complex **C** and **D** we do not observe a large difference in the X–Fe–N1 covalent angle, in contrast with the complexes **A** and **B**. This suggests that simplified models are not appropriate for resolving the issue of the sixth ligand of the Fe(II)BLM. Moreover, it also indicates that the disaccharide moiety plays a very important role in the metal bonding domain of BLM, showing that performing DFT calculations with the more realistic model is crucial.

4.3.4 Comparison with experiment

A quintet electronic ground state for Fe(II)BLM complex was established many years ago using EPR and Mösbauer spectroscopy [120-122]. Our B3LYP - DFT calculation is able to reproduce this spin state correctly. This is important since previous DFT calculations could not reproduce the proper experimental spin state [123]. In Ref. [123] the calculation were carried out using B3LYP functional, as we have done here, but on a simpler model of bleomycin with only a pentadonor ligand, and with a doubly deprotonated ligand at the histidine amide N4 and the N2 of the secondary amine of the β-aminoalanine. The concordance between the observed spin state and the one computed here emphasizes again the importance of using a realistic model including all six potential donor sites.

We wish now to compare our optimized geometries with the available NMR data. Table 4.3 summarizes the comparison between the iron – proton DFT optimized distances for **AH**, **A**, **BH** and **B** complexes in their electronic ground state with the Fe–H distances based on 2D NMR studies of the paramagnetic complex Fe(II)BLM [2]. The experimental iron – proton distances were obtained from relaxation times of the paramagnetically shifted protons. The calculated RMSD between experimental and theoretical data is 0.8 Å for both protonated models and 0.6 Å and 0.5 Å for deprotonated complexes **A** and **B**, respectively. However, if we consider only the mannose sugar moiety that is directly connected to the carbamoyl group, the calculated partial RMSD for complex **AH** is 0.8 Å while for **BH** it is equal to 1.2 Å. For their deprotonated analogues these values are 0.5 Å and 0.8 Å, respectively. This implies that the structures of complexes **AH** and **A**

Table 4.3 Comparison of our theoretical results and experimental Fe-H distances [Å] from NMR data

Assignments[a]	Experimental data[b]	$S = 2$ Complex		$S = 2$ Complex	
		A	AH	B	BH
aALA C$^\alpha$H	3.7	3.3	3.3	3.3	3.4
½ aALAC$^\beta$H$_2$	4.1	4.2	4.1	4.1	4.1
½ aALAC$^\beta$H$_2$	3.4	3.5	3.4	3.4	3.3
PYR C$^\beta$H	< 3.1	4.1	3.9	3.9	3.7
½ PYR C$^\alpha$H$_2$	4.8	5.0	4.9	5.0	5.1
½ PYR C$^\alpha$H$_2$	4.6	4.0	3.7	5.0	5.0
HIS C$^\alpha$H	3.6	4.1	4.6	4.0	4.2
HIS C$^\beta$H	4.4	4.6	4.9	4.6	4.6
HIS C2H	< 3.2	3.4	3.2	3.4	3.3
HIS C4H	5.0	5.4	5.5	5.3	5.4
GUL-1	5.5	4.9	5.6	5.8	6.2
GUL-2	6.8	6.8	7.1	7.2	7.8
GUL-3	7.0	7.5	7.3	6.8	7.6
GUL-4	7.4	8.2	8.1	7.0	7.7
GUL-5	5.4	6.7	6.7	5.6	6.2
GUL-6	7.4	8.3	8.6	7.7	8.0
GUL-6	7.4	8.5	8.9	8.3	8.8
MAN-1	7.1	7.0	7.2	7.7	8.4
MAN-4	7.5	7.0	5.8	7.4	7.4
MAN-5	5.8	5.0	5.0	6.3	6.8
MAN-6	7.7	7.3	7.3	8.8	9.1
MAN-6	7.8	7.5	7.5	9.0	9.3
Total RMSD[c]	-	0.6	0.8	0.5	0.8
Partial RMSD[d]	-	0.5	0.8	0.8	1.2

[a] Notation has been taken from Ref. [2] (see Fig. 4.2)

[b] Experimental data are from Ref. [2]

[c] RMSD calculated with the formula: $\sqrt{\dfrac{\sum\left(D_{exp} - D_{theor}\right)^2}{N}}$, where D_{exp} and D_{theor} are the proton – metal distances derived from NMR data and our model, respectively; N is the number of data points

[d] RMSD only for mannose sugar moiety ($N = 5$)

provide a better fit to the experimental data than complexes **BH** and **B**, with respect to resolving the sixth ligand. Indeed we find the smallest RMSD value for the deprotonated models, corresponding to the experimental conditions of pH = 6.7 in the NMR study. The deprotonation of the N4 equatorial ligand influences the position of the axial O$_{carbamoyl}$ ligand. The Fe–N4 distance decreases while at the same time the Fe–O$_{carbamoyl}$ distance increases (Tables 4.1 and 4.2). This provides a better fit to the experimental data. In additional, Loeb *et al.* concluded from their experimental work that the Fe(II)BLM complex is six-coordinate with at least five endogenous ligands and the sixth ligand being either the O$_{carbamoyl}$ substituent of the mannose sugar or a solvent molecule in this axial position [103].

Their study did not determine whether the $O_{carbamoyl}$ atom is directly coordinated to the metal or connected through H–bonding. Based on our results and the information from experimental work of Lehmann we can now say that the $O_{carbamoyl}$ atom is directly connected to the iron center, though in a highly distorted octahedral configuration.

Our findings are also in agreement with very recent data on Fe(II)-bleomycin obtained by X-ray absorption near edge structure spectroscopy [113]. In this work the best agreement between theoretical and experimental spectra is achieved for the bleomycin model complex with the primary amine and oxygen of the mannose sugar occupying the axial positions. They also observed that the coordination environment is characterized by serious distortions of the iron octahedron.

4.4 Conclusions

The DFT calculations have been performed on an extensive model of the Fe(II)BLM complex based on available NMR data [2, 102]. By comparing the model complex **A** and **B**, distinguished by having the oxygen or the nitrogen atom from the carbamoyl group as the sixth ligand, respectively, we can conclude that the oxygen atom is coordinated to the metal center. We find that complex **A** besides being energetically more favorable also gives a better fit to the experimental data [2, 102, 103, 120]. We observe that in complex **B** the carbamoyl group tends to move out of the coordination shell in the high spin configuration contrary to experimental evidence. We suggest that this behavior can be explained in terms of steric hindrance of the carbamoyl oxygen with the pyrimidine $C\gamma$. The energetically most stable complex **A** in the high spin configuration shows a highly distorted octahedral conformation, with an angle between axial ligands of 140.8º. This fact should facilitate the exchange of the $O_{carbamoyl}$ ligand with the oxygen molecule during the formation of ABLM.

We have investigated the effect of a different protonation state of the histidine amide on the geometric structure of the Fe(II)BLM complex. We observe that deprotonation of that equatorial ligand mainly influences the Fe–N4 distance, which is very long in the high spin configuration of complex **AH** and **BH**. It also affects the positions of axial ligands, *e.g.* the Fe–$O_{carbamoyl}$ distance increases from 2.12 Å in complex **AH** to 2.29 Å in complex **A** resulting in a better fit to experimental NMR distance estimates. The protonation state has no effect on the energetic spin order of our model complexes, nor does it influence the behavior of the carbamoyl group in complexes **B** and **BH**.

The study carried out with the simplified models shows that they are not appropriate for resolving the issue of the sixth ligand of the Fe(II)BLM. The results also indicate that the disaccharide moiety plays a very important role in the final structure of the metal bonding domain of BLM, showing that the use of a more realistic model in our DFT calculations is crucial. We also underline the importance of a theoretical method which directly takes into account the total spin of the system. This work provides a good starting point for further studies of the reaction mechanism of ABLM. One interesting issue is the self-inactivation of the drug in the absence of DNA. The activation of the O-O bond of the peroxide – the sixth, external ligand of ABLM – plays the main role in that process.

Chapter five

The Mechanism of the Bleomycin Suicide:
A Car-Parrinello Molecular Dynamics Investigation
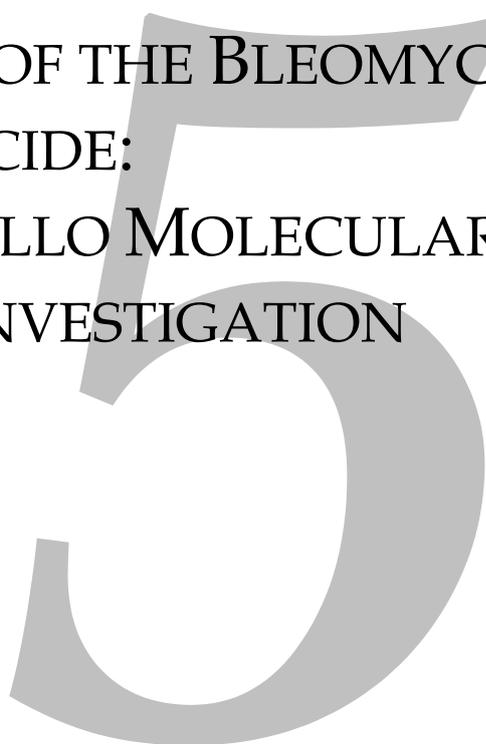
ABSTRACT

Using first-principles molecular dynamics simulations (Car-Parrinello method) we investigated the possible reaction pathways for decay of the active bleomycin-Fe(III)-OOH complex, the so-called bleomycin suicide. The theoretical model of activated bleomycin contains the whole metal bonding domain of the bleomycin ligand. Simulations performed both in vacuum and in water solution show that a facile decaying process involves a homolytic O-O bond cleavage with an almost simultaneous hydrogen atom abstraction. The formation of an intra- or intermolecular hydrogen bond appears to be crucial for the decay of the activated bleomycin. We did not observe any evidence of heterolytic cleavage of the O-O bond of the Fe(III)-OOH species.

5.1 Introduction

Bleomycin is a sixendogenous ligand, which is used as antitumor drug (Blenoxane) in the treatment of several types of cancer since the early eighties [17]. This drug is administrated to the patient in the metal free form. A species called activated bleomycin is formed in the body and it is identified as a low spin BLM-Fe(III)-OOH complex by electrospray ionization mass spectroscopy (ESI-MS) [124], extended X-ray absorption fine structure (EXAFS) spectroscopy [125], EPR spectroscopy [121], and Mössbauer spectroscopy [120]. The iron atom is coordinated by five nitrogen atoms coming from the bleomycin ligand (the secondary amine of the β-aminoalanine (A''aALA – N2) segment, the pyrimidine (PYR – N3), the imidazole ring (HIS-I – N5), the amide nitrogen (HIS-A – N4) of the β - hydroxyhistidine, and the primary amine of the β - aminoalanine residue (A'aALA – N1) as the axial ligand; see Fig. 1.1) and one end-on peroxide anion (OOH$^-$) as the second axial ligand. The activation of oxygen by transition-metal oxides is a fundamental mechanism that is of importance in organic synthesis, catalysis and biochemistry. ABLM is one natural example of such activation, similar to cytochrome P450 or methane monooxygenase [126]. The drug acts through peroxide group cleavage, which is believed to be the first step driving the DNA double strand degradation and *ipso facto* the disruption of the cancer cell. Figure 5.1 illustrates a well accepted cycle of BLM activation and deactivation which is broadly discussed in the review paper by Burger [1, 17]. It is interesting that the same products can be formed from DNA when BLM is activated with $Fe^{2+} + O_2$ or $Fe^{3+} + H_2O_2$. That is because both sets of activation cofactors yield the same active state BLM-Fe(III)-OOH as it is shown in Figure 5.1. As it appears from the presented cycle, ABLM turns into Fe(III)BLM after causing the oxidative damage of DNA and the Fe(III)BLM complex can be transformed into Fe(II)BLM and later ABLM in the presence of a reducer and oxygen. This reaction pathway has been verified experimentally [127]. However, the initial step in the conversion of ABLM into Fe(III)BLM can proceed both in the presence or absence of DNA

[121]. When ABLM decays in the absence of DNA, an irreversible change in BLM occurs. It undergoes a conversion to one or more Fe(III) complexes unable to attack DNA, even when aerobic Fe(II) is subsequently added [127]. The reaction is called activated bleomycin suicide because ABLM is a kinetic component in this reaction [121]. The suicide chemistry is very complex [128], yielding several products [124] and likely involving a modification of the bithiazole moiety [129].

Studies of the ABLM decay process apart from DNA may provide information about the formation and nature of the BLM intermediate species that may be responsible for the side effects of the drug. On the other hand the characterization of the suicide process may highlight some relevant aspects for DNA degradation by ABLM. One of the most common hypothesis for the initial step in the conversion of ABLM into Fe(III)BLM is the reaction analogous to that observed in cytochrome P450 with a heterolytic O-O bond cleavage [17] leading to a BLM-Fe(V)=O complex as intermediate state. However, an alternative scenario is a Fenton-like reaction, where a homolytic cleavage of the O-O bond takes place. In this case a hydroxyl radical and an intermediate iron complex BLM-Fe(IV)=O are formed[130]. This Fe(IV)=O complex is much more stable than the Fe(V)=O complex mentioned above, as shown by theoretical calculations [131]. The
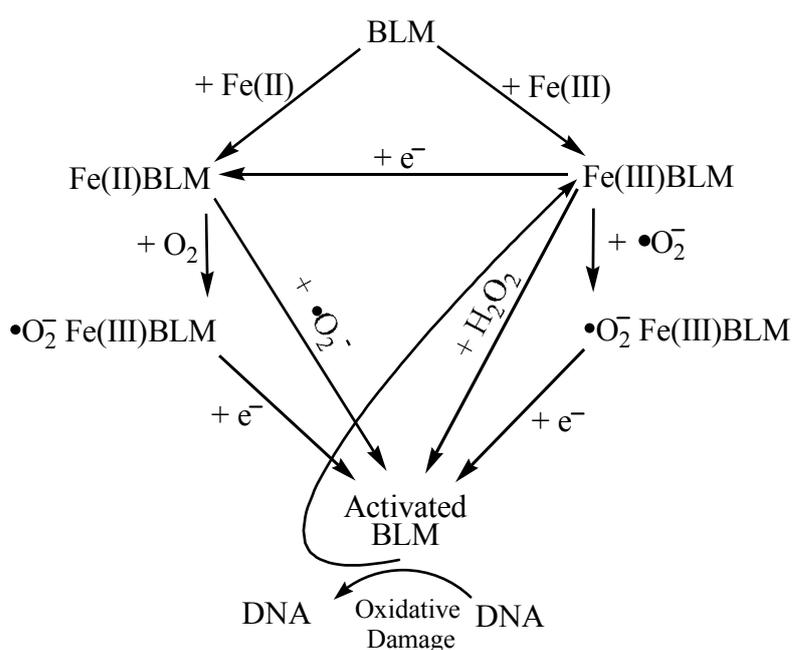

Figure 5.1 Activated bleomycin formation pathways adapted from Ref [1]. The ABLM can be formed both from Fe(II) and Fe(III) bleomycin complexes. During the degradation of ABLM in the absence or presence of DNA a Fe(III) complex is formed. Formally, the F(III)BLM complex can be transformed into F(II)BLM and ABLM as shown.

question we address here is which activation mechanism of the O-O bond plays a role in the self-inactivation of bleomycin. A similar problem is related to the mechanism of action of ABLM on DNA molecule that is subject of the study presented in the next chapter of the thesis.

Some experimental studies on the decay of ABLM in the absence of DNA suggest a heterolytic cleavage of the Fe(III)O-OH bond [132, 133]. In one of these studies a synthetic model for the metal bonding site of BLM has been treated with Fe(III)-H_2O_2 or Fe(II)-O_2 aiming to resolve the mechanism of the self-deactivation of ABLM. A breakdown at the N2 of the amino acid side chain of the metal binding site has been demonstrated [132]. Another study for the same purpose but with N-[bis(2-pyridylmethyl)aminoethyl]pyridine-2-carboxamide (HPaPy$_3$) as the ligand modeled BLM, demonstrated a modification at the carbon Cα to the amide (N1 in Fig. 5.2). The modification at this site was suggested to be the result of a two-electron oxidation of the ligand following the heterolytic cleavage of the O-O bond of the Fe(III)-OOH species [133]. However, studies performed by Padbury *et al.* [134] involving the interaction of 10 – hydroperoxy – 8,12 – ocatadecadienoic acid with Fe(II)BLM suggest that the homolytic O-O bond cleavage of the peroxide intermediate of ABLM can produce one equivalent of a hydroxyl radical in addition to the (•OH)-Fe(III)-BLM, which may be the DNA damaging species. Some other investigators in the field showed that a hydroxyl radical can be generated by Fe(II)BLM and proposed that these active radical species may be responsible for degradation of both ABLM and DNA [135, 136]. The Fe(II)-BLM mediated production of a hydroxyl radical was based on ESR spin-trapping techniques. The experimental models are therefore to some extent contradictory and do not provide a clear answer to how the O-O bond is activated in the process of BLM self-destruction.
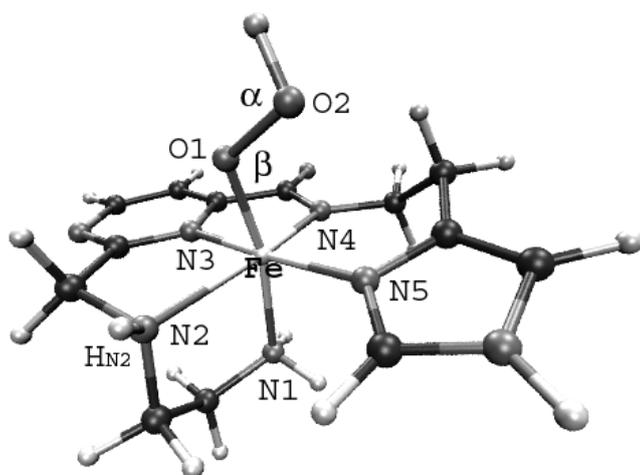
Figure 5.2 The optimized structure of the ABLM model used in this work. The five endogenous ligands of bleomycin (N1, N2, N3, N4, N5) and the hydroperoxo group (O1, O2, $\alpha = \angle$ H-O2-O1, $\beta = \angle$ O2-O1-Fe) are explicitly labeled. Additionally, the hydrogen atom H_{N2}, which is observed to be abstracted during the CPMD simulations in vacuum of ABLM, is also marked.

In this chapter we present a theoretical investigation of the possible mechanisms of the initial step of the ABLM suicide based on DFT calculations and Car-Parrinello molecular dynamics simulations. This is the first application of first principles molecular dynamics to the problem of the mechanism of ABLM self-inactivation that allows for an unbiased analysis of the initial chemical reaction steps, *i.e.* without a predefined reaction pathway.

5.2 Models and computational methods

Figure 1.1 shows a schematic structure of a metal-free BLM ligand. In our computational model the metal bonding domain is included without any modification. The bithiazole tail and the sugar moiety involved in DNA recognition and binding are neglected in our investigation. Their absence does not modify the coordination sphere of the metal that is relevant in the process of O-O bond activation. The presence of a bithiazole tail may be important in later steps of the ABLM suicide, beyond the scope of our present investigation. A realistic model of the coordination sphere of the metal is significant for our study since the reactivity of the Fe(III)-OOH species strongly depends on the nature of the metal

center ligands. For instance, in the case of the porphyrin ring ligand the reaction of the O-O bond breaking has a heterolytic character and the ligand stabilizes the peroxidase compound I [137, 138]. On the other hand, previous CPMD simulations on the $Fe^{III}(H_2O)_5$-OOH complex have shown that the complex is very stable and no O-O bond cleavage was observed [130]. Static theoretical calculations on the thermodynamics of the O-O bond homolytic cleavage also show that the energy barrier drops from 50 kcal/mol for the $(NH_3)_5$-Fe(III)-OOH complex to about 30 kcal/mol for the $(NH_3)_2$(Py)Fe(III)-OOH complex [5]. Our previous work on the structure determination of Fe(II)BLM [139] also highlights the importance of an accurate model in the theoretical investigation of bleomycin complexes.

The starting structure was based on NMR data for the Co^{III}BLM-OOH complex [4], with the corresponding Co(III) – to – Fe(III) substitution of the metal center. The model shown in Figure 5.2 was studied both in vacuum and in a water solvent. The calculations in vacuum allow to check the stability of the system and to observe the initial step in the ABLM suicide without any external perturbation or reactant.

The comparison between the simulations in vacuum and in water enabled us to check the influence of the water solvent in stabilizing the O-O bond of ABLM and to check the possible differences in the mechanism of bond cleavage in the initial step of ABLM self degradation in the gas phase and in aqueous environments.

The calculations were performed using the Car-Parrinello molecular dynamics method [47] as implemented in the CPMD code [63]. The Kohn-Sham orbitals are expanded in a PW basis set with a kinetic energy cutoff of 80 Ry. The test calculations described in detail in the next section, show that this cutoff is sufficient for achieving a good convergence of energies and structural properties for the considered systems. We employed *ab initio* norm-conserving pseudopotentials, generated within the Troullier-Martins scheme [140]. A generalized gradient approximation was used for the exchange-correlation functional, following the prescription of Becke and Perdew [53, 54]. All simulations are spin polarized with a total spin $S = 1/2$. Room temperature Car-

Parrinello MD simulations were performed by using a time step of 0.09 fs and a value of 400 a.u. for the fictitious electronic mass in the Car-Parrinello Lagrangian. For the simulation in vacuum we have used an isolated cubic supercell of size 17 x 17 x 17 Å^3. The simulation in water was performed using a hybrid quantum mechanics – molecular mechanics approach [29, 30]. In particular, the hybrid QMMM scheme used here is based on the CPMD method for the quantum part and on the Gromos96 MD program [38] for the classical molecular mechanics part. The quantum mechanics part includes the model of ABLM and 6 water molecules in a cubic box of 15 x 15 x 15 Å^3. The excess water molecules were described using the SPC model [141] with the force field implemented in the Gromos96 program. To prepare the initial configuration of ABLM in water we have performed a classical MD simulation using the Gromos96 program, in which the coordinates of the solute molecule were fixed while the 308 water molecules in the box have been relaxed. First a constant pressure simulation at 300 K and 1 bar was performed for 2 ns and then a constant volume simulation was performed for 1 ns. After relaxation of the solvent molecules, 6 waters were found within a distance of 3 Å from the –OOH ligand. These molecules were included in the quantum part of the system.

Both simulations in vacuum and in water start with 1000 steps of relaxation of the system at 100 K following by a slow temperature increase to 300 K.

5.3 Results

5.3.1 Test calculations

To validate the choice of the density functional, the kinetic energy cutoff of PW, and the pseudopotentials, a series of test geometry optimizations was performed with different programs using localized basis sets. In particular, we used the Amsterdam Density Functional code [142], in which the Kohn-Sham orbitals are expanded in a Slater-type orbital basis set, and the Gaussian03 code [112] using Gaussian-type orbitals. The model of ABLM shown in Figure 5.2 was optimized by: i) CPMD with BP functional and 80 Ry cutoff; ii) ADF with BLYP and BP

Table 5.1 The comparison of a few selected parameters of the optimized geometry of ABLM model (see Fig. 3) with different methods: the CPMD/BP calculations were performed with 80Ry cutoff; the ADF/BLYP and ADF/BP with triple zeta basis set with polarization functions; the Gaussian calculations with LanL2DZ basis set.

	O1-O2	O2-H	Fe-O1	Fe-N1	Fe-N2	Fe-N3	Fe-N4	Fe-N5	α	β	O1-Fe-N2	O1-Fe-N5	H-O1-O2-Fe
CPMD/BP	1.51	0.98	1.77	2.07	2.15	1.87	1.89	1.97	97.0	117.2	86.6	94.4	160.0
ADF/ BLYP	1.48	0.97	1.78	2.09	2.17	1.89	1.91	1.99	97.9	118.0	87.2	94.4	165.7
ADF/ BP	1.46	0.98	1.77	2.07	2.15	1.89	1.89	1.97	97.4	117.2	87.0	94.3	165.8
G03/ B3LYP	1.53	0.98	1.81	2.09	2.16	1.91	1.91	1.99	98.7	116.2	86.6	94.9	157.8
*G94/ B3LYP	1.50	0.99	1.83	2.00	2.00	1.99	1.90	2.01	102.0	117.6	67.6	90.4	-118.2

* The parameters of simplified model used in Ref [5]. The second axial ligand (N1) and one of the equatorial ligands (N2) are nitrogen atoms of ammonia molecules.

functionals with small core triple zeta plus a polarization functions basis set; iii) Gaussian03 with hybrid B3LYP functional and the effective core potential basis set LanL2DZ [110, 111]. In addition the results were compared with a simplified model of ABLM used in a previous DFT calculation [5]. The results are presented in Table 5.1. Relevant geometry parameters are very similar, and the largest differences appear between the parameters describing the simplified model of ABLM from the previous DFT calculations and our model. The Fe-N1 and Fe-N2 bonds lengths in particular are shorter than in model used here due to the presence of two ammonia molecules in the simplified model. These also affect the distance between the iron center and the pyrimidine ring (Fe – N3), which is about 0.1 Å longer in the simplified model, since here the pyrimidine ring is not tightly embedded within the whole ligand structure. There are some dissimilarities between the Fe-O1 bond length computed by the Gaussian program with LanL2DZ (~ 1.82 Å) and the other procedures (~ 1.77 Å), but more likely that is due to the small basis set that is used in Gaussian/B3LYP calculations. For a proper description of oxygen atoms a larger basis set is needed. It has been pointed out recently that not all functionals are appropriate for iron complexes to reproduce the proper ground state spin polarization [109]. Thus a

set of test calculations was performed to verify the accuracy of the BP functional for our systems. In all tested compounds (Fe^{II}(BLM)-O_2, $S = 0$; Fe^{IV}(BLM)=O, $S = 1$; Fe^{III}(BLM)-OOH, $S = 1/2$) the calculated spin ground state was in agreement with experimental data present in literature [17].

We have also checked the influence of the value of the PW energy cutoff on our system. The test calculations were made for the Fe^{III}(NH_3)$_5$-OOH complex, the HOOH molecule and the OOH anion changing the cutoff value from 70 Ry up to 100 Ry in steps of 10 Ry. The changes of the structural parameters when the energy cutoff is increased from 70 to 80 Ry are significant, in particular for the Fe-O and O-O bonds. With 70 Ry the O-O bond is 1.60 Å in the Fe^{III}(NH_3)$_5$-OOH complex and 1.56 Å in the OOH^- anion, which is too long and makes the bond weaker and predisposed to scission. Increasing the cutoff from 90 Ry to 100 Ry does not change the O-O distance, which converges to 1.50 Å and 1.47 Å for the Fe^{III}(NH_3)$_5$-OOH complex and the OOH^- anion, respectively. These results justify the choice of computational parameters for the CPMD simulations.

5.3.2 ABLM in vacuum

After the initial relaxation of the system the trajectory was followed for a total time of 290 fs at room temperature. Figure 5.3 shows the behavior of a few selected bond distances (Fe-O1, O1-O2, O2-H_{N2}) along the dynamics. By analyzing the O2 – H_{N2} distance (dashed line), three parts in the dynamics can be identified. In the first part (I) the –OOH group is rotating around the Fe-O1 axis. The distance between the O2 and the hydrogen atom bound to the secondary amine of the β-aminoalanine A''aALA – N2 is decreasing, while the O1-O2 bond is stable with an average bond length of 1.57 Å. After 100 fs the O2 oxygen atom forms a hydrogen bond with the H_{N2}. In the second part of the dynamics the intramolecular H-bond facilitates the O1-O2 bond cleavage of the hydroperoxo group (II in Fig. 5.3). After 50 fs from the moment when the H-bond was created, the O1-O2 bond starts increasing from the initial value of 1.58 Å and reaches a value of ≈2.00 Å at 200 fs. At the same time the Fe-O1 distance drops to 1.60 Å and oscillates around this shorter value. The just formed OH radical abstracts almost
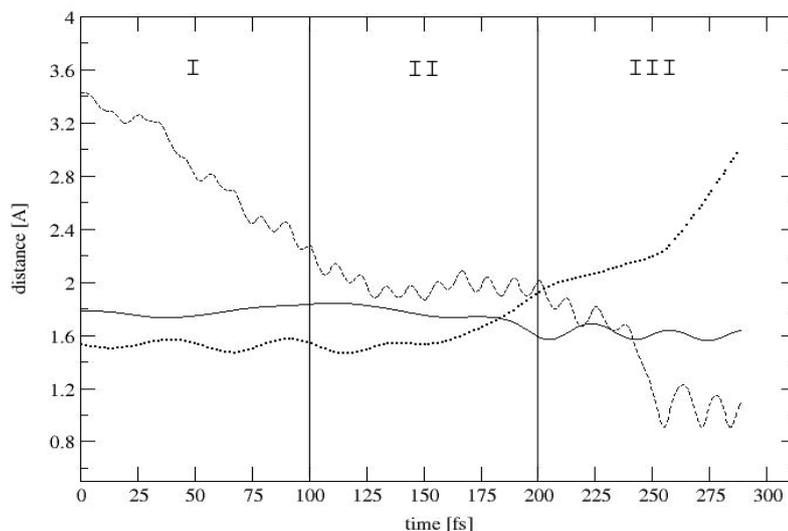
Figure 5.3 Dynamic evolution of a few selected distances along the molecular dynamics of ABLM in vacuum: i) full line: Fe – O1; ii) dashed line: O2 – H_{N2}; iii) dotted line: O1-O2.

immediately (within 50 fs) the H atom of the secondary amine to form a water molecule. In Figure 5.3 this is illustrated by the oscillation of the O2- H_{N2} bond around a value of 1.00 Å starting from 250 fs. The third part of the MD run illustrates how the formation of the water molecule from the OH radical occurs soon after the O1-O2 bond scission.

A second CPMD simulation was performed starting from slightly different initial conditions, leading to the same mechanism for the peroxide cleavage. The only difference observed during the second run was that the system took more time to find a proper configuration to form a hydrogen bond with a correspondingly longer first part of dynamics, ~ 170 fs. However, the same H-bond was eventually formed with H_{N2}, and once this internal H-bond was present, the next steps in the reaction mechanism II and III proceeded at about the same speed as in the first run in Figure 5.3.

To check the energetics of the observed O1-O2 bond cleavage singlet point calculations were performed on a few snapshots along the trajectory with O1-O2 distances of 1.47 Å, 1.53 Å, 1.64 Å, 1.88 Å and 2.05 Å. The potential energy along this path increases by ~12 kcal/mol before going down again with the formation

of the water molecule. This shows that the thermal energy is enough to overcome an energy barrier of this size. The observed energy maximum corresponds to the snapshot in the trajectory where the H-bond between O2 and H_{N2} is formed (dO1-O2 = 1.64 Å).

To identify the products of the O-O bond scission we took a snapshot at ~ 225 fs where the O1-O2 distance is 2.15 Å and using the ADF program we performed a single point calculation. The charge on the formed OH species was slightly negative with a value of -0.266 electronic equivalent according to the Mulliken analysis and the spin distribution analysis indicated the localization of -0.5 spin on the oxygen atom of OH species. Therefore we concluded that the OH species can be identified as an OH radical rather than an OH anion.

For comparison we calculate the energy profile of O1-O2 bond scission in vacuum at 0 K without creation of any H-bond using the geometry obtained for our ABLM model from the Gaussian calculation. The O1-O2 bond was elongated with steps of 0.05 Å from 1.50 Å till 2.4 Å. The structures were fully optimized with Gaussian03 as described in section Methods. We have noticed that the homolytic O1-O2 bond scission is an endothermic reaction with activation energy equal to 17 kcal/mol. The results suggest that the reaction is still possible, however the spontaneous creation of a H-bond in the molecular dynamics simulations facilitates the scission.

5.3.3 ABLM in H_2O

We now investigate the reaction mechanism of the ABLM suicide in water. This simulation will enable us to check the influence of the water solvent on stabilizing the O-O bond of ABLM. After the relaxation procedure, we have followed the trajectory for about 1 ps. We have illustrated the simulation in Figure 5.4, where the Fe-O1, O1-O2, and O2-H_{W1} distances are plotted. In this simulation we observed also the formation of a H-bond after about 500 fs, but in this case the H-bond was formed between the OOH ligand and one of the water molecules indicated as W1 in Figure 5.5. The following observed reaction steps were similar to that described for the simulation in vacuum. Simultaneously with the

formation of the H-bond, the O1-O2 distance starts increasing at ~500 fs. This is illustrated in Figure 5.4 with the dotted line. The released OH species remains H-bonded to W1 and finally after about 400 fs it forms a water molecule by abstraction of a hydrogen atom from the W1 water molecule at ~850 fs (Fig. 5.4, dashed line). Next we could observe a fast chain reaction where the OH species diffuses through the H-bond network of the water molecules. Figure 5.5 shows a few snapshots from the dynamics, where this process is well illustrated. In the first panel we show the initial configuration at $t = 0$ fs; at $t = 605$ fs we can see the formation of the H-bond with H_{W1}; at $t = 902$ fs we observe the breaking of the O1-O2 bond and simultaneous H abstraction and formation of a new water molecule; the last snapshot at $t = 950$ fs illustrates the fast propagation of the OH species through the water molecules facilitated by the existing hydrogen bond network. To identify the OH species we performed
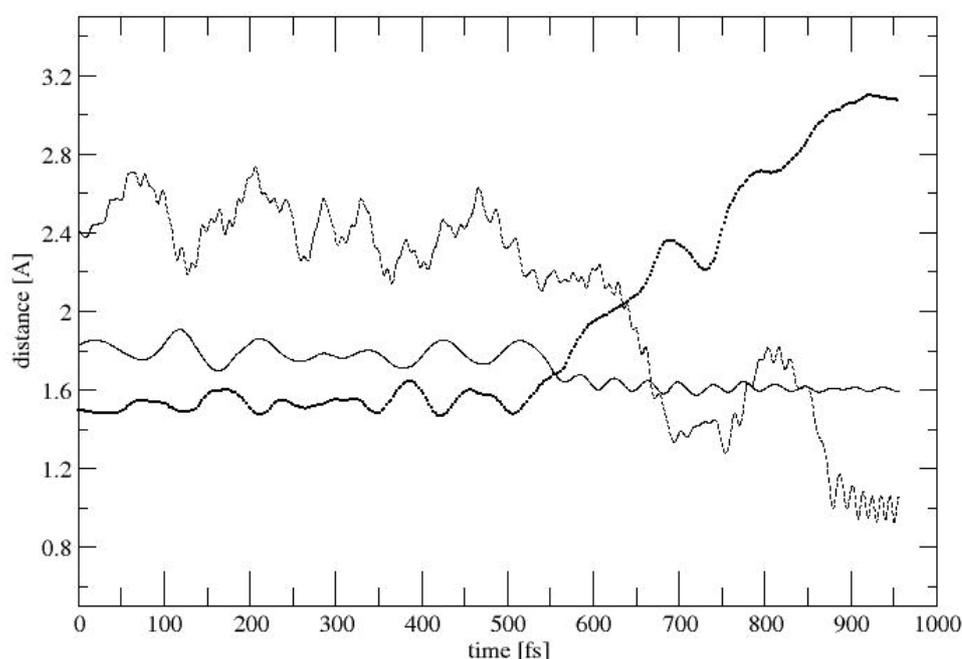

Figure 5.4 Dynamic evolution of a few selected distances along the molecular dynamics of ABLM in water: i) full line: Fe – O1; ii) dashed line: O2 – H_{W1}; iii) dotted line: O1 – O2. H_{W1} is a hydrogen atom coming from one of the water molecules included in the QM part, indicated also in Figure 5.5 (W1).

an ADF single point calculation for the configuration at 650 fs. As in the case of the vacuum simulations, the OH species was identified as a hydroxyl radical with the spin density localized on the oxygen atom and with a small negative charge. For a complete picture of the observed reaction mechanism and identification of the products, we have also performed a single point calculation for the same configuration of the system at t = 650 fs but with a total spin of 3/2. The total bonding energy was about 10 kcal/mol higher than for the system with total spin equal to 1/2. This result shows that no spin flip takes place and our implicit assumption about the conservation of the total spin during the reaction was correct. In the equation below we have also explicitly indicated the total spin of the particles involved to illustrate that the number of unpaired spins can change during the reaction although the total spin is conserved:

$$[Fe^{III}(BLM)\text{-}OOH]^+ \rightarrow [Fe^{IV}(BLM)\text{=}O]^+ + OH^\bullet$$
$$\uparrow \qquad\qquad \uparrow\uparrow \qquad\qquad \downarrow$$
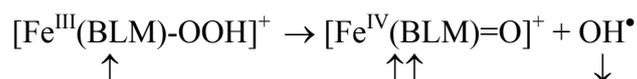
In both single point calculations with different total spin we observed the spin density localized mostly on the iron atom, the oxygen atom of the complex and on the oxygen atom of the OH species. The only difference is that the spin density located on the OH radical is β spin for S = 1/2 while it is α spin for S = 3/2. We do not see any significant spin density on the equatorial ligands of the iron center thus ruling out the presence of cytochrome P450 compound I-like $[Fe^V(BLM)\text{=}O]^+$ complex. Importantly, this high-valent species, like cytochrome P450 compound I, is better described as an Fe(IV)=O unit magnetically coupled to a ligand radical with the ligand hole essentially localized on the deprotonated amide. A detailed study of the electronic structure of this Fe(V) compound was performed by static DFT calculations in Ref. [131].

a)
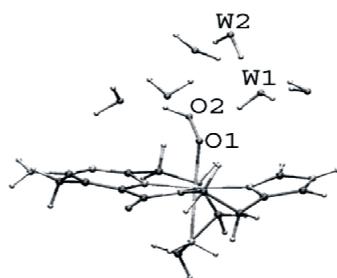
t = 0 fs
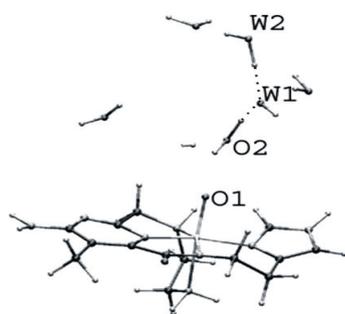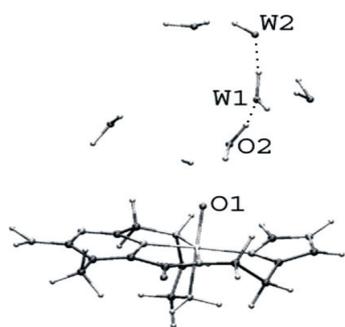
b)
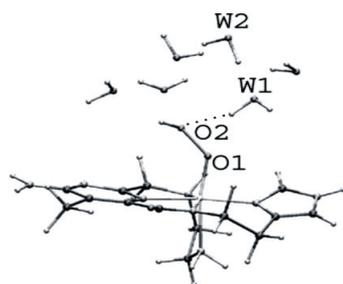
t = 605 fs

c)

t = 902 fs

d)

t = 950 fs

Figure 5.5 A few snapshots along the trajectory of ABLM in water illustrating the relevant steps in the reaction path. Only the quantum mechanics part of the system is shown for clarity.

a) The initial configuration of the system. O1 and O2 indicate the two oxygen atom of hydroperoxo ligand; W1 and W2 are the two water molecules which will participate in the diffusion of the hydroxyl radical through the hydrogen bond network.

b) The O2 atom from the OOH group creates a hydrogen bond with the H_{W1} hydrogen atom.

c) The O1 – O2 bond is already cleaved and the H_{W1} is abstracted. The hydroxyl radical (now W1) form two hydrogen bonds with the new formed water molecule (O2) and the solvent water W2.

d) The hydroxyl radical moves through the hydrogen bond network of the water solvent (now marked as W2).

5.4. Conclusions and discussion

In this work we presented a theoretical study of the mechanism of the first reaction step in the bleomycin suicide based on Car-Parrinello molecular dynamics simulations. We studied an extensive and realistic model of ABLM both in vacuum and in water solvent. We point out that a proper description of the ligand plays a key role in studying the chemistry of ABLM. Our energy profile for the homolytic O-O bond cleavage calculated in vacuum gives a reaction energy of 17 kcal/mol which is about 13 kcal/mol lower than reported previously [5]. The only difference between these two calculations is in the ligand model. As already mentioned in section 3.1, the previous model was simplified by introducing two ammonia molecules into the coordination sphere of the iron center.

During our simulations in vacuum and in water we observed a hydrogen bond formation prior to O-O bond scission. In vacuum an intramolecular H-bond is formed between the OOH group and the hydrogen atom belonging to the secondary amine of the β-aminoalanine (H_{N2}), while in the presence of water it is more likely to observe an intermolecular H-bond with a water molecule. However, in both cases the formation of the H-bond appears to be a necessary step to facilitate the O-O bond scission by lowering the reaction energy barrier. We have identified the OH species formed in that process as a short-lived hydroxyl radical. This radical immediately abstracts the hydrogen atom participating in the H-bond. The second product of the reaction was identified as the BLM-Fe(IV)=O complex ($S = 1$). Thus we can conclude that in the ABLM suicide process the O-O bond cleavage is homolytic. Furthermore, since the process of the OH radical formation and the H abstraction proceed in a concerted fashion, it is more appropriate to classify the whole mechanism as a direct hydrogen abstraction.

The degradation process of ABLM in the absence of the DNA is rather quick according to our simulations, being observed spontaneously in the ps time-scale at 300 K. Experimental data give a half-life time of ABLM between 2 min (6°C, pH = 7.0) [121] or 1.5 min (0°C, pH = 7.2) and < 15 s at pH = 5.8 [143]. The

experimental results clearly show that the half-life time of ABLM highly depends on the temperature and pH and possibly also other conditions. Even taking into account the somehow higher temperature of our simulations, our reaction rate is definitely much higher than expected. This is likely due to an underestimation of the energy barrier within the approximated exchange-correlation functional used here. It has been suggested that self-interaction corrections may improve the accuracy of chemical reactions involving radical species [144, 145] . Nevertheless, we think that the mechanism of self-deactivation observed here is realistic although accelerated.

In this first reaction step we observed the formation of two active species that can further react with a substrate, indicating that the whole process of the Bleomycin suicide is very complex. In the context of the observed homolytic cleavage of the O-O bond, one may easily explain the modification of the bithiazole moiety of the bleomycin [129]. Some of the drug side effects may also be associated with the formation of the active species during the degradation process.

A direct hydrogen abstraction was also suggested by static calculations [131] as the most likely mechanism for the radical formation at the C4' position of the DNA ribose sugar. In fact the reaction with participation of free hydroxyl radical is very unselective and this is why the direct hydrogen abstraction is much easier to accept. Our study shows that the border between homolytic cleavage and direct hydrogen atom abstraction is very thin. Moreover, we cannot exclude that the Fe(IV)=O complex has a direct role in the DNA attack and is not only an intermediate for regenerating the Fe(III)BLM complex. The BLM-Fe(IV)=O complex seems to be more stable than BLM-Fe(III)–OOH and it would have a chance to dock to the DNA molecule. However, at this point we cannot make a straight analogy to the reaction of ABLM with DNA and this issue still requires further investigations.

Chapter six

THE MECHANISM OF THE BLEOMYCIN ACTION: A CAR-PARRINELLO MOLECULAR DYNAMICS INVESTIGATION
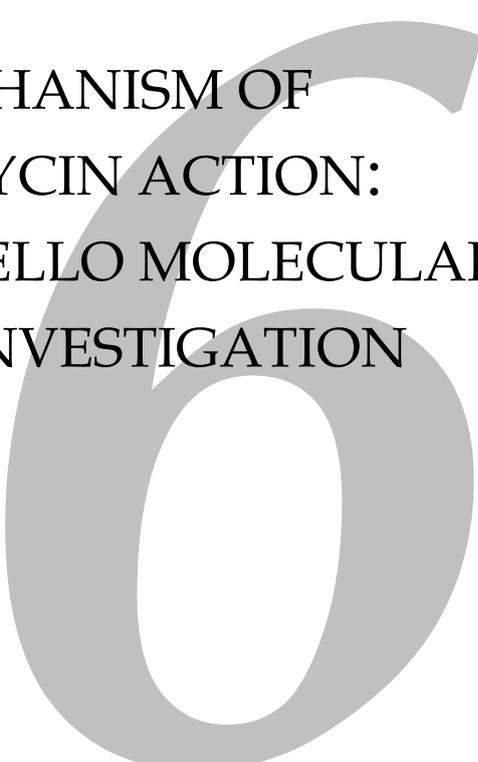
ABSTRACT

Using first-principles molecular dynamics simulations with the Car-Parrinello method, the possible reaction pathways of the active bleomycin-Fe(III)-OOH complex with the deoxyribose sugar of DNA were investigated. The theoretical model of activated bleomycin contains the whole metal bonding domain of the bleomycin ligand. Simulations were performed in a vacuum and in a natural environment by explicitly including a DNA fragment within a QM/MM approach. The simulations show that a facile decaying process involves a homolytic O-O bond cleavage with an almost simultaneous hydrogen atom abstraction. The formation of a hydrogen bond appears to be crucial for the O-O bond cleavage in the Fe(III)-OOH species. The inclusion of a realistic environment is therefore essential to study the selectivity of the mechanism of bleomycin action.

6.1 Introduction

Bleomycin is a drug that has been commonly used in chemotherapy since the early eighties against several types of cancer, including cervix and uterus cancer, head and neck cancer, testicle and penile cancer, and certain types of lymphoma [1]. The precise mechanism of action of bleomycin is not known. The drug is administrated to patients in its metal free form. *In vitro* studies suggest that it forms an Fe(II)BLM complex with an iron atom in the body and then binds an oxygen molecule to generate the O_2-Fe(II)-BLM complex. This complex is converted into activated bleomycin (BLM-Fe(III)-OOH) [120, 121, 124, 125] and is then bound selectively to DNA at the position 5′-GC-3′ or 5′-GT-3′. Bleomycin kills cells by binding with and then degrading genetic material. It carries out this process in the dividing phase of the cell's life cycle and so prevents the cell from growing. It is known from experiment that the degradation of DNA starts by forming a radical at 4′C position of a deoxyribose sugar [17]. Why the reaction occurs selectively at this position is however an open issue, especially given that the mentioned deoxyribose sugar has three hydrogens available for abstraction, *i.e.* 1′H, 2′H and 4′H. Moreover, the first step of the reaction cascade which involves the activation of the O-O bond is not known. Three scenarios are possible: (i) heterolytic cleavage of the O-O bond that would lead to a high-valent intermediate, formally BLM-Fe(V)=O that attacks DNA; (ii) homolytic cleavage of the O-O bond that would lead to a hydroxyl radical and a high-valent intermediate, formally BLM-Fe(IV)=O, both of which could attack DNA; (iii) direct H abstraction of the low-spin BLM-Fe(III)-OOH species with DNA to give a DNA radical, H_2O and a BLM-Fe(IV)=O product. The formally Fe(V)=O species, formed by heterolytic cleavage of the O-O bond, has been observed in heme systems, where it is stabilized by the oxidation of the porphyrin ligand (POR), leading to a $(POR^{+\bullet})Fe(IV)=O$ intermediate known as compound I. By analogy to the heme systems, the heterolysis of the O-O bond in ABLM has been accepted as the most likely first step in DNA degradation [17]. However, in the case of a non-

heme environment, the ligands are much more difficult to oxidize, and therefore not providing sufficient stabilization for the high-valent Fe(V)=O species. Indeed, theoretical calculations show that BLM-Fe(V)=O complex would be energetically unfavorable, given that the electron hole is localized on the deprotonated amide [126]. Thus the homolytic cleavage of the O-O bond started to be considered as a possible mechanism for the O-O bond activation, especially given that O-O homolysis has been shown to occur for low-spin analogues of ABLM [146-148]. On the other hand, this mechanism appears to be unlikely for bleomycin action since it is non-selective. The reactive hydroxyl radical could damage any part of DNA or ABLM itself. It is known that the DNA is not only the target but also protects ABLM from permanent self-destruction, as discussed in Chapter 5. The direct H-atom abstraction from C-H bond of the DNA sugar is closely related to the O-O homolysis, having identical overall reaction energy but a different reaction pathway. It remains to be elucidated whether and how DNA affects ABLM reactivity, what the intermediate reaction products are and why the selectivity occurs at 4'C position of the deoxyribose sugar. These issues are addressed in the present study.

In this chapter CPMD simulations are presented giving insight into the reaction of ABLM with deoxyribose sugar, which is attacked in the first step of the DNA degradation. Since the reaction involves bonds breaking and forming, the obvious choice of the computational method is a quantum mechanical one such as DFT. CPMD is one of the best methods to investigate the reaction pathways without *a priori* selection of the reaction coordinates. Another advantage of CPMD is that the thermal fluctuations are taken into account by performing the simulation at room temperature. Therefore application of the CPMD method to this study is especially important for solving the issue of the selectivity of the ABLM action. Since the role of DNA is questioned, the model should also include the DNA fragment explicitly. This is done here within a QM/MM framework.

6.2 Models and computational details
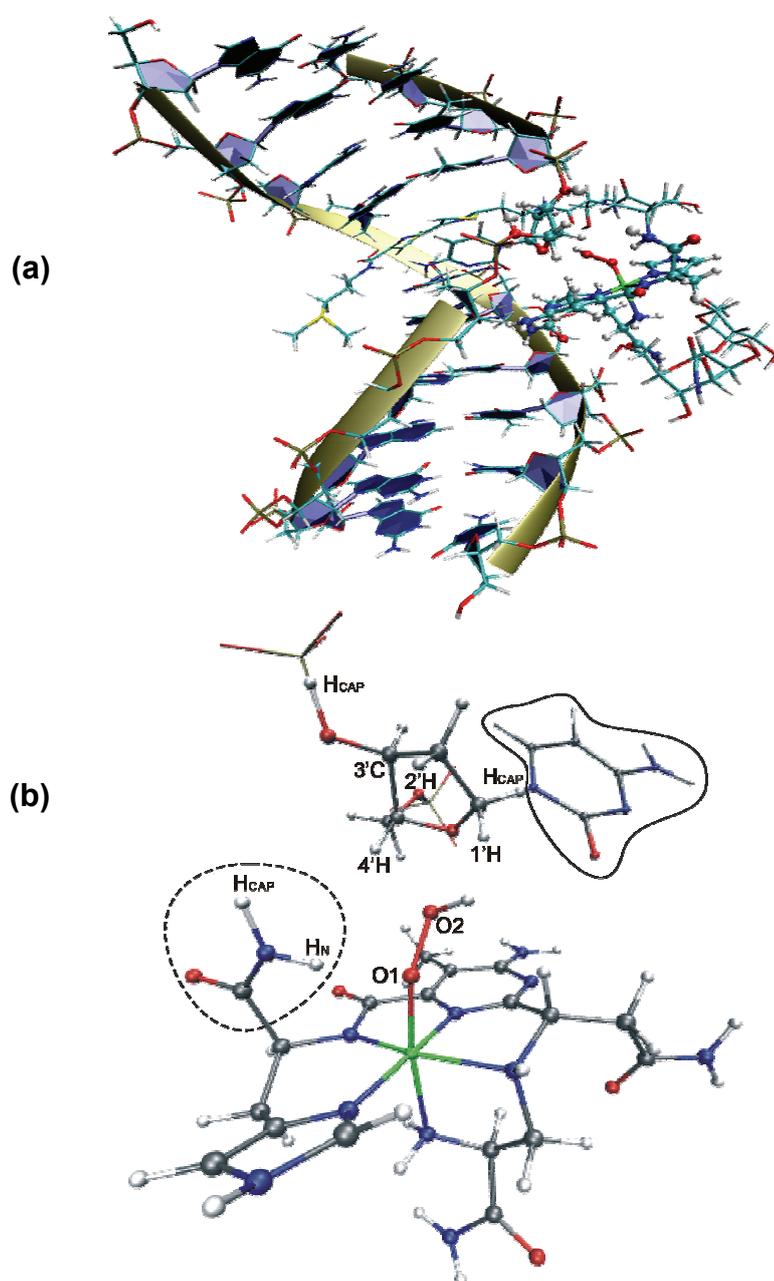
Figure 6.1 (a) The QM/MM model used in this study. The QM part is presented in Balls & Sticks representation. The 3986 water molecules and 16 Na⁺ counter ions included in the simulation are omitted here for clarity. **(b)** The model of ABLM with deoxycytidine in the orientation taken from Ref. [4]. Two models of the deoxyribose sugar are considered, one without and one including the cytosine attached to the sugar at the 1'C position (encircled in the solid line). The dashed line shows the fragment of the bithiazole tail added to the QM part of QM/MM calculations. Atoms represented by lines are present in the MM part of QM/MM calculations. The saturating hydrogens are denoted as H_{CAP}.

The models used in the present study are based on the NMR data for a BLM-Co(III)-OOH complex bonded to a fragment of DNA (D(GGAAGCTTCC)) [4]. The NMR structure has been already presented in Figure 1.2. In the bleomycin ligand we can distinguish three main parts: the metal bonding domain, the sugar moiety, and the bithiazole tail. While the sugar moiety and the bithiazole tail are responsible for recognition and docking to DNA, the metal bonding domain plays a main role in the metal coordination and the O-O bond activation. Thus, for studying the reaction mechanisms we have chosen the metal bonding domain (represented by stick and balls in Fig.1.2) to be described by quantum methods without further simplifications. The equatorial ligands are: the secondary amine of the β-aminoalanine (A''aALA – N2) segment, the pyrimidine (PYR – N3), the imidazole ring (HIS-I – N5) and the amide nitrogen (HIS-A – N4) of the β - hydroxyhistidine. The primary amine of the β - aminoalanine residue (A'aALA – N1) and the oxygen atom of the peroxy group are the two axial ligands.

The deoxyribose sugar (indicated in cyan in Fig 1.2) was included in the model to study the reaction of ABLM with the sugar in the position shown in Figure 1.2. The positions of the oxygen atoms bonded originally to phosphor were fixed during the simulations in vacuum. Two models of the deoxyribose sugar are considered, one without and one including the cytosine attached to the sugar at the 1'C position to check its effect on the selectivity of the reaction (see also Fig. 6.1) The positions of the atoms forming the aromatic ring of cytosine were fixed during the simulations in vacuum. In the QM/MM simulations the entire structure of the BLM-Fe(III)-OOH complex bonded to a fragment of DNA (D(GGAAGCTTCC)) was used (Fig. 6.1a). The complex of ABLM and DNA was placed in a box with 3986 water molecules and the total charge of the system was neutralized by 16 Na$^+$ counter ions. The system was first equilibrated under constant pressure at 300 K for 1 ns and then for another nanosecond at 300 K with constant volume with a box of size 49.88 x 44.17 x 59.22 Å^3. During the equilibration procedure the bleomycin was kept fixed.

Analogous to the simulations performed in vacuum there are two set-ups for the QM/MM simulations that differ in the size of the QM part. The first set-up

contains the metal bonding domain of ABLM and deoxyribose sugar while the second contains the metal bonding domain of ABLM and deoxycytidine. In both cases the QM/MM boundary cuts through several covalent bonds and the dangling bonds are saturated by hydrogen atoms. Two cuts take place in the DNA backbone between an oxygen atom and a phosphor atom in such a way that the entire deoxyribose sugar is present in the QM part (see Fig. 6.1). Since cuts of this type have not previously been reported, additional tests were necessary to validate this procedure. For this purpose only a short segment of the DNA helix, of the order of ten pairs, was included in the simulations with the deoxyribose sugar present in the QM part. The dangling bonds were saturated by hydrogen atoms which are not included in the MM Hamiltonian. The simulations were performed for 0.5 ps and the system was stable. The same set-up was used for simulations with saturating hydrogens explicitly added to the topology of the system as dummy atoms. All electrostatic interactions between the dummy atoms and the MM atoms within a given distance is then excluded from the QM/MM Hamiltonian. After 300 steps of simulations with the exclusions of 3 and 5 Å the O_{QM}-P_{MM} bond becomes very long and the two MM oxygen atoms of the phosphate group collapse together. Hence it was not possible to perform simulations with explicit exclusions for these types of cuts. Apparently, the polar backbone of the DNA is very sensitive to the introduction of additional dummy atoms and all calculations within the QM/MM approach were performed with saturating hydrogens in the QM part only.

CPMD was used for all molecular dynamics simulations described by quantum mechanics. The Gromos96 code [38] with the Amber/parm99 force field [39] is used to describe MM atoms within the QM/MM approach [30, 149]. The Kohn-Sham orbitals are expanded in a PW basis set with a kinetic energy cutoff of 80 Ry. We employed Martins-Troullier *ab initio* norm-conserving pseudopotentials [140]. A generalized gradient approximation for the exchange-correlation functional, following the prescription of Becke and Perdew (BP) was used [53, 54]. All simulations are spin polarized with a total spin $S = 1/2$. Room temperature Car-Parrinello molecular dynamics simulations were performed by using a time

step of 0.09 fs and a value of 400 a.u. for the fictitious electronic mass in the Car-Parrinello Lagrangian. For the simulation including the sugar in vacuum an isolated cubic supercell of size 18 x 18 x 18 Å^3 was used.

6.3 Results and discussion

6.3.1 ABLM with deoxyribose sugar in vacuum

This simulation was started at 20 K and then the temperature was slowly increased to 300 K. As mentioned in the Introduction the sugar molecule has three hydrogen atoms available for abstraction (1'H, 2'H, 4'H). During the simulation the formation of H-bonds is observed between the O2 atom of the –OOH group and all three hydrogens at different time intervals (see Fig. 6.2). We notice however that only the formation of a H-bond with 4'H has an influence on the amplitude of the O1-O2 bond length oscillation. Figure 6.2 shows that when the O2-1'H and O2-2'H distances are shorter than O2-4'H, the O1-O2 bond remains stable, oscillating with a small amplitude around an average value of 1.5 Å. After 750 fs the O2-4'H distance becomes the shortest and the O1-O2 bond
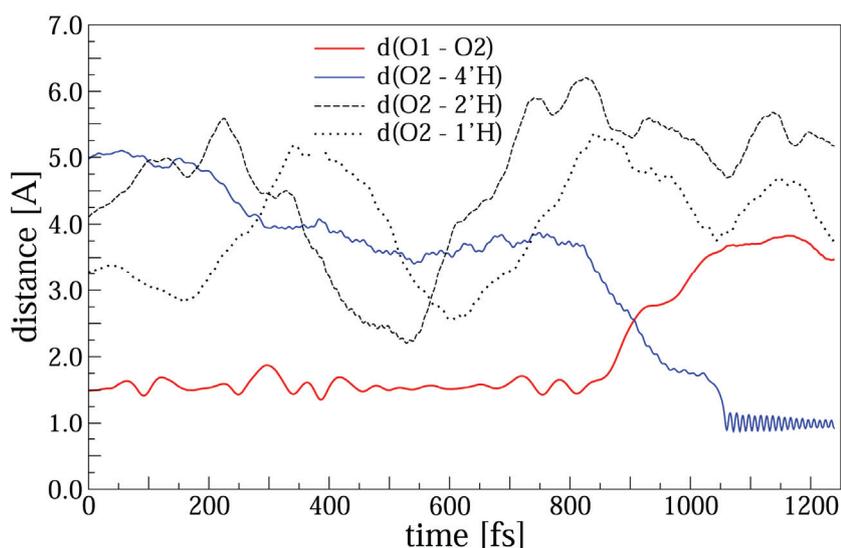

Figure 6.2 Dynamical evolutions of a few distances along the trajectory of the CPMD simulation of ABLM with deoxyribose sugar in vacuum.

distance immediately starts increasing. After a further 300 fs the O1-O2 bond is broken and the water molecule H-O2-4'H is formed by direct hydrogen atom abstraction. The Fe-O distance drops from an average value of 1.8 Å to 1.65 Å indicating the formation of a stronger Fe=O double bond. The selectivity of the observed reaction supports earlier experimental findings. The simulations show clearly the preference of H-bond formation with the 4'H atom which is known to be abstracted in the reaction with DNA.

6.3.2 Molecular orbitals analysis of the BLM-Fe(IV)=O complex

One of the products of the O-O homolysis is the BLM-Fe(IV)=O complex. The Fe(IV)=O component can appear in different spin states. Generally, non-heme enzymes are high spin, $S = 2$, and the heme enzymes are typically low spin, $S = 1$ [150, 151]. The difference is explained by the weaker ligand field in the non-heme systems. It is also well-recognized that the high spin complexes are more reactive than the low spin ones. The total energy of the optimized structure of the BLM-Fe(IV)=O complex in the triplet and quintet spin states indicate that the $S = 1$ complex is more stable by 18 kcal/mol than $S = 2$ configuration. An analysis of the MOs of the BLM-Fe(IV)=O complex was done in terms of perturbation of the basic FeO^{2+} pattern ($S = 2$) by the bleomycin ligand using the ADF code [142]. The results show that the two unpaired electrons occupy the HOMO and HOMO-1 orbitals that are constructed mainly by the $2\pi_x$ and $2\pi_y$ orbitals of FeO^{2+}. In the FeO^{2+} there are two degenerate δ orbitals derived by iron d_{xy} and $d_{x^2-y^2}$ orbitals. In the BLM-Fe(IV)=O complex that has a distorted octahedral coordination, the degeneracy of d_{xy} and $d_{x^2-y^2}$ orbitals is lifted. The $\delta_{x^2-y^2}$ orbital is strongly destabilized by anti-bonding interactions with the ligand and ends up as unoccupied $65a^{\uparrow}$, 2.7 eV above the δ_{xy} orbital ($61a^{\downarrow}$) which is doubly occupied. This yields a stable low spin complex.
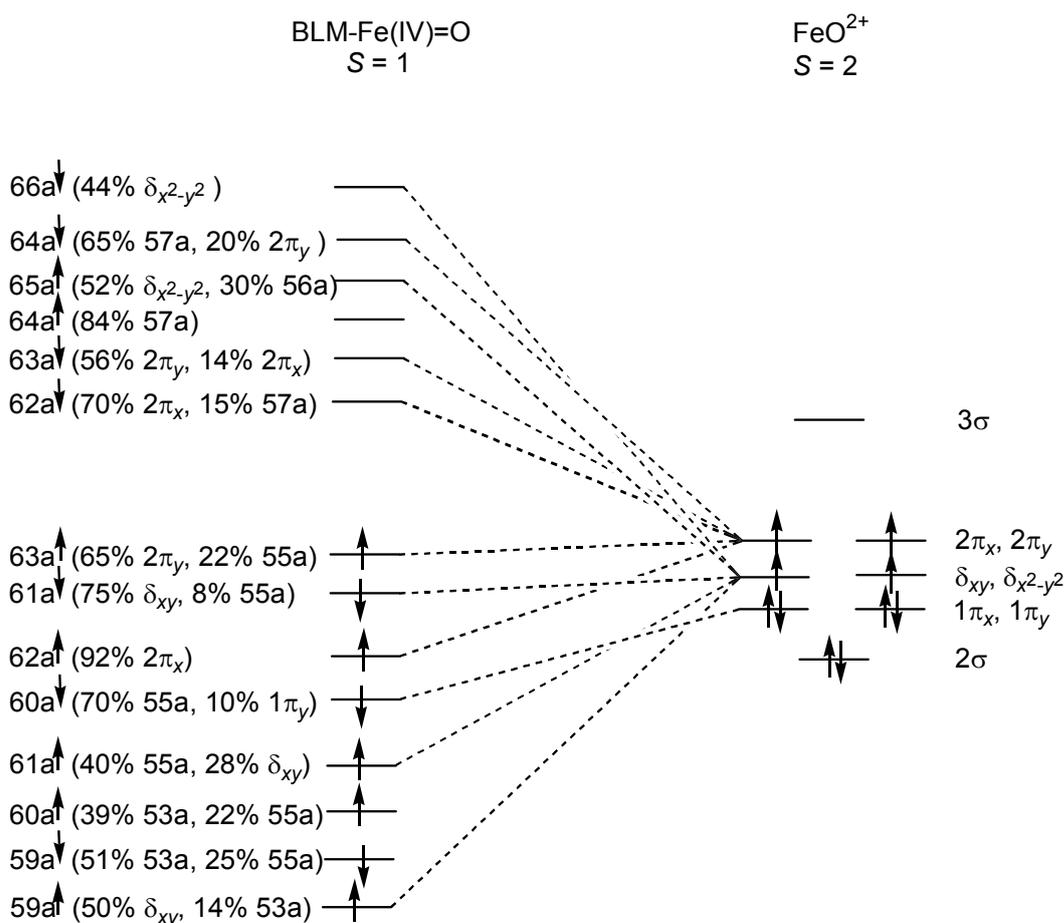
Figure 6.3 Molecular orbital diagram illustrating the perturbation of the FeO^{2+} molecular orbitals (right) to create the BLM-Fe(IV)=O complex (left). The percentage composition given in brackets is coming from the molecular orbitals of the BLM ligand and FeO^{2+} fragment.

6.3.3 Selectivity of the hydrogen atom abstraction

The stability of the three possible radicals of the deoxyribose sugar was checked to validate the observed selectivity of the hydrogen abstraction. A full energy minimization was therefore performed for each possible radical. It was observed that the most stable radical is indeed the 4'C radical, which is lower in energy by 2.9 kcal/mol and 8.4 kcal/mol than the 1'C and 2'C radicals, respectively. This result is in line with the observed selectivity during the dynamical simulations. On the other hand, in the natural system the 1'C atom is connected to the cytosine. The calculated energies for the radicals 1'C and 4'C for deoxycytidine indicate that the presence of the aromatic ring in that position stabilizes the 1'C radical.

However, the difference in energy is quite small, the 1'C radical being only 0.9 kcal/mol lower than the 4'C. In this situation other factors, such as steric and electrostatic interactions or the geometry of the formed H-bond, could influence the selectivity of the hydrogen abstraction to form a radical.

Since the issue of the reaction selectivity is important, CPMD simulations with an extended model including the cytosine molecule were performed. During the 2 ps of the molecular dynamics no spontaneous reaction was observed (see Fig. 6.4). The O-O bond did not break, but fluctuated largely around an average bond length of 1.47 Å. Formation of various H-bonds is also observed but none is able to facilitate the O-O bond scission. A few attempts are however observed, especially at 0.82 ps, 1.37 ps and 1.80 ps while the O-O distance increases to 1.82 Å, 1.86 Å and 1.93 Å, respectively. Finally the ABLM molecule drifts away from the sugar and this disables any reaction. As a consequence of that movement the peroxy group starts rotating around the Fe-O1 axis. This simulation clearly shows the need of using a QM/MM approach for studying the reaction mechanism between ABLM and deoxyribose sugar. The presence of DNA
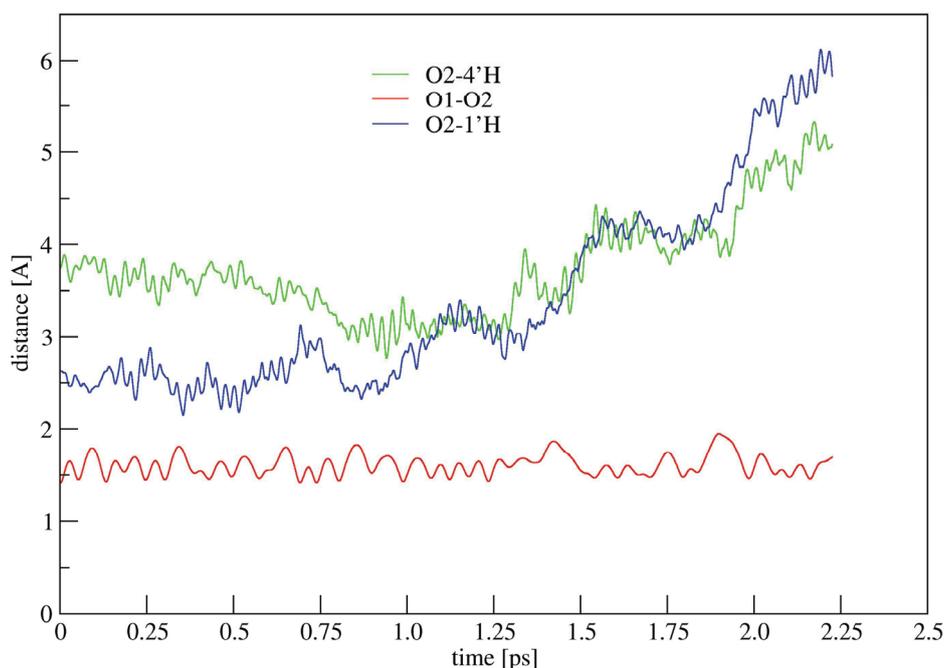

Figure 6.4 Dynamical evolutions of a few distances along the CPMD simulation of ABLM with deoxycytidine in a vacuum.

prevents the molecules from separation. In a natural setting DNA and ABLM create a well-defined reaction environment where not only the steric interactions are taken into account but also all the electrostatic interactions that could influence the selectivity.

6.3.4 ABLM with DNA: QM/MM study

To gain insight into the selectivity of the reaction of ABLM with DNA, an extensive model was set up within the QM/MM approach. The QM part was increased by adding atoms belonging to the bithiazole tail in order to realistically represent the active site and explore all possible scenarios of the H atom abstraction. In the active site there are 4 hydrogens available for abstraction: 1'H, 2'H, 4'H of deoxyribose sugar and H_N of the bithiazole tail. The dynamics has been followed for ~2.5 ps and relevant variables are plotted in Figure 6.5. As illustrated in the graph 6.5 c the temperature was increased during the run from 300 K up to 400 K in order to accelerate the reaction. A spontaneous O-O bond scission is observed during the simulation after increasing the temperature. Soon after the O-O bond breaking a water molecule is formed and the deoxyribose sugar is destabilized. The scission is facilitated by the formation of a H-bond between O2 and H_N with an average distance $dO2\text{-}H_N = 1.8$ Å and an angle $\angle$N-O2-H_N oscillating between 140 and 180 degree. These values provide evidence for a strong H-bond. It can be noted that formation of the O2$\cdots H_N$ H-bond affects the amplitude of the O-O bond oscillation already at 300 K as illustrated in Figure 6.5 at 0.5 ps and 1.1 ps (red line). Finally at 1.8 ps the O-O bond is broken and the OH species is formed. The newly formed fragment stays hydrogen bonded to the H_N atom of the bithiazole tail and the 4'H of the sugar molecule. The strong H-bonds are represented in Figure 6.5 by short distances 4'H-O2 and O2-H_N (in green and brown, respectively) and H-bond angles oscillating around 160 degree. A visualization of the total spin distribution for snapshots at 1.0 ps and 1.85 ps is presented in Figure 6.6. The first snapshot is representative of the system during the first part of the simulations from 0 up to 1.7 ps. As shown in figure 6.6a the unpaired α electron is situated on the iron-peroxy center of the ABLM.
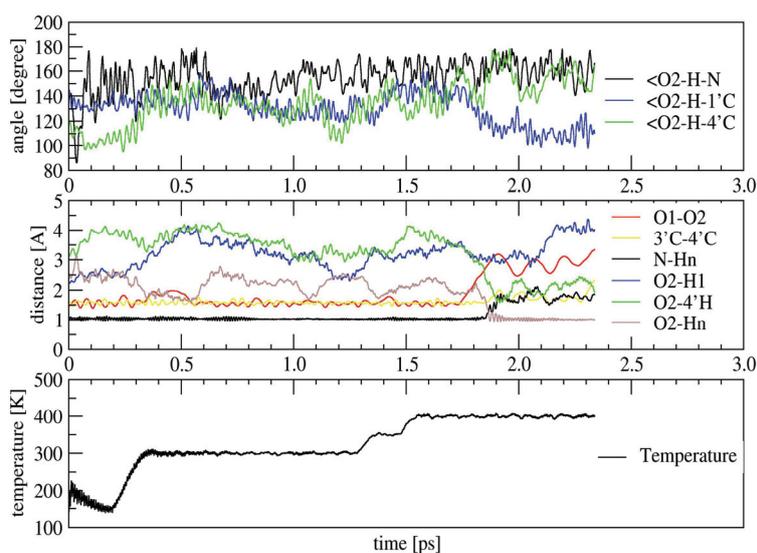
Figure 6.5 Dynamical evolution of relevant parameters along the trajectory of the QM/MM simulation of the ABLM with deoxyribose sugar.

The second snapshot is taken when the O-O distance is equal to 2.73 Å and the O2 atom is hydrogen bonded to H_N with a distance of 1.64 Å and 4'H with a distance of 2.35 Å. The two unpaired α electrons are localized on the Fe(IV)=O center while the unpaired β electron is localized mostly on the OH species. This clearly indicates that a hydroxyl radical is created as a consequence of the O1-O2 bond scission. In addition, the spin distribution in Figure 6.6b shows that a minor part of the spin density is on the sugar. It appears that electron density is transferred from the sugar to the hydroxyl radical through the strong H-bond. This effect further destabilizes the deoxyribose sugar and leads to the 3'C-4'C bond cleavage that is observed starting from 1.9 ps (see yellow curve in Fig. 6.5). Additional evidence for charge transfer from the sugar ring is the electrostatic potential map that shows the deoxyribose sugar being more electropositive after the formation of the water molecule (see Fig. 6.7). The water molecule formed after the H_N atom abstraction is bridging between the sugar molecule and a BLM-Fe(IV)=O complex. The results presented here provide a different perspective on the commonly accepted reaction mechanism of the ABLM action. Experimental work *in vitro* indicates that the degradation of the DNA starts with the 3'C-4'C bond
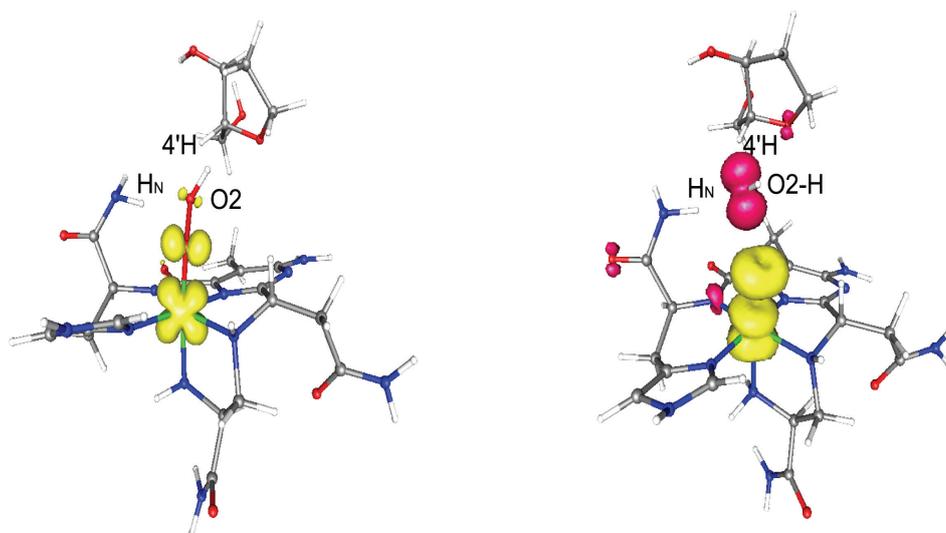
Figure 6.6 The total spin density distribution of the QM part snapshots taken at 1.0 ps **(a)** and 1.85 ps **(b)** of the QM/MM simulation. The α spin is indicated in yellow and β spin in pink.

scission. The same C-C bond is indeed broken during the present simulation at 400 K. However, no 4′H atom abstraction is observed during the simulation. It remains to be seen if the temperature increase from 300 K to 400 K has not only accelerated the process but has also affected the reaction mechanism. However, it should be pointed out that the Fe(IV)=O species is known to be able to abstract the H atom [152]. At the end of the simulation a water molecule is bridging between the sugar molecule and a BLM-Fe(IV)=O complex. This would be a very good starting point for a cascade reaction of H-atom abstraction involving the water molecule and the 4′H atom of the deoxyribose sugar. It is also likely that the bithiazole tail plays a role in quenching the radical formed by the homolysis of O-O bond and somehow facilitates the destruction of the second strand of the DNA helices thus explaining the cleavage of the double stranded DNA by only one ABLM complex as it is observed experimentally. The results presented here thereby pave the way to new avenues of investigation aimed at solving this complex puzzle.
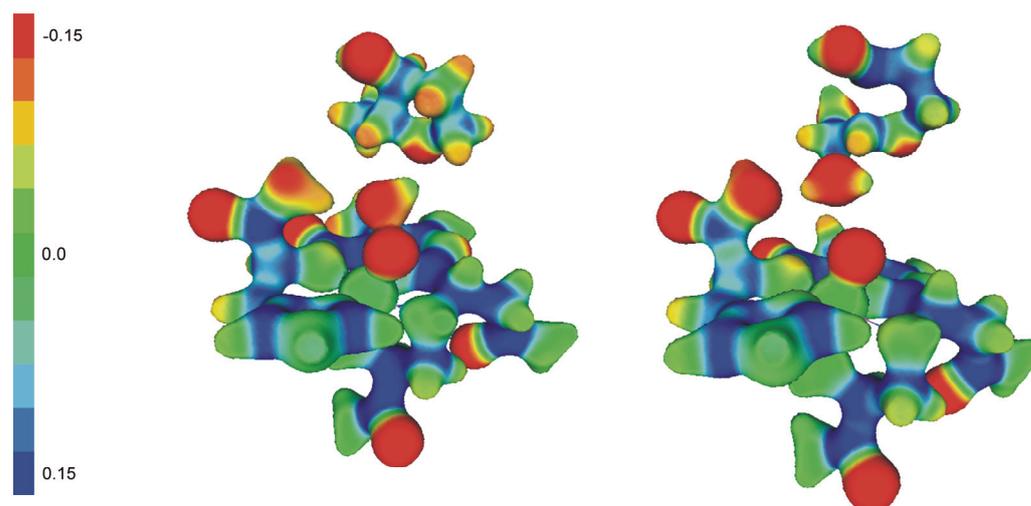
Figure 6.7 The electrostatic potential map at 1.0 ps and 2.3 ps

6.4 Conclusions and future outlook

First principles molecular dynamics simulations strongly support a homolytic cleavage as the mechanism of the O-O bond activation in ABLM complex. The reaction with sugar can be characterized as a direct hydrogen atom abstraction. The newly formed hydroxyl radical is hydrogen bonded with the 4'H of the sugar which is then abstracted to form a sugar radical and a water molecule. Thus we can conclude that the selectivity of the observed H abstraction derives from the fact that a hydrogen bond is formed prior to the bond scission and that the hydroxyl radical is somehow trapped by a specific hydrogen bond. However, the simplified model of the ABLM with deoxyribose sugar is not enough to draw a definitive conclusion about the selective reaction between ABLM and DNA. In addition, inclusion of cytosine in order to simulate a more realistic system in vacuum was not conclusive in this respect. The simulations do however underline the importance of the creation of the H-bond and emphasize the need to perform the simulations in a more realistic environment by including the double stranded DNA fragment. This type of simulation was done in the framework of the QM/MM approach. The simulations give evidence that the selectivity of the reaction is a consequence of the formed H-bond. The newly formed OH species is

H-bonded to H_N and $4'H$. The analysis of the electron spin density reveals that the observed OH species is a hydroxyl radical. This radical is able to destabilize the $3'C$-$4'C$ bond of the deoxyribose sugar causing its cleavage and to abstract the hydrogen from the bithiazole tail of BLM to form a water molecule.

The FPMD method is clearly a powerful approach to study the reaction mechanism. The presented investigations are able to answer some of the initial questions and, as it is often the case, they also raise new issues that deserve further detailed analysis. One appealing question concerns the free energy profile of the reaction where the hydrogen atom is abstracted by ABLM. Recently metadynamics method has been developed which allows calculating the free energy [153, 154]. This method allows exploring the free energy surface in the space defined by a few collective variables. The collective variables important for the problem under investigation would be the parameters of the created H-bonds. The FPMD simulations suggest that the H-atom abstraction mechanism is not more selective than the homolytic O-O bond cleavage, in contrast with references from previous studies [5, 126, 131]. The key to understand the cleavage of the double stranded DNA by only one ABLM complex may be the creation of the active BLM-Fe(IV)=O complex. This active species is formed during all presented simulations. The QM/MM simulation shows additionally that the created water molecule may be used as a bridge to facilitate the cascade reaction of H-atom abstraction. Unfortunately our simulations are too short to give a strong evidence for further steps of the reaction and longer QM/MM simulations with extended QM part of the cytosine have to be performed at 300K to draw stronger conclusions about the mechanism of ABLM action. Further studies are in progress to provide complementary knowledge for a complete picture of the ABLM action.

REFERENCES

1. R. M. Burger, *Nature of activated bleomycin.* Metal-Oxo and Metal-Peroxo Species in Catalytic Oxidations, **97,** 287 (2000)

2. T. E. Lehmann, L. J. Ming, M. E. Rosen and L. Que, *NMR studies of the paramagnetic complex Fe(II)-bleomycin.* Biochemistry, **36,** 2807 (1997)

3. S. J. Brown, P. K. Mascharak and D. W. Stephan, *Characterization of a crystalline synthetic analog of copper(II) bleomycin.* Journal of the American Chemical Society, **110,** 1996 (1988)

4. C. Q. Zhao, C. W. Xia, Q. K. Mao, H. Forsterling, E. DeRose, W. E. Antholine, W. K. Subczynski and D. H. Petering, *Structures of HO2-Co(III)bleomycin A(2) bound to d(GAGCTC)(2) and d(GGAAGCTTCC)(2): structure-reactivity relationships of Co and Fe bleomycins.* Journal of Inorganic Biochemistry, **91,** 259 (2002)

5. N. Lehnert, F. Neese, R. Y. N. Ho, L. Que and E. I. Solomon, *Electronic structure and reactivity of low-spin Fe(III)-hydroperoxo complexes: Comparison to activated bleomycin.* Journal of the American Chemical Society, **124,** 10810 (2002)

6. L. Balbes, S. Mascarella and D. B. Boyd, *A perspective of modern methods in computer-aided drug design.* Reviews in Computational Chemistry, **5,** 337–370 (1994)

7. T. Lengauer, *Algorithmic research problems in molecular bioinformatics.* Proceedings of the 2nd Israeli Symposium on the Theory of Computing and Systems, (1993)

8. D. B. Boyd, *Successes of computer-assisted molecular design.* Reviews in Computational Chemistry, **1,** 355–371 (1990)

9. P. W. Finn and L. E. Kavraki, *Computational Approaches to drug design.* Algorithmica, **25,** 347–371 (1999)

10. L. M. Brown, A. M. Malkinson, D. E. Rannels and S. R. Rannels, *Compensatory lung growth after partial pneumonectomy enhances lung tumorigenesis induced by 3-methylcholanthrene.* Cancer Research, **59,** 5089 (1999)

11. I. Haq, *Thermodynamics of drug-DNA interactions.* Archives of Biochemistry and Biophysics, **403,** 1 (2002)

12. S. Neidle and Z. Abraham, *Structural and sequence-dependent aspects of drug intercalation into nucleic-acids.* Crc Critical Reviews In Biochemistry, **17,** 73 (1984)

13. L. S. Lerman, *Structural considerations in interaction of DNA and acridines.* Journal of Molecular Biology, **3,** 18 (1961)

14. L. F. Povirk, Y. H. Han and R. J. Steighner, *Structure of bleomycin-induced DNA double-strand breaks - Predominance of blunt ends and single-base 5' extensions.* Biochemistry, **28,** 5808 (1989)

15. R. J. Steighner and L. F. Povirk, *Bleomycin-induced DNA lesions at mutational hot-spots - Implications for the mechanism of double-strand cleavage.* Proceedings of the National Academy of Sciences of the United States of America, **87,** 8350 (1990)

16. J. Stubbe, J. W. Kozarich, W. Wu and D. E. Vanderwall, *Bleomycins: A structural model for specificity, binding, and double strand cleavage.* Accounts of Chemical Research, **29,** 322 (1996)

17. R. M. Burger, *Cleavage of nucleic acids by bleomycin.* Chemical Reviews, **98,** 1153 (1998)

18. E. A. Sausville, R. W. Stein, J. Peisach and S. B. Horwitz, *Properties and products of degradation of DNA by Bleomycin and iron(II).* Biochemistry, **17,** 2746 (1978)

19. G. H. McGall, L. E. Rabow, G. W. Ashley, S. H. Wu, J. W. Kozarich and J. Stubbe, *New insight into the mechanism of base propenal formation during bleomycin-mediated DNA-degradation.* Journal of the American Chemical Society, **114,** 4958 (1992)

20. L. M. Fisher, R. Kuroda and T. T. Sakai, *Interaction of Bleomycin A2 with deoxyribonucleic-acid - DNA unwinding and inhibition of bleomycin-induced DNA breakage by cationic thiazole amides related to bleomycin A2.* Biochemistry, **24,** 3199 (1985)

21. M. R. Ciriolo, R. S. Magliozzo and J. Peisach, *Microsome-stimulated activation of ferrous bleomycin in the presence of DNA.* Journal of Biological Chemistry, **262,** 6290 (1987)

22. I. Mahmutoglu and H. Kappus, *Redox cycling of Bleomycin-Fe(III) by an Nadh-dependent enzyme, and DNA damage in isolated rat-liver nuclei.* Biochemical Pharmacology, **36,** 3677 (1987)

23. H. Umezawa, T. Takita, Y. Sugiura, M. Otsuka, S. Kobayashi and M. Ohno, *DNA-bleomycin interaction - Nucleotide sequence-specific binding and cleavage of DNA by bleomycin.* Tetrahedron, **40,** 501 (1984)

24. B. Giese, X. Beyrichgraf, P. Erdmann, M. Petretta and U. Schwitter, *The chemistry of single-stranded 4'-DNA radicals - Influence of the radical precursor on anaerobic and aerobic strand cleavage.* Chemistry & Biology, **2,** 367 (1995)

25. B. Giese, P. Erdmann, L. Giraud, T. Gobel, M. Petretta, T. Schafer and M. Vonraumer, *Heterolytic C-O bond cleavage of 4'-nucleotide radicals.* Tetrahedron Letters, **35,** 2683 (1994)

26. A. D. Dandrea and W. A. Haseltine, *Sequence specific cleavage of DNA by anti-tumor antibiotics neocarzinostatin and bleomycin.* Proceedings of the National Academy of Sciences of the United States of America, **75,** 3608 (1978)

27. C. Bailly and M. J. Waring, *The purine 2-amino group as a critical recognition element for specific DNA cleavage by bleomycin and calicheamicin.* Journal of the American Chemical Society, **117,** 7311 (1995)

28. D. L. Boger, T. M. Ramsey, H. Cai, S. T. Hoehn and J. Stubbe, *Definition of the effect and role of the bleomycin A(2) valerate substituents: Preorganization of a rigid, compact conformation implicated in sequence-selective DNA cleavage.* Journal of the American Chemical Society, **120,** 9149 (1998)

29. A. Laio, J. VandeVondele and U. Rothlisberger, *D-RESP: Dynamically generated electrostatic potential derived charges from quantum mechanics/molecular mechanics simulations.* Journal of Physical Chemistry B, **106,** 7300 (2002)

30. A. Laio, J. VandeVondele and U. Rothlisberger, *A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations.* Journal of Chemical Physics, **116,** 6941 (2002)

31. W. C. Swope, H. C. Andersen, P. H. Berens and K. R. Wilson, *A computer-simulation method for the calculation of equilibrium-constants for the formation of physical clusters of molecules - Application to small water clusters.* Journal of Chemical Physics, **76,** 637 (1982)

32. D. Frenkel and B. Smit, *Understanding molecular smulations.* Academic Press, (2002)

33. H. C. Andersen, *Molecular-dynamics simulations at constant pressure and/or temperature.* Journal of Chemical Physics, **72,** 2384 (1980)

34. W. G. Hoover, *Canonical dynamics - equilibrium phase-space distributions.* Physical Review A, **31,** 1695 (1985)

35. S. Nose, *A molecular-dynamics method for simulations in the canonical ensemble.* Molecular Physics, **52,** 255 (1984)

36. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *CHARMM - A program for macromolecular energy, minimization, and dynamics calculations.* Journal of Computational Chemistry, **4,** 187 (1983)

37. W. L. Jorgensen and J. Tiradorives, *The Opls potential functions for proteins - energy minimizations for crystals of cyclic-peptides and crambin.* Journal of the American Chemical Society, **110,** 1657 (1988)

38. W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Huenenberger, P. Krueger, A. E. Mark, W. R. P. Scott and I. G. Tironi, *Biomolecular simulation: The Gromos96 manual and user guide.* Hochschulverlag an der ETH Zurich, (1996)

39. D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *The Amber biomolecular simulation programs.* Journal of Computational Chemistry, **26,** 1668 (2005)

40. J. W. Ponder and D. A. Case, *Force fields for protein simulations.* **66** (2003)

41. J. M. Wang, P. Cieplak and P. A. Kollman, *How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?* Journal of Computational Chemistry, **21,** 1049 (2000)

42. C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *A Well-behaved electrostatic potential based method using charge restraints for deriving atomic charges - The RESP model.* Journal of Physical Chemistry, **97,** 10269 (1993)

43. W. D. Cornell, P. Cieplak, C. I. Bayly and P. A. Kollman, *Application of RESP charges to calculate conformational energies, hydrogen-bond energies, and free-energies of sSolvation.* Journal of the American Chemical Society, **115,** 9620 (1993)

44. D. M. York, T. A. Darden and L. G. Pedersen, *The effect of long-range electrostatic interactions in simulations of macromolecular crystals - A comparison of the Ewald and truncated list methods.* Journal of Chemical Physics, **99,** 8345 (1993)

45. J. W. Eastwood and R. W. Hockney, *Shaping force law in 2-dimensional particle-mesh models.* Journal of Comparative Physiology, **16,** 342 (1974)

46. W. Koch and H. C.Holthausen, *A chemist's guide to DFT.* VCH, (2000)

47. R. Car and M. Parrinello, *Unified approach for molecular-dynamics and density-functional theory.* Physical Review Letters, **55,** 2471 (1985)

48. P. Hohenberg and W. Kohn, *Inhomogeneous electron gas.* Physical Review B, **136,** B864 (1964)

49. W. Kohn and L. J. Sham, *Self-consistent equations including exchange and correlation effects.* Physical Review, **140,** 1133 (1965)

50. R. G. Parr and W. T. Yang, *Density functional theory of atoms and molecules.* Oxford University Press, (1986)

51. D. M. Ceperley and B. J. Alder, *Ground-state of the electron-gas by a stochastic method.* Physical Review Letters, **45,** 566 (1980)

52. S. H. Vosko, L. Wilk and M. Nusair, *Accurate spin-dependent electron liquid correlation energies for local spin-density calculations - A critical analysis.* Canadian Journal of Physics, **58,** 1200 (1980)

53. A. D. Becke, *Density functional calculations of molecular-bond energies.* Journal of Chemical Physics, **84,** 4524 (1986)

54. J. P. Perdew, *Density-functional approximation for the correlation-energy of the inhomogeneous electron-gas.* Physical Review B, **33,** 8822 (1986)

55. C. T. Lee, W. T. Yang and R. G. Parr, *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron-density.* Physical Review B, **37,** 785 (1988)

56. J. P. Perdew, K. Burke and M. Ernzerhof, *Generalized gradient approximation made simple.* Physical Review Letters, **77,** 3865 (1996)

57. J. M. Kriegl, K. Nienhaus, P. C. Deng, J. Fuchs and G. U. Nienhaus, *Ligand dynamics in a protein internal cavity.* Proceedings of the National Academy of Sciences of the United States of America, **100,** 7069 (2003)

58. A. D. Becke, *A new mixing of Hartree-Fock and local density-functional theories.* Journal of Chemical Physics, **98,** 1372 (1993)

59. D. Marx and J. Hutter, *Ab initio molecular dynamics: Theory and implementation.* Modern Methods and Algorithms of Quantum Chemistry, **1,** 301 (2000)

60. D. R. Hamann, M. Schluter and C. Chiang, *Norm-conserving pseudopotentials.* Physical Review Letters, **43,** 1494 (1979)

61. T. Z. Mordasini and W. Thiel, *Combined quantum mechanical and molecular mechanical approaches.* Chimia, **52,** 288 (1998)

62. A. Warshel and M. Levitt, *Theoretical studies of enzymic reactions - Dielectric, electrostatic and steric stabilization of carbonium-ion in reaction of lysozyme.* Journal of Molecular Biology, **103,** 227 (1976)

63. J. Hutter, A. Alavi, T. Deutsch, M. Bernasconi, S. T. Goedecker, D. Marx, M. Tuckerman and M. Parrinello, *CPMD,* Zurich (2000)

64. C. Manach, A. Scalbert, C. Morand, C. Remesy and L. Jimenez, *Polyphenols: Food sources and bioavailability.* American Journal of Clinical Nutrition, **79,** 727 (2004)

65. J. A. Luchsinger and R. Mayeux, *Dietary factors and Alzheimer's disease.* The Lancet Neurology, **3,** 579 (2004)

66. B. J. Morris, *A forkhead in the road to longevity: The molecular basis of lifespan becomes clearer.* Journal of Hypertension, **23,** 1285 (2005)

67. J. A. Vita, *Polyphenols and cardiovascular disease: effects on endothelial and platelet function.* The American Journal of Clinical Nutrition, **81,** 292S (2005)

68. E. Middleton, Jr., C. Kandaswami and T. C. Theoharides, *The effects of plant flavonoids on mammalian cells: implications for inflammation, heart disease, and cancer.* Pharmacological Reviews, **52,** 673 (2000)

69. C. D. Kanakis, P. A. Tarantilis, M. G. Polissiou, S. Diamantoglou and H. A. Tajmir-Riahi, *DNA interaction with naturally occurring antioxidant flavonoids quercetin, kaempferol, and delphinidin.* Journal of Biomolecular Structure & Dynamics, **22,** 719 (2005)

70. A. Rahman, F. Fazal, J. Greensill, K. Ainley, J. H. Parish and S. M. Hadi, *Strand scission in DNA induced by dietary flavonoids: role of Cu(I) and oxygen free radicals and biological consequences of scission.* Molecular and Cellular Biochemistry, **111,** 3 (1992)

71. R. Solimani, *The flavonols quercetin, rutin and morin in DNA solution: UV-vis dichroic (and mid-infrared) analysis explain the possible association between the biopolymer and a nucleophilic vegetable-dye.* Biochimica et Biophysica Acta (BBA) - General Subjects, **1336,** 281 (1997)

72. M. L. Calabro, S. Tommasini, P. Donato, D. Raneri, R. Stancanelli, P. Ficarra, R. Ficarra, C. Costa, S. Catania, C. Rustichelli and G. Gamberini, *Effects of [alpha]- and [beta]-cyclodextrin complexation on the physico-chemical properties and antioxidant activity of some 3-hydroxyflavones.* Journal of Pharmaceutical and Biomedical Analysis, **35,** 365 (2004)

73. Y. Zheng, I. S. Haworth, Z. Zuo, M. S. S. Chow and A. H. L. Chow, *Physicochemical and structural characterization of quercetin-beta-cyclodextrin complexes.* Journal of Pharmaceutical Sciences, **94,** 1079 (2005)

74. A. Di Pietro, G. Conseil, J. M. Perez-Victoria, G. Dayan, H. Baubichon-Cortay, D. Trompier, E. Steinfels, J. M. Jault, H. de Wet and a. Maitrejean et, *Modulation by flavonoids of cell multidrug resistance mediated by P-glycoprotein and related ABC transporters.* Cellular and Molecular Life Sciences: CMLS, **59,** 307 (2002)

75. F. Hollosy and G. Keri, *Plant-derived protein tyrosine kinase inhibitors as anticancer agents.* Current Medicinal Chemistry. Anti-Cancer Agents, **4,** 173 (2004)

76. M. N. Jacobs and D. F. V. Lewis, *Steroid hormone receptors and dietary ligands: a selected review.* The Proceedings of The Nutrition Society, **61,** 105 (2002)

77. R. J. Williams, J. P. E. Spencer and C. Rice-Evans, *Flavonoids: antioxidants or signalling molecules?* Free Radical Biology and Medicine, **36,** 838 (2004)

78. K. Ito, H. Suzuki, T. Horie and Y. Sugiyama, *Apical/basolateral surface expression of drug transporters and its role in vectorial drug transport.* Pharmaceutical Research, **22,** 1559 (2005)

79. M. H. Saier, Jr., *A functional-phylogenetic classification system for transmembrane solute transporters.* Microbiology and Molecular Biology Reviews: MMBR, **64,** 354 (2000)

80. G. L. Sottocasa, G. C. Lunazzi, C. Tiribelli and F. Sidney Fleischer *Isolation of bilitranslocase, the anion transporter from liver plasma membrane for bilirubin and other organic anions.* in Methods in Enzymology, **174,** 50 (1989)

81. L. Battiston, S. Passamonti, A. Macagno and G. L. Sottocasa, *The Bilirubin-Binding Motif of Bilitranslocase and Its Relation to Conserved Motifs in Ancient Biliproteins.* Biochemical and Biophysical Research Communications, **247,** 687 (1998)

82. S. Passamonti, M. Terdoslavich, A. Margon, A. Cocolo, N. Medic, F. Micali, G. Decorti and M. Franko, *Uptake of bilirubin into HepG2 cells assayed by thermal lens spectroscopy. Function of bilitranslocase.* The FEBS Journal, **272,** 5522 (2005)

83. S. Passamonti, L. Battiston and G. L. Sottocasa, *Bilitranslocase can exist in two metastable forms with different affinities for the substrates--evidence from cysteine and arginine modification.* European Journal of Biochemistry / FEBS, **253,** 84 (1998)

84. G. L. Sottocasa, G. Baldini, G. Sandri, G. Lunazzi and C. Tiribelli, *Reconstitution in vitro of sulfobromophthalein transport by bilitranslocase.* Biochimica et Biophysica Acta (BBA) - Biomembranes, **685,** 123 (1982)

85. S. Passamonti, U. Vrhovsek and F. Mattivi, *The interaction of anthocyanins with bilitranslocase.* Biochemical and Biophysical Research Communications, **296,** 631 (2002)

86. M. Novic and M. Vracko, *Comparison of spectrum-like representation of 3D chemical structure with other representations when used for modelling biological activity.* Chemometrics and Intelligent Laboratory Systems, **59,** 33 (2001)

87. A. Karawajczyk, V. Drgan, N. Medic, G. Oboh, S. Passamonti and M. Novic, *Properties of flavonoids influencing the binding to bilitranslocase investigated by neural network modelling.* Biochemical Pharmacology, **73,** 308 (2007)

88. J. Dayhof, *Neural network architectures, An Introduction.* Van Nostrand Reinhold, (1990)

89. J. Zupan, M. Novic and I. Ruisanchez, *Kohonen and counterpropagation artificial neural networks in analytical chemistry.* Chemometrics and Intelligent Laboratory Systems, **38,** 1 (1997)

90. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. P. Stewart, *AM1: A new general purpose quantum mechanical molecular model.* Journal of the American Chemical Society, **107,** 3902 (1985)

91.   A. R. Katritzy, V. S. Lobanov and M. Karelnos, *CODESSA 2.0,* University of Florida U.S.A. (1990)

92.   J. Zupan and M. Novic, *Optimisation of structure representation for QSAR studies.* Analytica Chimica Acta, **388,** 243 (1999)

93.   S. Passamonti and G. L. Sottocasa, *The quinoid structure is the molecular requirement for recognition of phthaleins by the organic anion carrier at the sinusoidal plasma membrane level in the liver.* Biochimica et Biophysica Acta (BBA) - Biomembranes, **943,** 119 (1988)

94.   K. Lemanska, H. Szymusiak, B. Tyrakowska, R. Zielinski, A. E. M. F. Soffers and I. M. C. M. Rietjens, *The influence of pH on antioxidant properties and the mechanism of antioxidant action of hydroxyflavones.* Free Radical Biology and Medicine, **31,** 869 (2001)

95.   S. Passamonti, A. Vanzo, U. Vrhovsek, M. Terdoslavich, A. Cocolo, G. Decorti and F. Mattivi, *Hepatic uptake of grape anthocyanins and the role of bilitranslocase.* Food Research International, **38,** 953 (2005)

96.   L. R. Huang, Y. Xie and J. W. Lown, *Bleomycin antibiotics and their role in cancer chemotherapy.* Expert Opinion on Therapeutic Patents, **6,** 893 (1996)

97.   S. J. Brown, S. E. Hudson, D. W. Stephan and P. K. Mascharak, *Syntheses, structures, and spectral properties of a synthetic analog of copper(II)-bleomycin and an intermediate in the process of its formation.* Inorganic Chemistry, **28,** 468 (1989)

98.   J. D. Tan, S. E. Hudson, S. J. Brown, M. M. Olmstead and P. K. Mascharak, *Syntheses, structures, and reactivities of synthetic analogs of the 3 forms of Co(III)-bleomycin - Proposed mode of light-induced DNA damage by the Co(III) chelate of the drug.* Journal of the American Chemical Society, **114,** 3841 (1992)

99.   H. Kurosaki, K. Hayashi, Y. Ishikawa and M. Goto, *The first trigonal-bipyramidal structure of zinc(II) complex of a bleomycin model.* Chemistry Letters, **8,** 691 (1995)

100.   S. J. Brown, M. M. Olmstead and P. K. Mascharak, *Iron(II) and iron(III) complexes of N-(2-(4- imidazolyl)ethyl)pyrimidine-4-carboxamide, a ligand resembling part of the metal-binding domain of bleomycin.* Inorganic Chemistry, **29,** 3229 (1990)

101.   T. E. Lehmann, *Molecular modeling of the three-dimensional structure of Fe(II)-bleomycin: are the Co(II) and Fe(II) adducts isostructural?* Journal of Biological Inorganic Chemistry, **7,** 305 (2002)

102.   T. E. Lehmann, M. L. Serrano and L. Que, *Coordination chemistry of Co(II)-bleomycin: Its investigation through NMR and molecular dynamics.* Biochemistry, **39,** 3886 (2000)

103.   K. E. Loeb, J. M. Zaleski, C. D. Hess, S. M. Hecht and E. I. Solomon, *Spectroscopic investigation of the metal ligation and reactivity of the ferrous active sites of bleomycin and bleomycin derivatives.* Journal of the American Chemical Society, **120,** 1249 (1998)

104.   K. E. Loeb, J. M. Zaleski, T. E. Westre, R. J. Guajardo, P. K. Mascharak, B. Hedman, K. O. Hodgson and E. I. Solomon, *Spectroscopic definition of the geometric and electronic- structure of the nonheme iron active-site in iron(II) bleomycin - correlation with oxygen reactivity.* Journal of the American Chemical Society, **117,** 4545 (1995)

105.   J. P. Piquemal, B. Williams-Hubbard, N. Fey, R. J. Deeth, N. Gresh and C. Giessner-Prettre, *Inclusion of the ligand field contribution in a polarizable molecular mechanics: SIBFA-LF (vol 24, pg 1963, 2003).* Journal of Computational Chemistry, **25,** 308 (2004)

106.   N. Oppenheimer, L. Rodriguez and S. Hecht, *Structural studies of "active complex" of bleomycin: Assignment of ligands to the ferrous ion in a ferrous-bleomycin-carbon monoxide complex.* Proc. Natl. Acad. Sci. USA, **76,** 5616 (1979)

107. M. A. J. Akkerman, E. Neijman, S. S. Wijmenga, C. W. Hilbers and W. Bermel, *Studies of the solution structure of the bleomycin-A2 iron(II) carbon-monoxide complex by means of 2-dimensional NMR- spectroscopy and distance geometry calculations.* Journal of the American Chemical Society, **112,** 7462 (1990)

108. A. Fouqueau, S. Mer, M. E. Casida, L. M. L. Daku, A. Hauser, T. Mineva and F. Neese, *Comparison of density functionals for energy and structural differences between the high-[T-5(2g):(t(2g))(4)(e(g))(2)] and low-[(1)A(1g):(t(2g))(6)(e(g))(0)] spin states of the hexaquoferrous cation [Fe(H₂O)(6)](2+).* Journal of Chemical Physics, **120,** 9473 (2004)

109. M. Swart, A. R. Groenhof, A. W. Ehlers and K. Lammertsma, *Validation of exchange - correlation functionals for spin states of iron complexes.* Journal of Physical Chemistry A, **108,** 5479 (2004)

110. P. J. Hay and W. R. Wadt, *Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg.* Journal of Chemical Physics, **82,** 270 (1985)

111. T. H. D. Jr and P. J. Hay, *Methods of Electronic Structure Theory.* Modern Theoretical Chemistry, **3,** 1 (1976)

112. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A.D. Daniels, K. N. Kudin, M. C. Strain, Ö. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komáromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle and J. A. Pople, *Gaussian 03*, Pittsburgh (2003)

113. G. Smolentsev, A. V. Soldatov, E. C. Wasinger and E. I. Solomon, *Axial ligation of Fe(II)-bleomycin probed by XANES spectroscopy.* Inorganic Chemistry, **43,** 1825 (2004)

114. I. Morgenstern-Badarau, F. Lambert, J. Philippe Renault, M. Cesario, J.-D. Marechal and F. Maseras, *Amine conformational change and spin conversion induced by metal-assisted ligand oxidation: from the seven-coordinate iron(II)-TPAA complex to the two oxidized iron(II)-(py)3tren isomers. Characterization, crystal structures, and density functional study.* Inorganica Chimica Acta, **297,** 338 (2000)

115. D. G. Lonnon, D. C. Craig and S. B. Colbran, *Monomeric and dimeric metal complexes of a simply prepared and versatile pentapyridyldiamine.* Inorganic Chemistry Communications, **5,** 958 (2002)

116. L. Sacconi and M. d. Vaira, *Correlation between spin state and nature of donor atoms for five-coordinate iron(II) complexes. X-ray structure determination of two iron(II) complexes with tris(2-diphenylphosphinoethyl)amine and tris(2-diphenylphosphinoethyl)phosphine.* Inorganic Chemistry, **17,** 810 (1978)

117. M. A. J. Akkerman, C. A. G. Haasnoot, U. K. Pandit and C. W. Hilbers, *Complete assignment of the C-13 NMR-spectra of bleomycin-A2 and its zinc complex by means of two-dimensional NMR-Spectroscopy.* Magnetic Resonance in Chemistry, **26,** 793 (1988)

118. F. Fedeles and M. Zimmer, *Conformational/configurational analysis of all the binding geometries of cobalt(III) bleomycin.* Inorganic Chemistry, **40,** 1557 (2001)

119. M. Sugiyama, T. Kumagai, M. Hayashida, M. Maruyama and Y. Matoba, *The 1.6-A crystal structure of the copper(II)-bound bleomycin complexed with the bleomycin-binding*

protein from bleomycin-producing Streptomyces verticillus. Journal of Biological Chemistry, **277,** 2311 (2002)

120. R. M. Burger, T. A. Kent, S. B. Horwitz, E. Munck and J. Peisach, *Mossbauer study of iron bleomycin and its activation intermediates.* Journal of Biological Chemistry, **258,** 1559 (1983)

121. R. M. Burger, J. Peisach and S. B. Horwitz, *Activated bleomycin - A transient complex of drug, iron, and oxygen that degrades DNA.* Journal of Biological Chemistry, **256,** 1636 (1981)

122. J. Peisach, R. M. Burger, S. Band Horwitz, T. A. Kent and E. Munck, *The electronic structure of iron complexes of bleomycin.* Inorganica Chimica Acta, **79,** 33 (1983)

123. M. Freindorf and P. M. Kozlowski, *DFT study of the metal coordination center domain of Fe(II)- bleomycin.* Journal of Physical Chemistry A, **105,** 7267 (2001)

124. J. W. Sam, X. J. Tang and J. Peisach, *Electrospray mass-spectrometry of iron bleomycin - Demonstration that activated bleomycin is a ferric peroxide complex.* Journal of the American Chemical Society, **116,** 5250 (1994)

125. T. E. Westre, K. E. Loeb, J. M. Zaleski, B. Hedman, K. O. Hodgson and E. I. Solomon, *Determination of the geometric and electronic-structure of activated bleomycin using X-Ray-absorption spectroscopy.* Journal of the American Chemical Society, **117,** 1309 (1995)

126. E. I. Solomon, T. C. Brunold, M. I. Davis, J. N. Kemsley, S. K. Lee, N. Lehnert, F. Neese, A. J. Skulan, Y. S. Yang and J. Zhou, *Geometric and electronic structure/function correlations in non-heme iron enzymes.* Chemical Reviews, **100,** 235 (2000)

127. R. M. Burger, J. Peisach, W. E. Blumberg and S. B. Horwitz, *Iron-bleomycin interactions with oxygen and oxygen analogs effects on spectra and drug activity.* Journal of Biological Chemistry, **254,** 906 (1979)

128. C. C. Cheng, J. G. Goll, G. A. Neyhart, T. W. Welch, P. Singh and H. H. Thorp, *Relative rates and potentials of competing redox processes during DNA cleavage - Oxidation mechanisms and sequence-specific catalysis of the self-inactivation of oxometal oxidants by DNA.* Journal of the American Chemical Society, **117,** 2970 (1995)

129. M. Nakamura and J. Peisach, *Self-inactivation of Fe(II)-bleomycin.* Journal of Antibiotics, **41,** 638 (1988)

130. B. Ensing, F. Buda and E. J. Baerends, *Fenton-like chemistry in water: Oxidation catalysis by Fe(III) and H_2O_2.* Journal of Physical Chemistry A, **107,** 5722 (2003)

131. F. Neese, J. M. Zaleski, K. L. Zaleski and E. I. Solomon, *Electronic structure of activated bleomycin: Oxygen intermediates in heme versus non-heme iron.* Journal of the American Chemical Society, **122,** 11703 (2000)

132. T. Owa, T. Sugiyama, M. Otsuka, M. Ohno and K. Maeda, *Synthetic studies on antitumor antibiotic bleomycin. A model study on the mechanism of the autoxidation of bleomycin.* Tetrahedron Letters, **31,** 6063 (1990)

133. M. R. Bukowski, S. R. Zhu, K. D. Koehntop, W. W. Brennessel and L. Que, *Characterization of an Fe-III-OOH species and its decomposition product in a bleomycin model system.* Journal of Biological Inorganic Chemistry, **9,** 39 (2004)

134. G. Padbury, S. G. Sligar, R. Labeque and L. J. Marnett, *Ferric bleomycin catalyzed reduction of 10-hydroperoxy-8,12-octadecadienoic acid - Evidence for homolytic O-O bond scission.* Biochemistry, **27,** 7846 (1988)

135. L. W. Oberley and G. R. Buettner, *Production of hydroxyl radical by bleomycin and iron(II).* FEBS Letters, **97,** 47 (1979)

136. Y. Sugiura and T. Kikuchi, *Formation of superoxide and hydroxy radicals in iron(II)-bleomycin-oxygen system - Electron-spin resonance detection by spin trapping.* Journal of Antibiotics, **31,** 1310 (1978)

137. A. J. Allentoff, J. L. Bolton, A. Wilks, J. A. Thompson and P. R. O. Demontellano, *Heterolytic versus homolytic peroxide bond-cleavage by sperm whale myoglobin and myoglobin mutants.* Journal of the American Chemical Society, **114,** 9744 (1992)

138. J. N. Rodriguez-Lopez, D. J. Lowe, J. Hernandez-Ruiz, A. N. P. Hiner, F. Garcia-Canovas and R. N. F. Thorneley, *Mechanism of reaction of hydrogen peroxide with horseradish peroxidase: Identification of intermediates in the catalytic cycle.* Journal of the American Chemical Society, **123,** 11838 (2001)

139. A. Karawajczyk and F. Buda, *The metal bonding domain of the antitumor drug Fe(II)-bleomycin: a DFT investigation.* Journal of Biological Inorganic Chemistry, **10,** 33 (2005)

140. N. Troullier and J. L. Martins, *Efficient pseudopotentials for plane-wave calculations.* Physical Review B, **43,** 1993 (1991)

141. H. J. C. Berendsen, G. Postma, W. F. van Gunsteren and J. Hermans, *Interactions models for water in relation to protein hydratation.* in Intermolecular Forces Journal, 331, Dodrecht (1981)

142. G. Velde te, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. Gisbergen van, J. G. Snijders and T. Ziegler, *Chemistry with ADF.* Journal of Computational Chemistry, **22,** 931 (2001)

143. A. Natrajan, S. M. Hecht, G. A. Vandermarel and J. H. Vanboom, *A study of O_2-supported versus H_2O_2-supported activation of Fe-bleomycin.* Journal of the American Chemical Society, **112,** 3997 (1990)

144. O. A. Vydrov and G. E. Scuseria, *A simple method to selectively scale down the self-interaction correction.* Journal of Chemical Physics, **124,** (2006)

145. J. VandeVondele and M. Sprik, *A molecular dynamics study of the hydroxyl radical in solution applying self-interaction-corrected density functional methods.* Physical Chemistry Chemical Physics, **7,** 1363 (2005)

146. I. Arends, K. U. Ingold and D. D. M. Wayner, *A mechanistic probe for oxygen activation by metal-complexes and hydroperoxides and its application to alkane functionalization by [Fe(III)Cl(2)Tris(2-Pyridinylmethyl)Amine](+)Bf4.* Journal of the American Chemical Society, **117,** 4710 (1995)

147. P. A. MacFaul, K. U. Ingold, D. D. M. Wayner and L. Que, *A putative monooxygenase mimic which functions via well-disguised free radical chemistry.* Journal of the American Chemical Society, **119,** 10594 (1997)

148. C. Nguyen, R. J. Guajardo and P. K. Mascharak, *[Fe-III(PMA)](2+): A mononuclear non-heme iron complex that catalyzes alkane oxidation.* Inorganic Chemistry, **35,** 6273 (1996)

149. A. Laio, J. VandeVondele and U. Rothlisberger, *D-RESP: Dynamically Generated Electrostatic Potential Derived Charges from Quantum Mechanics/Molecular Mechanics Simulations.* J. Phys. Chem. B, **106,** 7300 (2002)

150. A. Bassan, M. R. A. Blomberg, T. Borowski and P. E. M. Siegbahn, *Theoretical studies of enzyme mechanisms involving high-valent iron intermediates.* Journal of Inorganic Biochemistry, **100,** 727 (2006)

151. L. Bernasconi, M. J. Louwerse and E. J. Baerends, *The role of equatorial and axial ligands in promoting the activity of non-heme oxidoiron(IV) catalysts in alkane hydroxylation.* European Journal of Inorganic Chemistry, in press (2007)

152. B. Ensing, F. Buda, M. C. M. Gribnau and E. J. Baerends, *Methane-to-methanol oxidation by the hydrated iron(IV) oxo species in aqueous solution: A combined DFT and Car-Parrinello molecular dynamics study.* Journal of the American Chemical Society, **126,** 4355 (2004)

153. C. Michel, A. Laio, F. Mohamed, M. Krack, M. Parrinello and A. Milet, *Free energy ab initio metadynamics: A new tool for the theoretical study of organometallic reactivity?*

Example of the C-C and C-H reductive eliminations from platinum(IV) complexes. Organometallics, **26,** 1241 (2007)

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

SUMMARY

Bleomycin is a DNA-cleaving antibiotic successfully used in the chemotherapy against several types of cancer like head and neck cancer or certain lymphomas and testicular cancer. Although it has been in use for more than two decades, the mechanism of its action is not known. Thus the harmful side effects are difficult to eliminate. On the other hand the process of design or improvement of pharmaceuticals is extremely complex and expensive. Moreover, the chemical drug discovery is often associated with the synthesis of an excessive number of chemical compounds and tests on animals. In the era of human genome being deciphered and with 3-dimensional structures of proteins available in an increasing number, drug companies will have to increase the utilization of information technology and computer science within the traditional drug design process to improve quality and efficiency of their output. A new trend within drug discovery is emerging with the application of clean chemistry, by performing molecular modeling of new compounds and by running virtual tests to assess their suitability before an expensive synthesis attempt is made. Virtual experiments can help scientists to eliminate compounds that would not perform the required function. It is now very difficult to imagine how a new drug could be designed in the 21st century with absolutely no contribution from computer aided drug design, whether reported or not. Computer aided drug design should be further integrated with traditional medicinal and synthetic chemistry, as well as biology and pharmacology to maximize its impact.

In the introductory chapter the work on the bleomycin anticancer drug is put in the broader context of the drug design process. Here I discuss the contribution of

different computational methods into this field, emphasizing the growing role played by quantum mechanical methods. Special attention is given to first principles molecular dynamics (Car-Parrinello) simulations based on density functional theory that have emerged as a powerful tool for studying physical, chemical and biological systems. Its implementation into a QM/MM approach is especially attractive for the *in situ* investigation of chemical reactions that occur in a complex and heterogeneous environment. In Chapter 2 the theoretical backgrounds of all the computational techniques applied in the studies of the mechanism of BLM action, are presented.

Chapter 3 contains a complete study done by chemometrics, one of the most common tools in the drug design process. The system under investigation is the membrane protein bilitranslocase and its ligands. This chapter shows for a set of compounds how, by using a powerful method known as counter propagation artificial neural network, one can determine relevant features that correlate with their activity against bilitranslocase. As discussed in the introductory chapter a quantum mechanics based investigation will be a natural further step to verify the conclusions and analyse the activity of potential drugs.

The following chapters in the thesis focus on the application of quantum mechanical methods to the anticancer drug bleomycin. At first the investigation of the geometric and electronic structure of ferrous complexes of bleomycin (Fe(II)BLM) by means of density functional theory calculations is presented in Chapter 4. Because of the presence of the transition metal center in the complex, a simple force-field molecular mechanics approach is not accurate enough to capture the structural details dictated by different spin states of the complex. The active site of this antitumor drug turns out to be a highly distorted octahedral complex with the coordination sphere completed by the five known endogenous ligands including pyrimidine, imidazole, deprotonated amide, and secondary and primary amine. The controversial issue of the nature of the sixth axial ligand has been addressed in this chapter and it has been identified as the oxygen of the carbamoyl group. In addition it is discussed here how the structure of the Fe(II)BLM complex depends on the total spin of the complex and the advantage of

using DFT calculations in this investigation. The calculations provide evidence for an easy substitution of the oxygen axial ligand due to the specific geometry of the complex. This leads to the formation of the activated bleomycin complex that is believed to be an active species in DNA degradation. The second axial ligand is substituted first by an oxygen molecule and then the BLM-Fe(III)-OOH complex is formed.

The active form of bleomycin is unstable and subject to self-inactivation. Therefore the possible reaction pathways for the decay of the active bleomycin-Fe(III)-OOH complex, the bleomycin suicide, were investigated using first-principles molecular dynamics simulations (Car-Parrinello method). The outcome is presented in Chapter 5. The theoretical model of activated bleomycin contains the whole metal bonding domain of the bleomycin ligand. Simulations performed both in a vacuum and in water solution show that a facile decaying process involves a homolytic O-O bond cleavage with an almost simultaneous hydrogen atom abstraction. The formation of an intra- or intermolecular hydrogen bond appears to be crucial for the decay of the activated bleomycin. No evidence of heterolytic cleavage of the O-O bond of the Fe(III)-OOH species has been observed, contrary to previously proposed models.

In Chapter 6 the reaction between activated BLM and the deoxyribose sugar in DNA was investigated in various environments by systematically extending the model. The simulations performed in vacuum and within a QM/MM approach where the DNA and aqueous solution are explicitly included in the model confirm the mechanism for O-O bond activation observed already in the ABLM suicide study. The important finding here is that the selectivity of the H-atom abstraction depends directly on the pre-formed hydrogen bond facilitating the homolytic O-O bond cleavage. Finally, new insights into the mechanism of double stranded DNA degradations are presented in Chapter 6 where it is shown how quantum mechanical tools may provide an answer on how one ABLM molecule destroys both strands of DNA.

SAMENVATTING

Bleomycine is een DNA-splitsend antibioticum dat met succes wordt gebruikt in de chemotherapie tegen verschillende typen van kanker zoals hoofd- en nek-kanker, bepaalde lymfomen en teelbalkanker. Hoewel het al een twintigtal jaar in gebruik is, is het reactiemechanisme nog steeds onopgehelderd. Het is daarom moeilijk schadelijke neveneffecten te elimineren. Daarnaast is het ontwikkelen en optimaliseren van medicijnen een buitengewoon complex en duur proces. Medicijnontwikkeling gaat veelal gepaard met de synthese van een buitensporig aantal van chemische verbindingen en dierproeven. In het tijdperk van de oplossing van het humane genoom en het in toenemende mate beschikbaar komen van 3-dimensionale eiwitstructuren, worden farmaceutische bedrijven gedwongen meer informatietechnologie en computerwetenschappen in te zetten bij het traditionele medicijnontwikkelingsproces om zo tot een verbeterde kwaliteit en efficiëntie van de productie te komen. Een nieuwe trend binnen de medicijnontwikkeling is de toepassing van zogenaamde schone chemie, door het gebruik van molecular modeling en het doen van virtuele tests voor het bepalen van de geschiktheid van nieuwe verbindingen alvorens een poging tot synthese te ondernemen. Virtuele experimenten kunnen behulpzaam zijn bij het elimineren van verbindingen die de vereiste werking ontberen. Het is inmiddels moeilijk voor te stellen hoe nieuwe medicijnen in de eenentwintigste eeuw ontwikkeld zouden kunnen worden zonder de inzet van computerondersteunde medicijn ontwikkeling (computer aided drug design). Computer aided drug design zou verder geïntegreerd moeten worden met de traditionele medicinale en

synthetische chemie, als ook biologie en farmacologie ten einde zijn volle invloed te laten gelden.

In het inleidende hoofdstuk, wordt het werk aan bleomycine (BLM), een medicijn tegen kanker, gezet in de bredere context van het medicijnontwikkelinsproces. Hier worden de bijdragen besproken van verschillende rekenmethoden aan dit veld, waarbij de nadruk wordt gelegd op de groeiende rol die kwantummechanische methoden spelen. Speciale aandacht wordt besteed aan de moleculaire dynamica simulaties gebaseerd op 'first principles' density functional theory (Car-Parrinello) welke hun nut hebben bewezen als krachtig gereedschap voor het bestuderen van fysische, chemische en biologische systemen. De implementatie in de QM/MM benadering is in het bijzonder aantrekkelijk voor het *in situ* bestuderen van chemische reacties die optreden in een complexe en heterogene omgeving. In het tweede hoofdstuk wordt de theoretische achtergrond gepresenteerd van alle bij de bestudering van het werkingsmechanisme van BLM toegepaste rekentechnieken.

Hoofdstuk 3 bevat een volledige studie met behulp van chemometrie, een van de gangbare tools in het medicijnontwikkelingsproces. Het bestudeerde systeem is het membraaneiwit bilitranslocase en haar liganden. Dit hoofdstuk laat voor een set verbindingen zien hoe bepaald wordt welke van hun eigenschappen correleren met hun biologische activiteit tegen bilitranslocase door het gebruik van de krachtige methode als een counter propagation artificial neural network. Zoals besproken in het inleidende hoofdstuk is de bestudering met een kwantummechanische methode de volgende logische stap bij de verificatie van de conclusies en bij de analyse van de werkzaamheid van potentiële medicijnen.

De volgende hoofdstukken in het proefschrift besteden aandacht aan de toepassing van kwantummechanische methoden op het kankermedicijn bleomycine. Eerst wordt het onderzoek naar de geometrische en elektronische structuur van IJzer(II) complexen van bleomycine (Fe(II)BLM) door middel van de density functional theory berekeningen gepresenteerd in Hoofdstuk 4. Door de aanwezigheid van een overgangsmetaal centrum in het complex, is de aanpak met eenvoudige force-field moleculaire mechanica niet nauwkeurig genoeg om de

invloed van verschillende spintoestanden op de structuur van het complex te beschrijven. De active site van deze antitumordrug blijkt een sterk vervormd octahedraal complex te zijn waarbij de coördinatieplaatsen bezet worden door de vijf bekende endogene liganden waaronder pyrimidine, imidazole, gedeprotoneerd amide, secondaire en primaire amine. Het controversiële punt van de identiteit van het zesde axiale ligand wordt in dit hoofdstuk behandeld en wordt geïdentificeerd als de zuurstof van de carbamoyl groep. Tevens wordt hier besproken hoe de structuur van het Fe(II)BLM complex afhangt van de totale spin van het complex en het voordeel van het gebruik van DFT berekeningen in dit onderzoek. De berekeningen ondersteunen een makkelijke vervanging van het axiale zuurstofligand ten gevolge van de specifieke geometrie van het complex. Dit leidt tot de vorming van een geactiveerd bleomycine complex dat verondersteld wordt een actieve rol te spelen bij de DNA degradatie. Het tweede axiale ligand wordt eerst vervangen door een zuurstofmolecuul en daarna wordt het BLM-Fe(III)-OOH complex gevormd.

De actieve vorm van bleomycine is instabiel en onderhevig aan zelfinactivatie. Daarom werden de mogelijke reactiepaden voor het verval van het actieve bleomycine-Fe(III)-OOH complex, de bleomycine zelfmoord, onderzocht met behulp van first-principles moleculaire dynamica simulaties (Car-Parrinello methode). Het resultaat wordt gepresenteerd in Hoofdstuk 5. Het theoretische model van geactiveerd bleomycine bevat het gehele metaal bindend domein van het bleomycineligand. Simulaties uitgevoerd zowel in vacuüm als in wateroplossing laten een gemakkelijk vervalproces zien waarbij een O-O binding homolytisch wordt verbroken gepaard gaande met een vrijwel onmiddellijke waterstofatoomabstractie. De vorming van een intra- of inter- moleculaire waterstofbrug blijkt cruciaal te zijn voor het verval van het geactiveerd bleomycine. Er is geen bewijs gevonden van een heterolytische verbreken van de O-O binding van het Fe(III)-OOH complex, in tegenstelling tot eerder voorgestelde modellen.

In Hoofdstuk 6 wordt de reactie tussen geactiveerd BLM en de deoxyribose suiker in DNA onderzocht in verschillende omgevingen door systematische uitbreiding

van het model. De simulaties uitgevoerd in vacuüm en binnen een QM/MM benadering, waarbij het DNA en de oplossing in water expliciet worden meegenomen in het model, bevestigen het mechanisme voor O-O bindingsactivering dat al eerder is vastgesteld in the studie van de ABLM-zelfmoord. De belangrijke bevinding hier is dat de selectiviteit van de H-atoom abstractie direct afhangt van de eerder gevormde waterstofbrug die vervolgens de homolytische splitsing van de O-O binding mogelijk maakt. Ten slotte worden nieuwe inzichten in het mechanisme van de degradatie van double stranded DNA gepresenteerd in Hoofdstuk 6 waar behandeld wordt hoe kwantummechanische tools een antwoord kunnen leveren op de vraag hoe een ABLM-molecuul beide strengen van het DNA kan afbreken.

PODSUMOWANIE

Bleomycyna jest antybiotykiem z powodzeniem stosowanym w chemoterapii przeciwko różnego rodzaju chorobom nowotworowym takim jak nowotwory w obrębie głowy i karku, ziarnica złośliwa czy też chłoniaki nieziarnicze. Chociaż jest stosowana od 1980 roku, mechanizm jej działania nie jest do końca wyjaśniony. Tym samym efekty uboczne są trudne do wyeliminowania. Z drugiej strony, proces opracowywania nowego leku lub ulepszenia istniejącego już na rynku jest bardzo skomplikowany i kosztowny. Chemiczny proces opracowywania leków jest kojarzony z syntezą niezliczonej liczby związków i doświadczeniami przeprowadzanymi na zwierzętach. Współcześnie obraz ten nie odpowiada jednak rzeczywistości. W wieku, kiedy ludzki genom został zmapowany, a ilość dostępnych 3-D struktur białek wciąż rośnie, firmy farmaceutyczne zwiększają wykorzystywanie komputerów i technologii informatycznych w procesie opracowywania nowych leków, aby zwiększyć jakość i wydajność procesu. Ten nowy trend w dziedzinie opracowywania nowych farmaceutyków jest związany z zastosowaniem tak zwanej czystej chemii, czyli modelowania molekularnego nowych związków i ich właściwości zanim jeszcze zostaną podjęte próby drogiej eksperymentalnej syntezy. Wirtualne eksperymenty mogą pomóc naukowcom w wyeliminowaniu związków, które nie spełniałyby stawianych wymagań, we wczesnym stadium procesu. Trudno jest sobie wyobrazić w jaki sposób nowy lek mógłby być wynaleziony i zoptymalizowany w XXI wieku bez udziału modelowania komputerowego. Komputerowo wspomagany proces opracowywania nowych leków powinien być

ściśle zintegrowany z tradycyjną medyczną i syntetyczną chemią jak również z biologią i farmakologią, aby zmaksymalizować jego użyteczność.

W rozdziale wprowadzającym praca dotycząca przeciwnowotworowego leku bleomycyny została włączona w szerszy kontekst procesu opracowywania nowego leku. Tutaj to dyskutowany jest wkład różnych metod obliczeniowych używanych w tej dziedzinie, podkreślając rosnącą role odgrywaną przez metody kwantowomechaniczne. Szczególna uwaga skierowana jest na dynamikę molekularną *ab initio* (Car-Parrinello) bazującą na teorii funkcjonałów gęstości, które, ostatnimi czasy, stają się potężnym narzędziem do studiowania fizycznych, chemicznych i biologicznych układów. Uczynienie tej metody częścią hybrydy łączącej mechanikę kwantową z mechaniką molekularną (QM/MM) jest wyjątkowo atrakcyjne dla badania *in situ* reakcji chemicznych, które mają miejsce w kompleksowym i heterogenicznym środowisku. W kolejnym rozdziale pracy zaprezentowane są podstawy teoretyczne wszystkich obliczeniowych technik wykorzystywanych do przeprowadzenia badań nad mechanizmem działania bleomycyny.

Rozdział trzeci to kompletne studium wykonane przy użyciu chemometrii, jednej z najpowszechniej wykorzystywanych metod w procesie opracowywania nowych leków. Układ poddany badaniom to bilitranslokaza i jej ligandy. Rozdział ten ukazuje dla serii związków, jak przy użyciu skutecznej metody znanej jako counter propagation artificial neural network, można ustalić ważne właściwości, które mają odniesienie do aktywności wykazywanej w stosunku do bitranslokazy. Jak zostało to przedstawione we wprowadzeniu, badania bazujące na mechanice kwantowej byłyby naturalną kontynuacją tego projektu. Dzięki nim można byłoby zweryfikować konkluzje i zanalizować aktywność potencjalnych leków.

Następne trzy rozdziały pracy skupiają się na wykorzystywaniu metod kwantowochemicznych do badań nad przeciwnowotworowym lekiem bleomycyną. Na początek badania geometrycznej i elektronowej struktury kompleksu Fe(II)BLM przedstawione są w rozdziale czwartym. Ze względu na obecność metalu przejściowego w kompleksie, proste podejście bazujące na polu siłowym nie jest wystarczająco dokładne, aby uchwycić szczegóły strukturalne

zależne od stanu spinowego kompleksu, dlatego też obliczenia przeprowadzono głównie za pomocą metod bazujących na teorii funkcjonałów gęstości. Aktywna forma tego przeciwnowotworowego leku okazała się być wysoce zdeformowanym oktaedrycznym kompleksem ze sferą koordynacyjną zawierającą 5 endogenous ligandów w tym pirimidinę, imidazol, deprotonowany amid, oraz drugo- i pierwszorzędową aminę. W tym rozdziale dyskutowana jest kontrowersyjna kwestia o naturę szóstego liganda, który ostatecznie został zidentyfikowany jako atom tlenu pochodzący od karbomylowj grupy. Dodatkowo jest tutaj przedyskutowana kwestia, w jaki sposób struktura kompleksu Fe(II)BLM zależy od całkowitego spinu kompleksu oraz korzyści płynące z użycia DFT obliczeń do prezentowanych badań. Obliczenia dostarczyły dowodów na łatwe zastąpienie aksjalnego liganda tlenu pochodzącego od karbomylowej grupy przez cząsteczkę tlenu dzięki specyficznej geometrii kompleksu. Prowadzi to do powstania aktywnego kompleksu bleomycyny, która z kolei jest odpowiedzialna za degradację DNA.

Aktywna forma bleomycyny jest niestabilna i predysponowana do samo-inaktywacji. Dlatego też możliwe mechanizmy reakcji na inaktywowanie aktywnego kompleksu bleomycyna-Fe(III)-OOH, tak zwanego samobójstwa bleomycyny, były badane przy użyciu dynamiki molekularnej *ab initio* (metoda Car-Parrinello). Wyniki tych symulacji komputerowych są zaprezentowane w rozdziale piątym. Model aktywnej bleomycyny zawiera całą domenę odpowiedzialną za wiązanie metalu. Symulacje wykonane zarówno w próżni jak i w wodzie wykazały ze łatwy proces degradacji związany jest z homolitycznym rozpadem wiązania O-O wraz z prawie spontaniczną abstrakcją atomu wodoru. Formacja intra- lub intermolekularnego wiązania wodorowego okazała się być istotna dla rozpadu aktywnego kompleksu bleomycyny. Nie zostały zaobserwowane żadne przesłanki dla heterolitycznego rozpadu wiązania O-O we fragmencie Fe(III)-OOH, w przeciwieństwie do wcześniej proponowanych modeli.

W rozdziale szóstym przedstawione są badania reakcji między aktywnym kompleksem bleomycyny i deoksyrybozy cukrem DNA w różnych środowiskach.

Podsumowanie

Symulacje wykonane w próżni i przy użyciu QM/MM, gdzie DNA i cząsteczki wody są bezpośrednio włączone do modelu, potwierdzają mechanizm aktywacji wiązania O-O obserwowany już wcześniej w badaniach nad samoinaktywacją ABLM. Ważne odkrycie stanowi fakt, że selektywność abstrakcji atomu wodoru zależy od selektywności utworzonego wcześniej wiązania wodorowego, które umożliwia homolityczny rozpad wiązania O-O. Ostatecznie, nowe spojrzenie na mechanizm zniszczenia obu nici DNA jest zaprezentowany w rozdziale szóstym, gdzie ukazana jest również rola kwantowomechanicznych obliczeń w sformułowaniu hipotezy o tym jak jedna cząsteczka ABLM może zniszczyć obie nici DNA.

PUBLICATIONS

1. A. Karawajczyk & F. Buda, (2005) *The metal bonding domain of the antitumor drug Fe(II)-bleomycin: a DFT investigation,* Journal of Biological Inorganic Chemistry **10**(1): 33-40.

2. A. Karawajczyk, *et al.* (2005) *Properties of flavonoids influencing the binding to bilitranslocase investigated by neural network modelling*, Biochemical Pharmacology **73**(2): 308-320

3. A. Karawajczyk, C. Gossens, U. Roethlisberger & F. Buda (2006), *The mechanism of the bleomycin suicide: a Car-Parrinello molecular dynamics study*, J. Phys. Chem. B, **110**(42): 21245-21250.

4. A. Karawajczyk, & F. Buda, (2006), *The anticancer drug bleomycin investigated by density functional theory*, Molecular Simulation. **32**(15): 1233-1239.

5. A. Karawajczyk & F. Buda, *The mechanism of DNA degradation by activated bleomycin,* manuscript in preparation

CURRICULUM VITAE

Since October 1996 I was a Chemistry student at the University of Wroclaw. As one of the best chemistry students I followed the Individual Track of Study since 2000. At the same time I joined the research group of the Department of Theoretical Modeling of Chemical Processes and started the master project under the supervision of Prof. dr. Zdzislaw Latajka. The subject of my master work was: "The theoretical studies of interaction between copper and ammonia molecules as a simple model of hemocyanine". During my study I was awarded a special studentship of the Foundation of University of Wroclaw for the best chemistry student. In June 2002, I received the masters degree in General Chemistry.

In January 2003, I started my PhD research at the Leiden Institute of Chemistry under the supervision of Dr. Francesco Buda and Prof. dr. Huub J. M. de Groot. The research described in this thesis was presented at various meetings and conferences. This includes "Molecular Modelling 2006" (Australia, 2006); "Multi-scale modelling: Electrons, Molecules and (Bio)Materials",(Royal Netherlands Academy of Arts and Sciences, Amsterdam 2006); "Holland Research School of Molecular Chemistry Symposium" (Amsterdam, 2005); CPMD 2005 Workshop "*Ab-initio* Molecular Dynamics Simulations - from Solid State Physics to Chemistry and Biology" and the Symposium in Honor of Professor Michele Parrinello on the Occasion of his 60th Birthday, (Switzerland, 2005). During the PhD study I was granted an EMBO fellowship at the EPFL (2004, Switzerland) in the group of Prof. U. Rothlisberger and a Marie Curie fellowship at the Host Training Site no. HPMT-CT-2001-00240 of the National Institute of Chemistry, Ljubljana, Slovenia.

NAWOORD

A PhD project is like exploring the landscape of a free energy surface. It is not trivial to find an optimal pathway. Fortunately, during that time I met many precious people, who made this journey exiting and interesting. No acknowledgment is necessary if you have no friend but the point is that I have many of them and everyone has had a small or big contribution to my graduation. First, my husband Tomek should be mentioned who not only is beside me in every moment of my life and scientific work but also took care of the graphics in this thesis. Thank you, Tomek, for believing that everything is possible. Without your acceptance, patience and dedication the thesis wouldn't exist.

An educative experience for me was the short visit to the Prof. Ursula Röthlisberger's group at the EPFL in Lausanne. Thank you, Ursula, for giving me the opportunity to learn from you and to experience a real team work atmosphere in your group. Ute Rörig and Christofer Gossen are people who will remain in my memory and heart for a long time. Thank you for your time, help and friendship.

Dr. Marjana Nović introduced me to another dimension of modeling in a very efficient way. I very much appreciate the time I spent under your supervision, Marjana. I also need to say that Slovenia is the most beautiful and friendly country I know.

As a theoretician I was a bit like an island in the ssNMR group but the social atmosphere among the AIOs created a nice environment for a scientific work. Anna, Shipra, Esha, Niels, Piotr, Eugenio, Karthick and many other thank you very much. I enjoyed the time we spent together playing board games, chatting

and organizing BBQs. Special acknowledgments go to Alia and Jörg. It is a real pleasure to be your friend.

To work on the PhD project I left many people I love in Poland and here I want to thank them for their support and understanding. Mamo, Babciu, Marto, Łukaszu, Agnieszko jesteście i będziecie dla mnie przystanią i opoką. Bez Was mój świat byłby płytszy. Dziękuję Wam za modlitwę i za to że we mnie wierzycie i nie pozwalacie się poddawać. Dziekuję za Waszą pomoc. Kocham Was bardzo.

Arnhem, 24th of June   Anna Karawajczyk