Metallo-β-lactamase:
reactivity and site directed evolution pathways addressed by computational approaches

Thesis submitted for the degree of

Doctor Philosophiæ

Candidate:
Fabio Simona

Supervisors:
Prof. Paolo Carloni
Dr. Alessandra Magistrato

November, 2008

Contents

Introduction

The indiscriminate prescription of antibiotics by physicians, along with their incorrect use [1] [1] has increased the exposition of bacteria to antibiotics, and thus has created a favorable environment for the Darwinian evolution of resistant strains [2]. Further increase of drug resistance is caused by the unnecessary massive use of antibiotics (70% of the total market is in the US!) to animals crammed into the unhygienic crowded quarters of factories [3, 4].

Diseases like tubercolosis, gonorrhea, malaria, and childhood ear infections, are increasingly becoming hard to treat with antibiotic drugs, posing serious concern in the human public health [5, 6].

The problem is even more serious if one considers that already in 70's and 80's that modification of the chemical structure of the already known antibiotics turned out to be exhausted and, at the same time, pharmacological companies decided not design of totally new antibiotics [2].

As a result, the number of clinically useful antibiotics today remains almost the same of twenty years ago [2] and scientists in universities and in public research institutions are practically the only ones investigating bacterial drug resistance.

The most serious issues are connected to β-lactams antibiotics (Fig. 1.9 1.10 1.11), which are the most used class of drugs against pathologies connected to bacteria [7]. Since they are indicated for the prophylaxis and treatment of bacterial infections, they are used also for all surgeon interventions [8].

One of the primary β-lactam resistance strategy is the bacterial biosynthesis of the β-lactamases enzymes [9], which catalyze the hydrolytic cleavage of the β-lactam ring (Fig. 1).
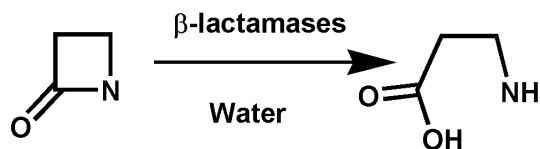
Table 1: General and schematic hydrolysis reaction of β-lactam ring catalyzed by β-lactamases.

---

[1]i.e. sporadic use of these drugs and/or incomplete treatment of bacterial infections.

These enzymes are divided into four classes: A, B, C and D. in which, according to the type of active site, one could identify two groups. Enzymes which belong to class A, C and D perform the hydrolytic reaction using a serine residue, while class B are Zn-dependent enzymes (called metallo-β-lactamases, MβLs). Although MβLs constitute only the 10% of the overall lactamase family [10], they represent a major concern in the clinical settings [11]. In fact: i) they feature a broad spectrum profile, which encompasses most β-lactam antibiotics; ii) there is the possibility of horizontal transfer (for details see section 1.1.1) among bacteria species of genes which encode the resistance to chemotherapy [12]; iii) clinically useful inhibitors are lacking.

MβLs exist in different subclasses (B1, B2 and B3), which differ in Zn content (one or two Zn ions) and substrate diversity [10]. B1 and B3 members are catalytically active *in vitro* in both the mono- and di-zinc forms, with different coordination polyhedra (Fig. 1). B2 members, on the other way, are mononuclear
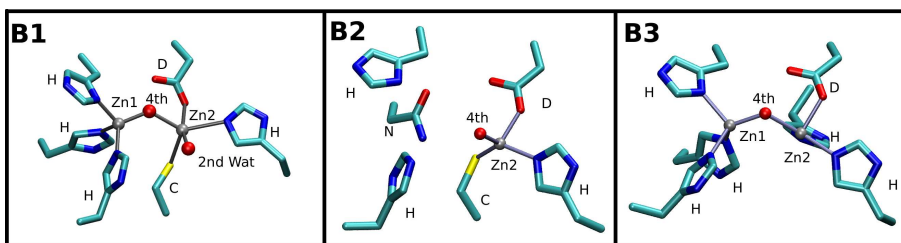
Figure 1: 'Zn1' and 'Zn2' binding sites in MβLs, as observed in the B1 BCII enzyme from *Bacillus cereus* [13] (residues are cut at the Cα, the B2 CphA enzyme from *Aeromonas hydrophila* [14] and the B3 L1 enzyme from *Stenotrophomonas maltophilia* [15]. In the case of the B1 enzyme, a second solvent molecule binds to the metal ion. The backbone units of the residues are not shown for clarity.

enzymes only (Fig. 1), and they show a selective carbapenemase activity. [2]

A large effort is being therefore made to develop MβLs inhibitors. Two major problems are associated with such drug development. The first is that eventual MβLs inhibitors which manage to block one of those may not be able to inhibit the other subclasses, with consequent little beneficial effect. Second, bacteria may express MβLs variants to counteract drug action, as already happened for serine-β-lactamases [19].

In this thesis, I have used molecular simulations approaches, to enlarge our understanding of MβLs structure and function. This in turn might help to address these very serious issues.

---

[2]Carbapenem are latest generation of β-lactam antibiotics, which have shown the broadest spectrum of activity within this antibiotic class. They exhibit *in vitro* bactericidal activity against numerous pathogens, including Gram-positive and Gram-negative aerobes and anaerobes (Fig. 1.11). In addition carbapenems are stable to almost all β-lactamases [16, 17, 18].

Concerning the first issue, I have formulated the following working hypothesis: it is possible that MβLs, in spite their differences in active site residues, metal content and activity, share common aspects in the catalytic reaction? If that would be the case, one might try to design transition state analogues of the enzymatic reaction, which may inhibit some or all MβLs subclasses, with increased beneficial effects.

This hypothesis has been prompted by a large body of work in our Sector few years ago. This work showed that a very large class of enzymes, the proteases, share striking functional similarities in spite of the different fold and active site architecture [20, 21].

At the time I had started this thesis, mechanistic studies had been carried out only for B1 subclass, by us [22, 23], and other research groups [24, 25, 26, 27, 28].

To make comparison across different subclasses, here I then focused on the CphA from *Aeromonas hydrophila*, the first enzyme belonging to B2 subclass whose X-ray structure was solved in complex with a partially hydrolyzed substrate [14]. I have used exactly the same computational setup, which consists of DFT calculations, classical MD calculations and hybrid Car-Parrinello QM/MM calculations, as that used for B1 [22, 23]: this was done to make our comparison meaningful.

Anticipating our results, we find the B1 and B2 subclasses share strikingly similar mechanisms of action, suggesting that the architecture and Zn content may vary across the enzymes, but mechanistic features may be conserved. This could be of help to design transition-state analogue inhibitors which block two of the three subclasses. The work has been done in collaboration with one of the crystallographers (Dr. G. Garau) and with Prof. A. Vila, Rosario, Argentina, one of the leading scientists investigating MβLs.

Several discussions with Vila's lab also prompted us to address a second issue. In a very important contribution [29], Vila's lab has challenged the B1 MβL, BcII from *Bacillus cereus*, to a direct evolution scheme, this has allowed to identify mutations, which may cause drug resistance, and yet they have not been isolated in the clinics.

The isolated BcII protein of the evolved *Bacillus cereus* bacteria was characterized by five point mutations. We investigated the structural and functional role of some of these mutations. Using classical MD simulations of the evolved BcII variants, we were able to provide a rational (at qualitative level) for the experimentally observed change in the catalytic rate.

A first mutation (G262S), close to the cleavage site, affects the reaction because it tunes the metal charge, leading to stabilization of a catalytically relevant intermediate. The second mutation (N70S-G262S), located far from the active site, causes an increase of the shape and flexibility of the active site cavity, which in turn may have an impact on the enzyme reactivity.

These findings could be exploited to propose a strategy for designing a mechanism-based inhibitor with an evolutionary perspective, namely anionic transition state analogues targeted to bind the "Zn2" site.

Part I

Biological overview and methods

Chapter 1

Bacteria

The phenomenon of bacterial resistance against pharmacological treatment with antibiotics has recently gained interest in scientific research, because it is becoming a serious public health issue.

Penicillin was the first chemical compound able to inhibit bacterial growth, and in some cases it is also able to cause lysis (named for that *antibiotics*). It was isolated in 1928 by the biologist and pharmacologist Sir Alexander Fleming, who immediately noticed the ability of bacteria to quickly develop resistance against these drugs. The discovery of new similar compounds, and the modification of the chemical structures of the already known antibiotics, allowed to initially circumvent the problem of resistance. Today a number of diseases, like tuberculosis, gonorrhea, malaria, and childhood ear infections, have become hard to treat with antibiotic drugs (http://www.fda.gov/oc/opacom/hottopics/anti_resist.html)

During the 2002, the U.S. Centers for Disease Control and Prevention (CDC) estimated that at least 90000 deaths a year only in the U.S. could be attributed to bacterial infections, more than half caused by bugs resistant to at least one commonly used antibiotic [30]

Disease-causing microbes that have become resistant to drug therapy are an increasing public health problem, that should be addressed soon.

1.1 Bacteria

In order to understand the biochemistry of resistance mechanism it is essential to give a biological introduction on bacteria.

Not all the bacteria are dangerous organisms for the human health, on the countrary, pathogenic bacteria are only a small portion of the bacteria as a whole. The interest toward these microorganisms is vast and can span from control of disease to biotechnology development.

Bacteria are microorganisms whose dimensions could range from (~ 0.1 μm to 500 μm), and which could have many different shapes (Fig. 1.1).

Bacteria are *prokaryotic* organisms, which differ from *eukaryotic* cells of animal and plants in many aspects: the most striking difference is that the former do not posses a nuclear membrane, and the "DNA material" is therefore in direct contact with the cytoplasm, while in the latter a double-layered nuclear membrane separates the genetic material from the cytosolic environment.

Another important difference is that all bacteria posses a biochemical structure, called cell wall, that surrounds the external cytoplasmic membrane and provides structural integrity to the cell, while this is rarely present in *eukaryotic* cells.

1.1.1   Bacteria growth and reproduction

In contrast to most eukaryotes, prokaryotes reproduce asexually. While sexual reproduction in eukaryotes results in offspring with genetic material which is a mixture of the parents X Y genome, a prokaryote will reproduce its own clones. During reproduction, eukaryotes generate genetic variation by sexual recombination. Genetic variation mechanisms of prokaryotes are not tied to reproduction.

Prokaryotes are generally smaller than eukaryotes. The smallness has, among other things, consequences for growth rates and generation times. Diffusion limitation generally restricts the maximal size of prokaryotic cells. Because of the asexual reproduction and short generation time relative to larger organisms, prokaryotes pass the genome rapidly to subsequent generations. Therefore genetically changed genomes are also rapidly transferred, and therefore prokaryotes swiftly adapt and colonize new niches and a wide range of habitats.

Bacterial cell growth involves a *co-ordinated* increase in the mass of its constituent parts, and usually leads to the division of a cell into two similar or identical cells.
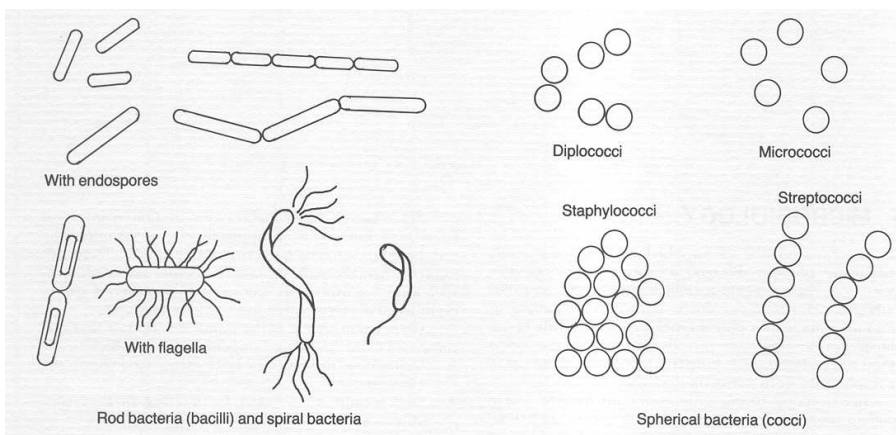

Figure 1.1: Schematic representation of various possible shapes of bacteria (not drawn in scale).

Bacteria grow only if their environment is suitable; if it is not optimal, growth may occur at a lower rate or not at all, or in the worst case, they may die. Essential requirements for growth include a supply of suitable nutriments, a source of energy, water and appropriate temperature, pH value and level of oxygen.

In a growing cell there is a co-ordinated increase in the mass of the component parts, where co-ordinated means that not all parts are made simultaneously: some are synthesized more or less continually, but others are made in a definite sequence during certain fixed periods. The cycle of events in which a cell grows and divides into two daughter cells, is called the *cell cycle*.

Events that belong to this phase are: i) growth of cell's dimension; ii) synthesis of the cell envelope; iii) chromosome (DNA) replication; iv) septum formation (the growing cell divides into two cells by the development of a cross-wall, which is the septum).

All these represent the asexual reproduction of bacteria, also called binary fission (vertical gene transfer). In one bacterium, the single circular chromosome duplicates. Then, the two resulting chromosomes attach to the inside of the plasma membrane. The cell elongates and separates into two strands. Finally, the cell membrane grows inward, the cell wall forms separating two daughter cells each with a chromosome.

In binary fission, the genetic information is copied almost identically to cell daughter. Mutations, and therefore evolution, are introduced randomly into the genetic material. Other ways by which bacteria can blend their genes exhist, in which new DNA elements may enter in a bacterium (*horizontal gene transfer*):

Transformation

In transformation, a bacterium takes up from its environment a piece of DNA; one strand of this *donor* DNA is internalized and may genetically transform the recipient cell by recombining with a homologous region of the chromosome (see Fig. 1.2). Transformation occurs naturally in various Gram-positive and Gram-negative bacteria, and **it may contribute to the spread of antibiotic resistance in pathogenic bacteria**.

Conjugation

Certain (conjugative) plasmids, and conjugative transposons (mobile genetic elements which can move from one DNA duplex to another) confer on their host cells the ability to transfer DNA to other cells by *conjugation*; in this process, a *donor* (male) cell transfers DNA to a *recipient* (female) cell, while the cells are in physical contact (see Fig. 1.2).

Like transformation, conjugation can result in important phenotypic changes in the recipient cell, like **the acquisition of plasmid-encoded resistance to antibiotics**.

Transduction

The transfer of DNA from one bacterial cell to another, via a bacteriophages vector (a virus), is called *transduction* (see Fig. 1.2).
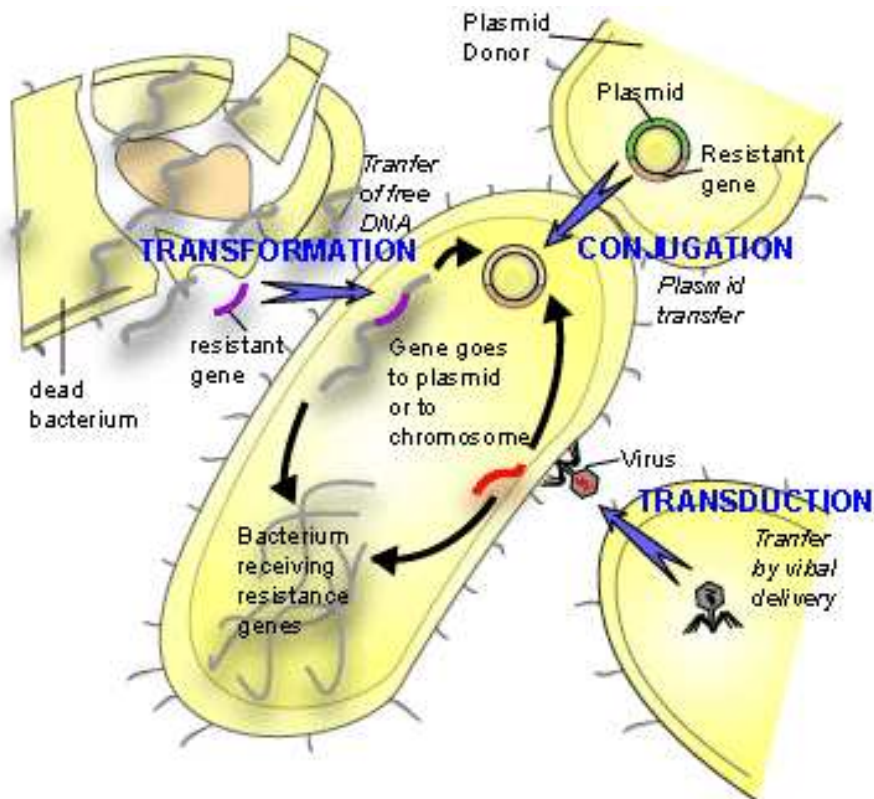
Figure 1.2: Horizontal gene transfer representation.

1.2 Cell wall: the main target of antibiotics.

The typical cell envelope of a bacteria is made by two components, the cytoplasmic membrane and, around it, the cell wall (Fig. 1.3).
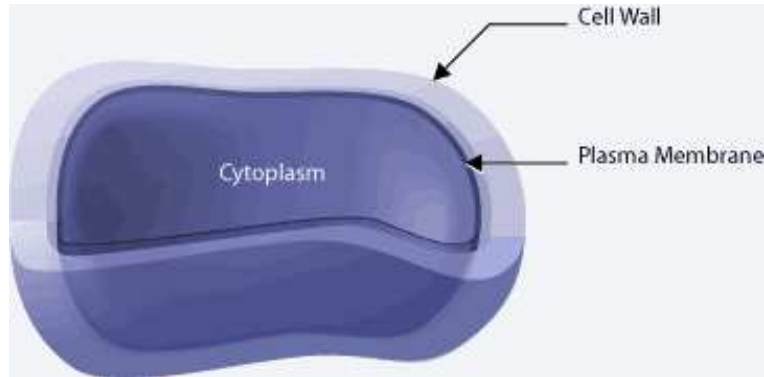
Figure 1.3: Schematic organization of a typical bacterial cell

If a *prokaryotic* cell loses its cell wall, the result is a structure called *protoplast*, and it can survive in a medium with the same salt concentration of the cytosolic medium, and even carry out many of the normal processes of a normal living cell. However, when immersed in a medium more diluted than the cytosol, water starts to enter into *protoplast*, which starts to swell until it burst. Such phenomena is called *osmotic lysis*. So, in an intact bacterium, the osmotic lysis is avoided by mechanical strength of the cell wall.

The cell wall also prevents the protoplast by mechanical damage linked to movement of bacterium in its environment, and it is responsible for its shape.

It also acts as a permeability barrier, regulating the transport of ions and molecule (among which antibiotics) to move across the membrane.

Every species posses a cell wall that differs from the one of other species in thickness, structure and composition. It is possible, however, to divide bacteria in two major groups, according to the reaction of the cell to treatment of certain dyes: one kind of cell wall, after being stained with, for example, crystal violet or iodine, retains the colour even after wash of ethanol or aceton, the other kind becomes decolorized. Based on this empirical procedure, bacteria that retain dyes are called *Gram positive*, in honour of the scientist Christian Gram, who discovered this procedure in 1880s, while the remaining ones are called *Gram negative* (Tab. 1.4).

The Gram-positive-type wall is made by a thick (from 30 to 100 nm) complex polymer called *peptidoglycan* (*murein*), while the Gram-negative-type (20-30 nm thick) posseses, in addition, another outer membrane (Fig. 1.4).

Gram-positive	Gram-negative
Bacillus	*Aeromonas*
Clostridium	*Bacteroides*
Lactobacillus	*Bordetella*
Listeria	*Brucella*
Propionibacterium	*Escherichia*
Staphylococcus	*Haemophilus*
Streptococcus	*Pseudomonas*
Streptomyces	*Salmonella*
Streptoverticillium	*Thiobacillus*

Table 1.1: Major *genera* of Gram-positive and Gram-negative bacteria.
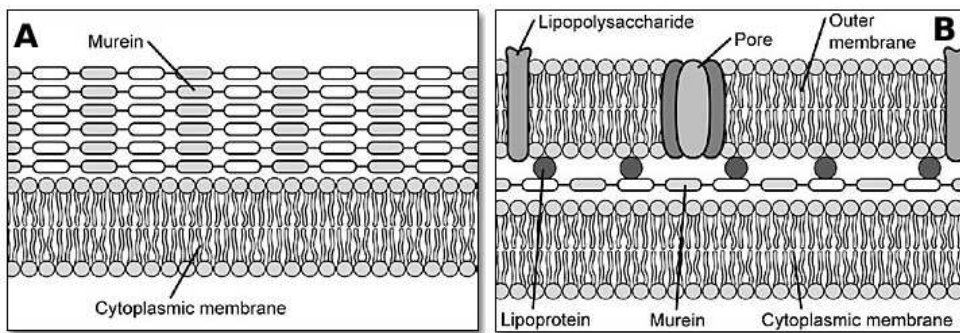
Figure 1.4: In Gram-positive bacteria (**A**), the cell wall is composed of a thick layer of murein, while in Gram-negative bacteria (**B**) the murein layer is thinner but it is surrounded by a second, outer lipid bilayer membrane.

The peptidoglycan sacculus is a unique and essential structural element in the cell of most bacteria, forming a closed, bag-shaped structure surrounding the cytoplasmic membrane.

The peptidoglycan scaffold of the bacterial cell wall is a repeating chain of two aminosugars: N-acetylglucosamine (NAG) -N-acetylmuramic disaccharide (NAM). A pentapeptide is attached to the D-lactyl moiety of each NAM (Fig. 1.5), and there is a high diversity in the composition and sequence of the peptides from different species, but it is worth to mention that in this biochemical structure one could find one of the few examples by which nature exploits the D-chirality of aminoacids, probably to escape the action of proteases degrading enzymes.
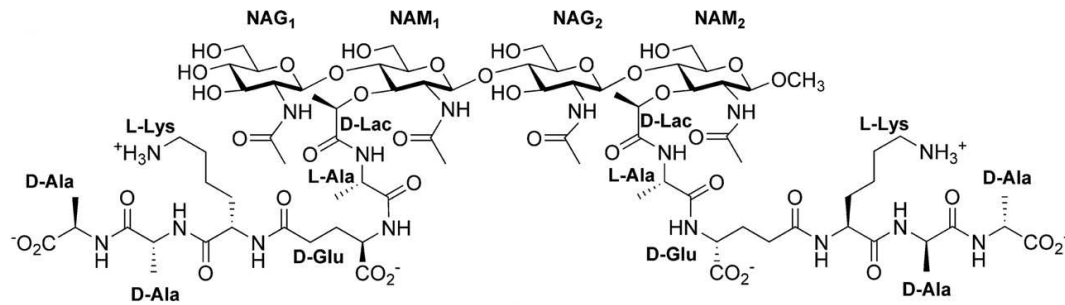
Figure 1.5: Chemical structure of the two building blocks of the peptidoglycan chain, namely N-acetylglucosamine (NAG) and N-acetylmuramic disaccharide (NAM), and the peptidic stem attached to NAM.

Each peptidoglycan monomer is synthesized in the cytosol of the bacterium, and then attached to a membrane carrier molecule, called bactoprenol, which brings the single monomers across the membrane and helps a number of other protein to insert the new monomers into the existing peptidoglycan chain, thus enabling regeneration of murein, and the growth of bacterium following binary fission.

The long sugar chains are then joined to one another by a peptide cross-link in a net like structure with the required high strength to counterbalance the internal osmotic pressure of the bacterium.

Murein posseses some striking biophysical properties: from one side it is a powerfull structure able to withstand the internal turgor of bacteria up to 25 atm, but from the other side the sacculi are flexible structures, allowing reversible expansion under pressure and possesing wide pores that enable diffusion of large molecules such as proteins.

1.3   Penicillin-binding proteins (PBP)

The polymerization of glycan chains (transglycosylation) and cross-linking of the the glycan chains (transpeptidation) are catalyzed by the penicillin-binding proteins

(PBPs) [31]

Bacteria posses a variable number of PBPs, which are divided into two main categories, the high molecular mass and the low molecular mass PBPs.

PBPs share a common DD-peptidase activity (whether a DD-transpeptidase, a DD-carboxypeptidase or a DD-endopeptidase activity) and are localized on the outside of the cell membrane where peptidoglycan shynthesis takes place [32]. The carboxypeptidation and transpeptidation reactions catalyzed by PBPs follow a three steps mechanism: (i) the rapid, reversible formation of a non covalent Henri-Michaelis complex between the enzyme and a peptidoglycan stem pentapeptide (L-Ala-γ-DGlu-A$_2$pm (or L-Lys)-D-Ala-D-Ala), called the donor strand; (ii) subsequent nucleophilic attack of the active serine on the carbonyl carbon atom of the C-terminal D-Ala-D-Ala peptide bond, leading to the formation of an acyl-enzyme intermediate and the concomitant release of the C-terminal D-Ala (acylation); (iii) the final step (deacylation) consists of either hydrolysis, with release of the shortened peptide (carboxypeptidation), or cross-link formation with a second peptidoglycan stem peptide called the acceptor strand (transpeptidation). The DD-endopeptidase activity of PBPs consists of the hydrolysis of the cross bridge resulting from the DD-transpeptidase activity [33]. The DD-peptidase activity of PBPs is catalyzed by a common PB domain which binds β-lactam antibiotics. These drugs are aimed to competitively inhibit the enzymatic DD-peptidase activity of this domain by playing on the structural similarity between penicillin and the D-Ala-D-Ala dipeptide [34]. This ending the natural substrate of PBPs, the pentapeptide precursor of the peptidoglycan.
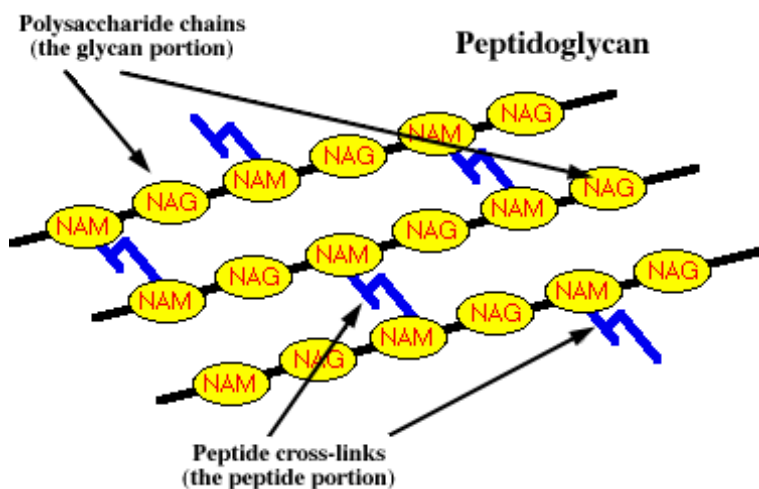
Figure 1.6: The peptidoglycan net.

1.4 β-lactam antibiotics

The common chemical structure among all antibiotics belonging to the so-called class of β-lactams, is a 4 member cyclic amide, resulting from the fusion of a lactone and an amide (Fig. 1.8).

β-lactam antibiotics comprise six different structural subtypes: penams, cephems, monobactam, clavams, penems and barbapenem.

The penams include benzylpenicillin and ampicillin (Fig. 1.9).

The cephems include classical cephalosporins such as cephaloridine, nitrocefin and cefotaxime, as well as cephamycins, which are 7-α-methoxy-cephalosporins (Fig. 1.10). The monobactams are monocyclic β-lactams and include aztreonam.

1.4.1 Carbapenem

The main subject of this work is a selective carbapenemase enzyme, and therefore an introduction to this class of antibiotics is required. Latest generation of β-lactam antibiotics are carbapenems, which have shown the broadest spectrum of activity within this antibiotic class. They exhibit *in vitro* bactericidal activity
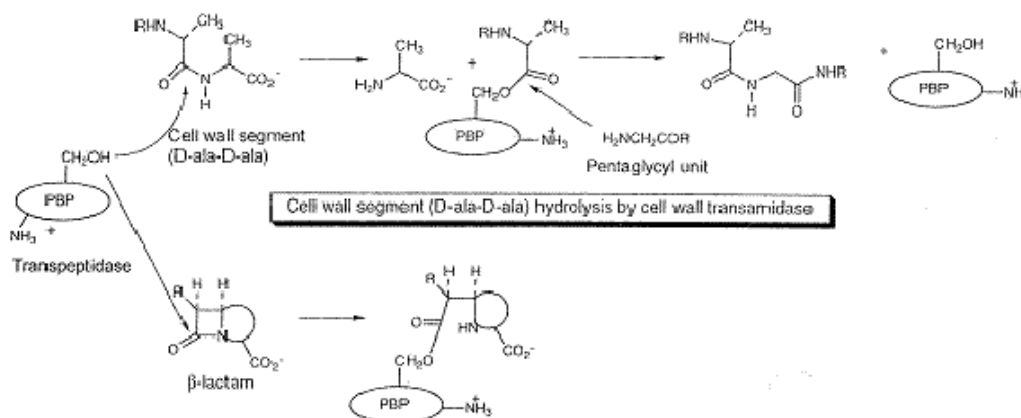
Figure 1.7: The reaction of transpeptidation catalyzed by PBP is represented on the top part of the figure, while competitive inhibition by action of β-lactam antibiotics is depicted on the bottom part.
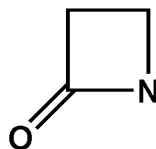
Figure 1.8: The β-lactam ring.
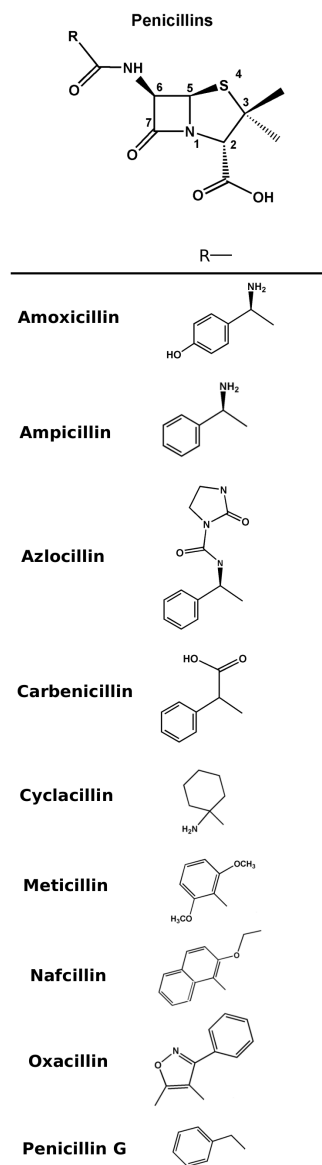
Penicillins

Figure 1.9: Some structures of most used penicillin antibiotics.
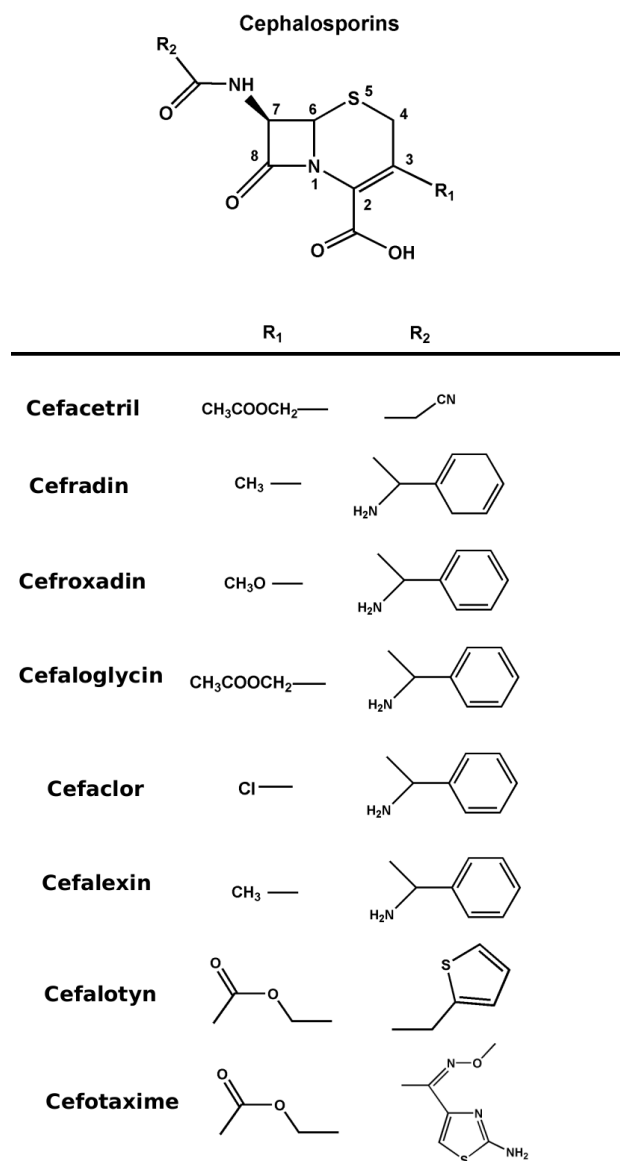
Figure 1.10: Some structures of most used cephalosporins.

against numerous pathogens, including Gram-positive and Gram-negative aerobes and anaerobes (Fig. 1.11). In addition, carbapenems are stable to almost all β-lactamases [16, 17, 18].
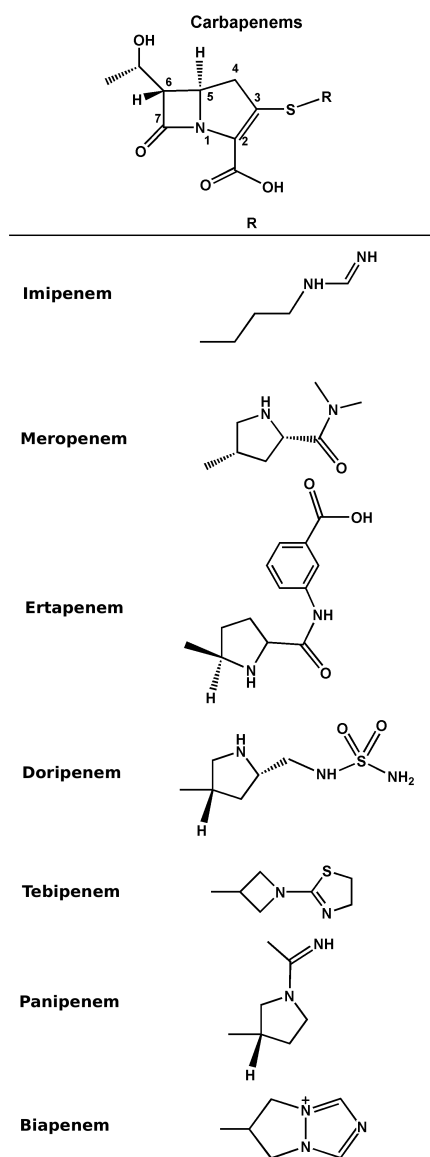
Figure 1.11: Carbapenem structures, in cronological order of discovery.

The carbapenems differ from penicillins (penams) in the presence of a carbon atom replacing the sulphur at position 1, and an unsaturated bond between C2 and C3 in the five membered ring structure.

The broad spectrum activity of carbapenems is associated with their intrinsic resistance to nearly all β-lactamases, which is due to the trans-α-1-hydroxyethyl subtituent at the 6 position of carbapenems. This is unique when compared with the side shains of penicillins and cephalosporins, which have *cis* configurations.

Chapter 2

Metallo-β-lactamase (MβL) enzymes

2.1 MβL: the zinc dependent lactamases

β-lactamases are the major cause of resistance of bacteria to β-lactam antibiotics. These enzymes cleave the amide bond of the β-lactam ring, thus inactivating the antibiotic. These enzyme fall in 4 classes: class A, C and D are serine-β-lactamases which employ an active-site serine to catalyze hydrolysis, while metallo-β-lactamases, or class B, are enzymes requiring one or two zinc ions for their activity. Metallo-β-lactamases can degrade all classes of β-lactams, except monobactams and are special for their constant and efficient carbapenemase activity. This is a most worrisome characteristic because carbapenems, which are stable against the vast majority of serine-β-lactamases produced by resistant pathogens, are the antibiotics with the broadest spectrum of activity. Moreover, metallo-β-lactamases are not susceptible to therapeutic β-lactamase inhibitors. The first of these enzymes, the *Bacillus cereus* metallo-β-lactamase, was identified in 1996 when Sabbath and Abraham showed that cephalosporinase activity produced by a *B. cereus* was inhibited by EDTA [35]. During the two following decades it was the only known example of metallo-β-lactamase, and was merely considered as a biochemical curiosity. The situation changed after 1980. Indeed, zinc β-lactamases have been discovered in an increasing number of nosocomial strains such as *Bactroides fragilis* [36], *Pseudomonas aeruginosa* [37], *Aeromonas hydrophila* [38], *Serratia marcescens* [39] and *Elizabethkinggia meningoseptica* [40]. The emergence and dissemination of acquired metallo-β-lactamases, encoded by genes carried on mobile DNA elements, among major Gram-negative pathogens, made the situation even more worrying [11]. Moreover, many metallo-β-lactamase genes are present in environmental species with thus constitute reservoirs of β-lactam resistance genes [41, 42, 43, 44].

The class B β-lactamases are classified into three subclasses B1, B2 and B3 on the basis of sequence alignments [45]. The classification task was complicated by

the generally low degree of similarity between subclass sequence, but facilitated by the availability of X-ray structures (Fig. 2.2), this allowed the identification of corresponding secondary structures elements, even when the sequence similarity was not obvious. The standard numbering of class B β-lactamases (BBL) allows easy comparison of the sequence of these enzymes [45]. Subclass B1 enzymes share more than 23% sequence identity. These enzymes includes the prototypical BcII from *B. cereus* [46], BlaB from *E. meningoseptica* [40] and EBR-1 from *Empedobacter brevis* [47]. The acquired IMP-type MβLs found in some clinical isolates of *P. aeruginosa* [48], *S. marcescens* [39], *Klebsiella pneumoniase* and *Acinetobacter baumannii* [17], as the VIM-type or SPM-1 enzymes produced by different species of *Aeromonas* (the most studied are CphA produced by *A. hydrophila* [38] and Imis produced by *A. veronii* [49]) as well as the Sfh-I enzyme produced by *Serratia fonticola* [42].

Subclass B3 metallo-β-lactamases has only nine conserved residues when compared with the other metallo-β-lactamases. Subclass B3 inlcude the L1 and GOB-1 metallo-β-lactamases produced by clinical strains of *S. maltophilia* [50] and *E. meningoseptica* and FEZ-1 isolated from *Legionella gormanii* [51]. To date, 18 variants of GOB-1 have been reported, but none outside the *E. menincoseptica* species. In the GOB enzymes, His116 is replaced by a glutamine. With the exception of L1 which is a homo-tetramer [15], all these enzymes are monomeric as the B1 and B2 enzymes. B1 and B3 members are broad spectrum enzymes and hydrolyze most β-lactam antibiotics including carbapenems [52], only the monobactams are neither hydrolyzed nor recognized by these enzymes, which moreover are not inhibited by classical inhibitors of serine-active β-lactamases such as clavulanic acid and tazobactam. These compounds, in fact, behave as poor substrate.

Antibiotic	*Aeromonas hydrophila* A2 enzyme			*Pseudomonas maltophilia* ULA-511 L-1 enzyme			*Bacillus cereus* β-lactamase II			*Bacteroides fragilis* β-lactamase ccrA		
	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (M$^{-1}\cdot$s^{-1})	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (M$^{-1}\cdot$s^{-1})	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/k_{cat} (M$^{-1}\cdot$s^{-1})	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (M$^{-1}\cdot$s^{-1})
Ampicillin	> 2000	> 500	$2.70 \times 10^{5*}$	40 ± 3	175	4.40×10^6	1530 ± 115	1105	7.20×10^5	120	190	1.60×10^6
Carbenicillin	450 ± 30	16.00	3.60×10^4	194 ± 15	277	1.40×10^6	4200 ± 300	755	1.80×10^5	240	190	8.00×10^5
Penicillin G	700 ± 50	5.21	7.40×10^3	50 ± 6	1100	2.20×10^7	1500 ± 200	678	4.50×10^5	40	193	4.80×10^6
Oxacillin	25 ± 3	0.75	3.00×10^4	27 ± 3	285	1.10×10^7	1750 ± 20	325	1.90×10^5	–	–	–
Cephaloridin	190 ± 8	0.12	6.30×10^2	300 ± 50	28	9.30 ± 10^4	1350 ± 140	25	1.90×10^4	6	42	7.00×10^6
Cephalexin	230 ± 20	0.65	2.90×10^2	100 ± 8	37	3.70×10^5	300 ± 26	1.7	5.70×10^3	49	95	1.90×10^6
Cephaloglycin	280 ± 30	0.20	7.10×10^2	36 ± 2	19	5.30×10^5	190 ± 5	61	3.20×10^5	–	–	–
Nitrocefin	100 ± 8	0.31	3.10×10^2	7 ± 1	20	2.90×10^6	70 ± 5	45	6.40×10^5	16	200	1.25×10^7
Cefuroxime	20 ± 3	0.07	3.50×10^3	30 ± 3	80	2.70×10^6	45 ± 5	35	7.80×10^5	4	29	7.30×10^6
Cefotaxime	26 ± 3	0.07	2.70×10^3	26 ± 3	66	2.60×10^6	90 ± 10	60	6.70×10^5	27	98	3.60×10^6
Imipenem	80 ± 8	168.00	2.10×10^6	90 ± 8	65	7.30×10^5	> 1000	> 100	$1.20 \times 10^{5*}$	270	200	7.40×10^5
6-β-IP	530 ± 50	0.64	1.20×10^3	640 ± 30	648	1.00×10^6	7460 ± 590	450	6.00×10^4	–	–	–
Sulbactam	37 ± 4	0.12	3.24×10^3	76 ± 4	210	2.80×10^6	5200 ± 325	10	1.90×10^3	–	–	–
Cefoxitin	Inact.	Inact.	Inact.	2 ± 0.05	1.10	5.50×10^5	2100 ± 465	0.2	9.50×10^1	110	10	9.00×10^4
Moxalactam	Inact.	Inact.	Inact.	1 ± 0.05	0.29	2.90×10^5	644 ± 90	1.2	1.90×10^3	160	(30)	$(1.90 \times 10^5)\ddagger$
Aztreonam	N.I.†	N.I.†	N.I.†	N.I.†	N.I.†	N.I.†	N.I.†	N.I.†	N.I.†	N.H.	N.H.	N.H.

* Only the k_{cat}/K_m ratio was determined because v_0 remained proportional to S_0 up to the highest substrate concentration which could be used.
† Up to 0.5 mM aztreonam.
‡ Cuchural et al. (1986). Values in parentheses were computed by us from their data.

Table 2.1: Kinetic parameters determined for MβLs [52].

The subclass B2 enzymes are high selective carbapenemases [52, 53, 49]. They only hydrolyze carbapenems efficiently and show a very weak activity, if any, towards penicillins and cephalosporins. In B2 enzymes, the conserved His116 of B1 and B3 enzymes, is replaced by an asparagine. The presence of this asparagine in position 116 is far from being the only reason which explains the narrow spectrum of B2 metallo-β-lactamases, since, although increased, the activity of the H116N CphA mutant towards penicillins and cephalosporins remains low [54].

A mutant of CphA with a considerably broadened activity spectrum, N116H-N220G was recently obtained by site-directed mutagenesis [55]. In contrast to the wild-type enzyme, this mutant is able to efficiently hydrolyze penicillins and cephalosporins in addition to carbapenems, although with a reduced efficiency towards the latter.

Subclass B1 and B2 are descended from a common ancestor, and really form a single group within which the sequences exhibit significant similarities. Subclass B3 shares structural but not sequence similatrities with B1/B2 groups [56]. A phylogenetic tree with representative members of all the three subclasses is given, to revise the classification of metallo-β-lactamases [57].

The three dimensional structure of the mono-zinc form of BcII was the first to be solved [58]. Thereafter, the structures of other subclass B1 enzymes, CcrA [59], the di-zinc form of BcII [60], IMP-1 [61], VIM-2, SPM-1 [62] and subclass B3 enzymes, L1 [15], FEZ-1 [63] and BJP-1 were determined. More recently, the structure of a subclass B2 enzyme, CphA, was solved [14]. Despite of the low level of identity between the amino acid sequences, the general fold of these enzymes is similar and consists of an $\alpha\beta\beta\alpha$ structure, composed by two central β-sheets with five solvent exposed α-helices. The N-terminal and C-terminal parts of the molecule, each of them comprising a β-strand and two α helices, can be superposed by a 180 deg rotation around a central axis, suggesting that the complete structure might have arisen from the duplication of a gene [60]. In all known structures, the active site is located at the external edge of the $\beta\beta$ sandwich. The N-terminal domain of B1 metallo-β-lactamases incorporates a loop (61-65) that can interact with substrate or inhibitor molecules which possess hydrophobic side-chains. This loop is very flexible in the native form of the enzyme. When the substrate or the inhibitor diffuses into the active site, the loop moves to block the molecule (the side chain of the W64 residue interacts with the hydrophobic side chain of the substrate [64]). The loop is also stabilized upon binding of inhibitors [59, 63] and transforms the active site groove into a tunnel-shaped cavity. The deletion of this loop (named flap) seriously affects the enzymatic activity, with the only exception of thje activity towards imipenem. Actually, the kinetic parameters of this carbapenem compound are barely affected by the absence of the active site loop. The main difference between the carbapenems relies on the fact that they usually have bulky aromatic ring substitutions which are absent in the former.

The 61-65 flap is absent in β-lactamases of subclasses B2 and B3.

2.2 CphA: the first crystal structure of a mononuclear B2 MβL

2.2.1 The free enzyme

Garau and coworkers [14] solved the first crystallographic structure of the enzyme CphA, from *Aeromonas hydrophila* in 2004.

This enzyme consists of 227 amino acid residues with a calculated molecular mass of 25.2 kDa. The model of the structure includes 224 amino acid residues (41-304, using the standard class B β-lactamase numbering [45]), the catalytically essential zinc ion, one carbonate anion located in the active site, one sulphate ion, and 178 water molecules.

CphA is a monomeric globular protein having the general $\alpha\beta\beta\alpha$ fold as described in Fig. 2.1
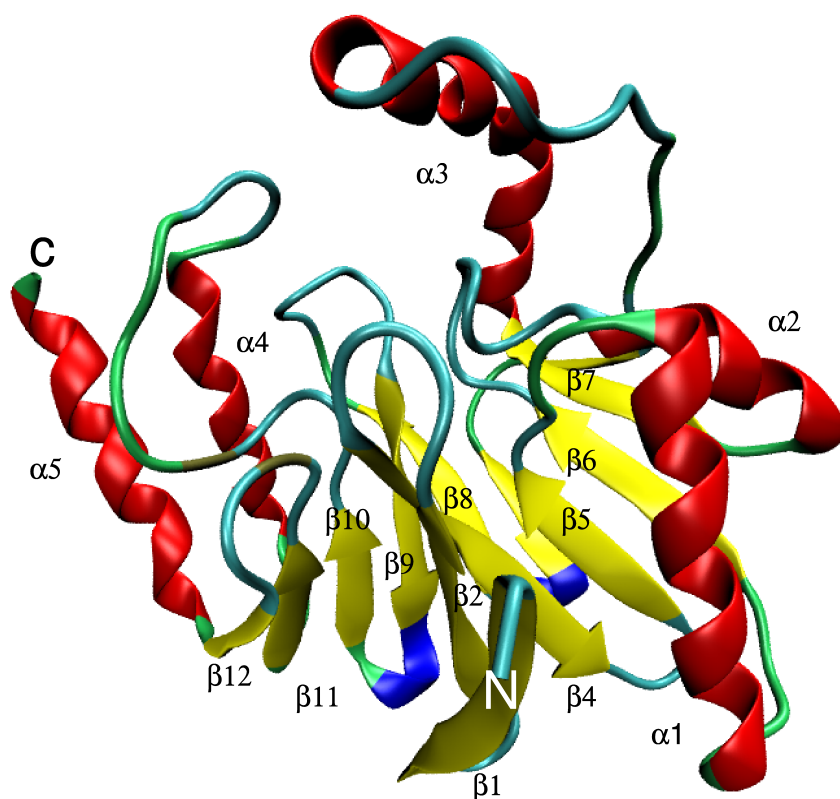

Figure 2.1: X-ray secondary structure of CphA from *Aeromonas hydrophila*.

Four β-strands (β1, β2, β3 and β4), two helix-strand elements (α1β5 and α2β6), one long α-helix (α3, and the β7 strand form the N-doman, while four β-

strands (β8, β9, β10 and β11), one 3_{10} helix, the α4 helix, the β12 strand, and the α5 helix form the C-domain. The N and C-domains face each other through sheets β1-7 and β8-12. The α1 and α2 helices of the N-domain are located externally on opposite faces of the β-strand core. In contrast to all reported B1 and B3 structures, a long α3 helix, formed by residues Arg140-Leu161, has a kink, which allows the helix to follow the curvature of the protein. The helix is followed immediately by an unusual proline-rich loop (Pro162, Pro165, Pro168 and Pro172) (Fig. 2.2).
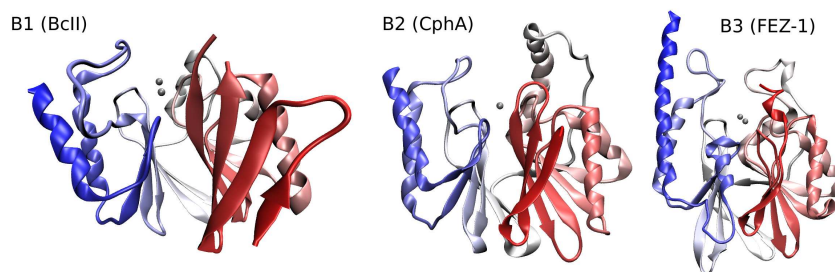
Figure 2.2: Secondary structure of members of each MβL subclass.

The ψ, ϕ angles of Tyr 60, Thr86, Ala195, Asn220 and Asp264 are located in the disallowed areas of the Ramachandran plot. The last three are adjacent to an amino acid involved in one of the zinc-binding sites ("Zn1" or "Zn2") and, in most other sequences, are glycine residues. Tyr60 is on a loop connecting two β-strands (β2 and β3), which forces it to adopt a strained conformation ($\psi = 58$ deg; $\phi = -123$ deg). Thr86 is on a loop connecting β4 to α1, and is adjacent to Trp87. A group of hydrophobic side-chains (Tyr59, Val67, Trp87, Tyr164, Pro165 and Phe236), together with other side-chains located on the α3 helix (Ile153, Phe156 and Leu161), appears to form a "hydrophobic wall", which defines the active site. Ala196 ($\psi = -160$ deg; $\phi = -160$ deg) is located on a loop between strands β8 and β9, which, at its apex bears His196, which is involved in the active-site hydrogen bond network. Asn220 is the first residue of the 3_{10} helix and precedes the zinc-coordinating residue, Cys221, while Asp264 follows another zinc-coordinating residue, His263.

Secondary structure elements bear two of the three residues involved in zinc coordination, Asp120 in the α2 helix and Cys221 in the 3_{10} helix. The third residue, His263, is located on the loop connecting β12 and β13. The distances betwen the Zn atom and the Asp120 carboxyl oxygen atom, the Cys221 sulfur atom, and His263 side-chain nitrogen atom are 1.96Å, 2.20Å and 2.05Å, respectively. Finally, a carbonate ion is bonded to the zinc (2.09 Å), and the four coordinating atoms, from Asp120, Cys221, His263, and O1 of CO_3^{2-}, generate a tetrahedral geometry around the metal ion.

2.2.2 The CphA-biapenem complex
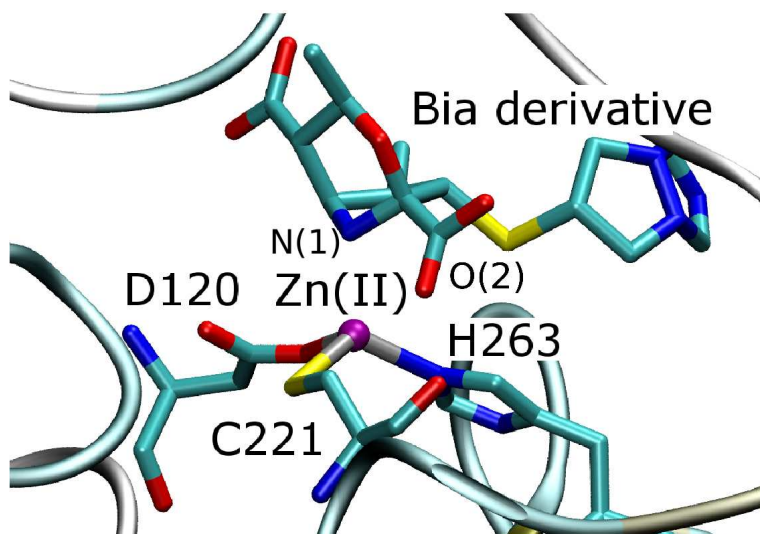
Figure 2.3: Close up of the X-ray structure of CphA from *Aeromonas hydrophila*, in complex with biapenem derivative.

The importance of experimental data of Garau *et al.* relies not only to have solved for the first time a B2 MβL structure, but to succeed also in obtaining a structure of the complex with an antibiotic (Fig. 2.3). The refined CphA-biapenem complex structure comprises 227 protein residues, one zinc ion, six sulphate anions, and the biapenem molecule located in the groove of the active site with full occupancy. In [14] the author called this modified biapenem "intermediate" of the reaction, but it is still not clear if this is in fact an intermediate, or a modification of the reaction products that back-acts as inhibitors in the catalytic pocket. In this work I will refer to it as "biapenem derivative" (Fig. and Fig. 2.4).

The biapenem derivative is bonded to the zinc ion located at the WT position. The distance from the nitrogen atom of the β-lactam ring (N4) to the Zn^{2+} is 2.22 Å. N4 also forms a hydrogen bond with Wat1 (2.9 Å), which in turn, forms hydrogen bonds with Asp120 and His118. The distance between Wat1 and the zinc atom is 3.39 Å. The β-lactam ring, as emerged from electron density map, has been cleaved between N4 and C7. The C3 carboxyl group forms strong hydrogen bonds wih the Lys224 side-chain nitrogen atom and the Asn233 backbone nitrogen atom. The C7 carboxyl group formed after antibiotic hydrolysis makes two hydrogen bonds with the two hydroxyl groups of Thr119 and Thr157. C2 and C3 of the biapenem derivative exhibit sp^3 hybridization, showing that the molecule in the active site has lost the C2=C3 double bond. The side-chain containing the sulphur atom and the C3 carboxyl group are in *cis* configuration. The biapenem is hydrolyzed and has lost the double bond, and it has an internal molecular rearrangement. The

biapenem derivative in the active site is stabilized by several hydrophobic contacts. The biapenem sulfur atom interacts with the His263 ring, inducing a small rotation of its plane; the 1β-mehyl group is in van der Waals contact with the side-chain of Trp87 and Val67, and the methyl group of the C6 1-hydroxyethyl forms hydrophobic contacts with Phe156 and Phe236 and the Asn233 sidechains carbon atoms.

2.2.3 Zinc coordination

Metallo-β-lactamases posses two potential Zn^{2+} binding sites (Fig. 2.5).
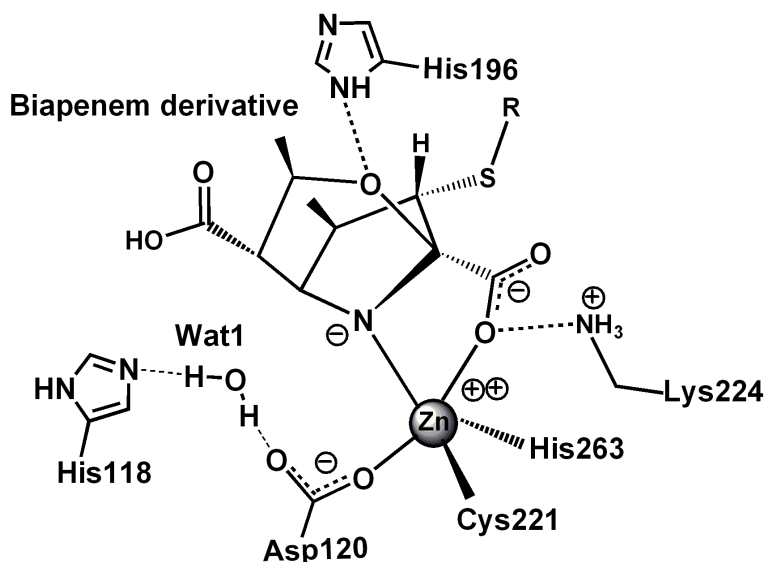
Figure 2.4: Schematic representation of the active site of CphA in complex with biapenem derivative (PDB 1X8I [14]).
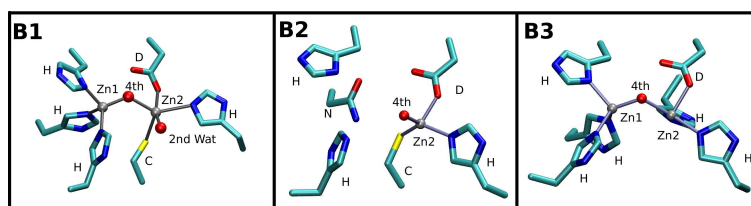
Figure 2.5: "Zn1" and "Zn2" binding sites in MβLs, as observed in the B1 BCII enzyme from *Bacillus cereus* [13] (residues are cut at the Cα), the B2 CphA enzyme from *Aeromonas hydrophila* [14] and the B3 L1 enzyme from *Stenotrophomonas maltophilia* [15]. In the case of the B1 enzyme, a second solvent molecule binds to the metal ion. The backbone units of the residues are not shown for clarity.

In the case of B1 enzymes, one zinc ion possesses a tetrahedral coordination sphere and is coordinated by His116, His118, His196 and a water molecule or OH$^-$ ion. The other metal ion has a trigonal-pyramidal coordination sphere which invloves Asp120, Cys221, His263 and two water molecules. One water/hydroxide molecule servers as a ligand for both metal ions. The two binding sites are named "Zn1" and "Zn2" binding site, or with a letter code which indicates the amminoacidic composition of the site (so that one is called 3H site, while the second DCH site).

In B3 enzymes the 3H site is the same as that found in B1 enzymes. Cys221 is replaced by a serine and the second zinc ion is coordinated by Asp120, His121, His263 and the nucleophilic water molecule which forms a bridge between the two metal ions. Ser221 does not directly interact with the zinc ion but with a second water molecule located in the active site and which could serve as proton donor in the catalytic process [15]. In contrast to all other related metallo-β-lactamases, the subclass B3 GOB-18 enzyme is fully active against a broad range of β-lactam substrate as a mono-zinc species. Different spectroscopic techniques, 3D modelling and mutagenesis experiments, reveal that the zinc ion is bound to Asp120, His121, His263 and a solvent molecule, which means in the canonical "Zn2" site of B3 subclass [65].

While the B1 and B3 enzymes exhibit maximum activity as di-zinc species, the B2 β-lactamases are inhibited upon binding of a second Zn^{2+} ion (for CphA, the dissociation constant K_i is 46 μM at pH 6.5 [66]).

In agreement with EXAFS spectroscopy studies [67] and site directed mutagenesis [54], the crystallographic structure of CphA shows that the zinc ion is in the "Zn2" site, namely a DCH site [14]. Currently, the identification of the second binding site in subclass B2 enzymes remains unsettled since even in the presence of a large excess of zinc (up to 10 mM) and in spite of the low value of the dissociation constant, Garau and his coworkers did not succeed in obtaining crystals of the di-zinc form. On the basis of spectroscopic studies with the cobalt substituted ImiS enzyme, Crawford *et al.* postulate that in subclass B2 the second metal binding site is not the traditional "histidine" site [68]. Costello *et al.* propose that in ImiS the inhibitory zinc ion binds to both His118 and Met146 [69]. Indeed, their EXAFS data show that the inhibitory zinc is bound by a sulphur. Such sulphur in the inhibitory site should come from a methionine, because the only other sulphur-containing residue is Cys221, yet involved in the bind of Zn(II) in the catalytic site. Moreover, M146I mutation abolishes inhibition induced by zinc. However, mutations of other CphA residues, which have not been implicated as "Zn2" ligands can have similar effects (lower inhibition or even activation by excess of zinc) [54, 55].

2.2.4   Mono vs di nuclear enzymes

The subclass B2 CphA enzyme exists mainly in the mono-zinc form as the value of the dissociation constants are 200 nM and 46 μM, respectively for the binding of the first and the second zinc ion [66]. Subclass B1 metallo-β-lactamases show

distinct stereochemical similarities in their active sites, with identical residues contributing to metal coordination. However, the corresponding zinc affinities appear to be different, so that CcrA and IMP-1 show two high-affinity metal-binding sites, whereas BcII binds "Zn1" and "Zn2" with very different affinities. Fluorescence spectroscopy studies with a chromophoric chelator show that BcII has very different dissociation constant for the two metal-binding sites. For the loss of metal-ion from the mononuclear enzyme, the dissociation constant K$_{mono}$ is 6.2 x 10^{-10} M and that for the loss of one metal-ion from the dinuclear MβL, K$_{di}$ is 1.5 x 10^{-6} M [70]. K$_{mono}$ decreases significantly, from nanomolar to picomolar values, in the presence of substrate, whereas K$_{di}$ decreases only twofold [71]. This suggests that the mono-zinc enzyme is responsible for the catalytic activity under physiological conditions, where the concentration of free zinc ions is found by crystallographic methods in the 3H site, at pH 5.6, in presence of 100 μM zinc [58]. At this pH, the activity is only three- to fourfold lower than that at pH 7.5 [72]. But the existence of the mononuclear form, where the metal ion would be shared between the two binding sites is supported by several kinetic results, PAC (Cd) and spectroscopic (Co) data [70]. Conversely, and despite the very close similarity with BcII, the B1 enzyme CcrA binds both zinc ions very tightly. Kinetic studies of CcrA with nitrocefin have shown that only the di-nuclear species is active and that a previously observed mono-zinc CcrA enzyme [73] was a mixture of the di-zinc and the apo-enzyme [74]. BcII is active in the presence of either one or two zinc ions, although full activity is observed when two zinc ions are bound. It has been proposed that the arginine 121 present in BcII, Vim-2, BlaB and in subclass B2 enzymes is responsible for the lower affinities. The CcrA and IMP-1 enzymes, showing high affinity for both zinc ions, have cysteine and serine, respectively, in the corresponding position. The replacement of the positive Arg121 side chain in CphA by a neutral one enhances the affinity for Zn^{2+} [55], as was already suggested for BcII [58, 75]. Rasia and Vila [76] have shown that R121H mutation increases the affinity for the second Zn^{2+}. The C121R mutant of CcrA was isolated in the di-zinc form, although the removal of the second zinc appeared to be facilitated [74]. All these data indicate that the replacementt of Arg121 by a neutral residue (a Cys, Ser or His, respectively) in CcrA, IMP-1, and the subclass B3 enzymes increases the affinity for the second zinc. Until recently, no mono-zinc form of subclass B3 enzymes was observed. However, the subclass B3 GOB-18 was shown to be a mononuclear enzyme [65].

As shown by the data described above, the activity and even the existence of mono-zinc forms of B1 and B3 enzymes remain controversial. It seems impossibile to supply a clear answer to these issues and it is possible that minor sequence differences might result in different behaviours. For B2 enzymes, the mono-zinc form is clearly the active one, but here the location of the inhibitory second zinc remains unknown.

2.2.5   Catalytic mechanism of MβLs

It is supposed that the β-lactam hydrolysis occurs in two steps: the nucleophilic attack on the carbonyl group and C-N bond cleavage, this second step being generally triggered upon protonation of the nitrogen. The bicyclic shape of β-lactams requires that the two steps take place on the same face of the antibiotic [72].

The nature of the attacking nucleophile in the three subclasses is dependent on the nature of Zn ligand in the catalytic pocket. This is related to the open issue of the exact location of the Zn in the mononuclear form, but it seems most likely to be a "Zn1"-bound hydroxide in mono and dinuclear B1 and B3 MβLs.

2.2.6   Dinuclear MβLs

The best understood dinuclear MβLs are CcrA, IMP-1 and L1. Nitrocefin was used as substrate in most of the studies. This is a chromophoric cephalosporin that, although with no therapeutic utility, undergoes intense spectroscopic changes upon hydrolysis making it a useful mechanistic probe.

Stopped-flow kinetic studies of nitrocefin hydrolysis by L1 and CcrA revealed the accumulation of ring-opened, anioinc intermediate [77, 78]. This species has been observed in nitrocefin hydrolysis by a diZn(II) model complex [79] and its existence is supported by computational studies [26]. Taken together these data indicate that Zn2 stabilizes negative charge on the nitrogen atom after C-N bond cleavage. Rapid-freeze quench EPR studies on Co(II)-substituted L1 demonstrate that changes in the geometry of Co2 occur when this intermediate is formed and that similar changes come along with the other β-lactams [80]. However, the degree of accumulation of this intermediate varies markedly between enzymes and depends upon the protein environment. Nitrocefin differs from the majority of therapeutic β-lactams in possession of an extended π-conjugated system that contributes to the stabilization of an anioinc nitrogen. Studies of a wider range of substrates have led Spencer and coworkers to suggest that, in most cases, the amide bond is still intact in the populated enzyme-bound species [81]. Nitrogen protonation and C-N bond cleavage may not be distinguishable as two separate steps with other substrates (Fig. 2.6).

The general reaction mechanism for dizinc MβLs, emerging from all previous works (Fig. 2.6) could be described with the following events: after a rapid-equilibrium binding of substrate [81], the bridging hydroxide becomes terminal as the β-lactam carbonyl interacts with "Zn1" [82, 27].

The β-lactam carbonyl is polarized by the metal and also by other residues, and it is therefore susceptible to nucleophilic attack [26]. Deprotonated Asp120 orients the hydroxide to attack the β-lactam carbonyl carbon, generating a tetrahedral species [26, 27, 83, 84] that, in contrast to the serine catalyzed β-lactamases, does not seem to accumulate.

This species collapses into a nitrogen-anionic intermediate that yields to product upon protonation [85]. This intermediate was postulated to be the species de-
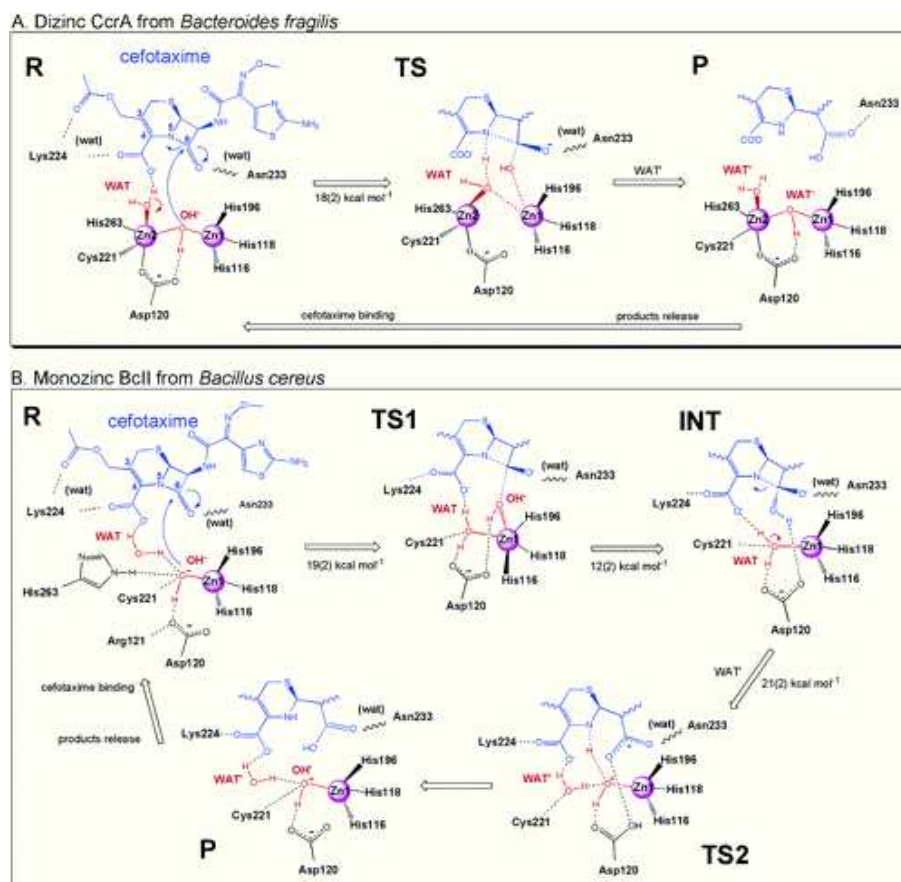
Figure 2.6: Cephalosporin hydrolysis catalyzed by (A) binuclear MβL CcrA from *B. fragilis* and (B) Mononuclear BcII from *B. cereus* [23].

tected in stopped-flow UV-Vis studies of nitrocefin hydrolysis [86]. Breakdown
of either the tetrahedral or the nitrogen-anionic intermediate may be rate-limiting,
depending on the substrate and the identity of the amino acids in the active site
[74, 81, 87, 88]. The crystal structure of the L1 in complex with hydrolyzed mox-
alactam complex supports a role for "Zn2" in stabilizing the transient development
of negative charge on the bridging nitrogen [89]. Asp120 and a water molecule
have been suggested as possible proton donors in the rate-determining step, al-
though this issue is still matter of debate. Mutagenesis studies of Asp120 reveal
that this residue is essential in CcrA, L1, IMP-1 and BcII [87, 90, 84, 91]

2.2.7   Mononuclear B1 MβLs

In the proposed mechanism for mono-Zn(II) BcII, the resting form of the enzyme
is assumed to be the tetrahedral "Zn1" site. The Zn(II)-bound water/OH- is hy-
drogen bonded to deprotonated Asp120, His263 and Cys221 [92, 93, 83]. Asp120
is oriented by Arg121 via two hydrogen bonds. Both kinetic studies of the pH
dependence of β-lactam hydrolysis [72] and crystallographic results [93] support
the hypothesis that the pK_a of the Zn(II)-bound water is 5.6. As in dinuclear
MβLs, this hydroxide serves as the reaction nucleophile, being oriented by Asp120
[94, 22]. During breakdown of the tetrahedral intermediate, proton transfer to the
β-lactam nitrogen is required. The source of this proton is not clear; however,
His263 [28, 24], Asp120 [72], and a water molecule [22] have been implicated as
potential proton donors. As recently predicted with L1 [84], Asp120 may play a
role in orienting water for proton donation. Mono-Zn(II) BcII is not thought to
stabilize anionic intermediates.

Mononuclear B2 MβLs: CphA

Based on structural information gained by X-ray structure of the complex CphA-
biapenem [14], Garau *et al.* proposed a specific reaction pathway, outlined in Fig.
2.7.

In the structure of the complex, the C3 carboxyl group makes a bond with the
zinc ion, and interacts with the NH_3^+ of Lys224, and the Asn233 nitrogen back-
bone. Thr119 and Thr157 form hydrogen bonds with, and Phe156 and Phe236
form hydrophobic interactions with, the hydroxyethyl group, which helps to stabi-
lize and orient the antibiotic molecule. Substrate binding promotes the polarization
of the β-lactam ring carbonyl oxygen atom by His196. The Asn116 side-chain ni-
trogen atom and Thr197 hydroxyl group form hydrogen bonds with the His196
sidechain nitrogen atom. This hydrogen bonding network allows His196 to par-
tially donate its hydrogen atom to the β-lactam ring carbonyl oxygen atom. A
water molecule, activated by His118, attacks the carbonyl carbon and cleaves the
β-lactam bond. As a consequence N4 of carbapenem bonds to the zinc. The Zn^{2+}
also helps in the correct orientation of the β-lactam ring during hydrolysis. The car-
boxyl group, generated by the hydrolytic process, is stabilized initially by His196,
Asp120 and His118, but after rotation around the C5-C6 bond, these interactions
are replaced by these observed in the complex between the carboxyl group and

Thr119, Thr157. The water molecule Wat1 is located in a position suitable for promoting the protonation of the lactam nitrogen atom, which will weaken the interactions with the Zn(II). At the last step of the catalytic reaction, the hydrolyzed carbapenem molecule can leave the active site.

Recent SCC-DFTB calculations [95] on the CphA enzyme were consistent with the postulated first step of reaction, identifying Asp120 as general base [96, 97]. The calculated free energy barrier for this step would be consistent with experimental data [14], if such step is the rate-limiting one, along the enzymatic cycle. This at present is not known, since the second step has not been yet investigated. In the same computational study, it was speculated that the β-lactam N(1) would accept a proton directly from Asp120 (**path II** in Fig. 5.1) with consequent breaking of the C-N bond.
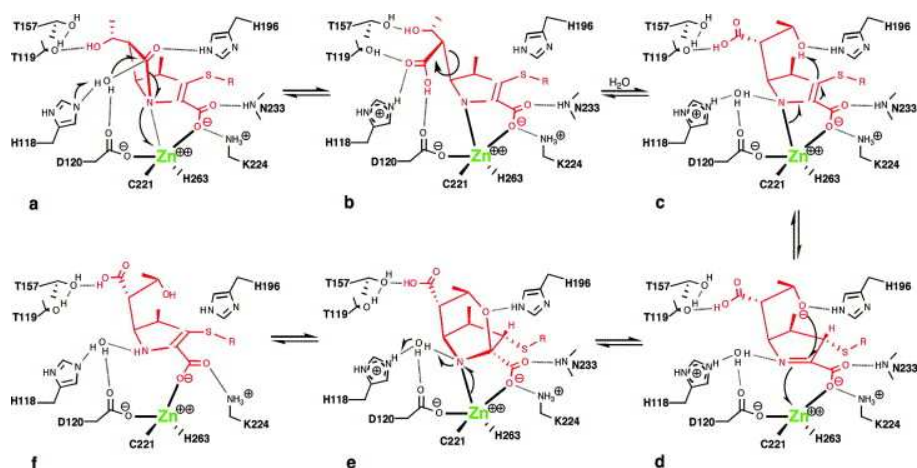
Figure 2.7: Reaction pathway proposed based on X-ray structural data of CphA in complex with biapenem derivative [14].

2.3 Enzyme Kinetics Theory

A quick introduction on the theory of enzymatic kinetics is given in order to explain the two experimental quantities k_{cat} and K_M with which we have compared our computational results.

We will always assume, in every derivation of kinetic equations, that the concentration of the substrate [S] is larger than [E], or that the latter is negligible with respect to the first. Such assumption is really close to reality, because of the high catalytic efficiency of the enzymes.

2.3.1   Michaelis-Menten mechanism

The Michaelis-Menten is the most simple mechanism that one could assume for an enzymatic reaction, and proceeds according to the following the scheme:

$$E + S \underset{\rightleftharpoons}{\overset{K_S}{}} ES \overset{k_{cat}}{\rightarrow} E + P \tag{2.1}$$

The above equation assumes that the substrate and enzyme combine rapidly and reversibly, via non-covalent interactions, to form the complex ES, also called *Michaelis-complex*. The chemical reaction occurs in a second step, with a first-order rate constant k_{cat} (the *turnover* number). From 2.1

$$\frac{[E][S]}{[ES]} = K_S \tag{2.2}$$

and

$$v = k_{cat}[ES] \tag{2.3}$$

We define the total enzyme concentration $[E]_0$, which is related to the free enzyme [E] by:

$$[E] = [E]_0 - [ES] \tag{2.4}$$

thus

$$[ES] = \frac{[E]_0[S]}{K_S + [S]} \tag{2.5}$$

and finally

$$v = \frac{k_{cat}[E]_0[S]}{K_S + [S]} \tag{2.6}$$

2.3.2   Experimental basis

It is found experimentally that there is a direct proportionality between v and $[E]_0$, that is, as we have already seen, the total enzyme concentration.

Usually it is observed that v follows a *saturation kinetics* with respect to [S], similar to the one plotted in 2.8.

All this is expressed by the experimental equation:

$$v = \frac{k_{cat}[E]_0[S]}{K_M + [S]} \tag{2.7}$$

The saturating limit towards which v tends as the substrate concentration increases, is $V_{max} = k_{cat}[E]_0$. We also define K_M to be the substrate at which $v = \frac{1}{2}V_{max}$, and finally we note that at very low [S], or at least at $[S] \ll K_M$, we have:

$$v = \frac{k_{cat}[E]_0[S]}{K_M} \tag{2.8}$$

.
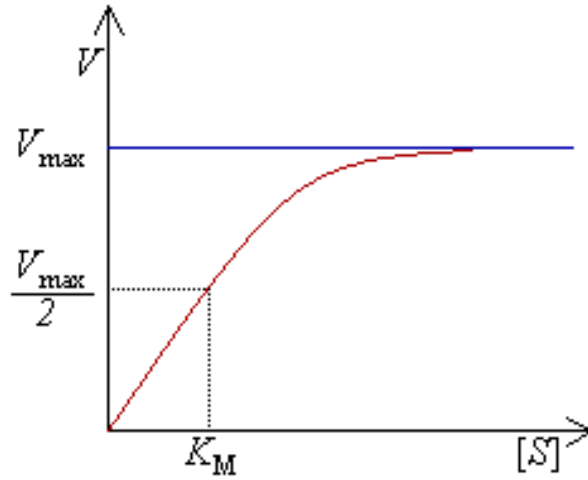
Figure 2.8: Reaction rate plotted against the substrate concentration

2.3.3 Extension of the Michaelis-Menten mechanism

The Michaelis-Menten mechanism derived before is just a particular case of the general mechanism that is valid when $k_2 \ll k_{-1}$:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_2}{\rightarrow} E + P \tag{2.9}$$

.

We solve this equation applying the steady state approximation to [ES]

$$\frac{d[ES]}{dt} = 0 = k_1[E][S] - k_{-1}[ES] - k_2[ES], \tag{2.10}$$

we remember that $[ES] = [E]_0 - [E]$,
which leads to

$$[ES] = \frac{[E]_0[S]}{\frac{k_2 + k_{-1}}{k_1} + [S]}, \tag{2.11}$$

and, since $v = k_2[ES]$

$$v = \frac{[E]_0[S]k_2}{\frac{k_2 + k_{-1}}{k_1} + [S]} \tag{2.12}$$

we call $K_M = \frac{k_2 + k_{-1}}{k_1}$, and remembering also that $K_S = \frac{k_{-1}}{K_1}$ we have

$$K_M = K_S + \frac{k_2}{k_1} \tag{2.13}$$

Now it is clear that in the case of $k_{-1} \gg k_2$ we have the simplification $K_M = K_S$

Chapter 3

Methods

The computational tools, used in this work to address the study of the important class of MβL enzymes, will be presented in this chapter.

3.1 A protein floating in water: the physical system

From the physical point of view, the system we have faced in this study (a protein immersed in a water medium), can be considered as a collection of interacting atoms (i.e. atomic nuclei surrounded by electronic clouds). Movements and reciprocal interactions can be described by quantum mechanics formalism: the system of M nuclei and N electrons is described by a wave-function $\Psi(\mathbf{R}, \mathbf{r}, t)$, which is a function of the 3M nuclear ($\mathbf{R}$), 3N electronic ($\mathbf{r}$) coordinates, and of time (for simplicity we will not consider nuclear and electronic spin degree of freedom, and spin coupling effects).

In a non-relativistic description, this wave function should solve the Schrödinger Equation:

$$i\hbar \frac{\partial}{\partial t} \Psi = \hat{\mathscr{H}} \Psi, \tag{3.1}$$

where $\hat{\mathscr{H}}$, the Hamiltonian, in the absence of external forces, is given by the general form:

$$
\begin{aligned}
\mathscr{H} \;=\; & -\sum_{I=1}^{M} \frac{\hbar^2}{2M_I} \nabla_I^2 - \sum_{i=1}^{N} \frac{\hbar^2}{2m} \nabla_i^2 + \frac{e^2}{2} \sum_{I=1}^{M} \sum_{J \neq I}^{M} \frac{Z_I Z_J}{|R_I - R_J|} \\
& + \frac{e^2}{2} \sum_{i=1}^{N} \sum_{j \neq i}^{N} \frac{1}{|r_i - r_j|} - \frac{e^2}{2} \sum_{I=1}^{P} \sum_{i=1}^{N} \frac{Z_I}{|R_I - r_i|},
\end{aligned}
\tag{3.2}
$$

where Z and M are the nuclear charges and masses, m is the electronic mass.

Temporal and spatial coordinates could be separated, and 3.1 could be simplified into a time-independent formulation, which is an eigenvalue equation:

$$\hat{\mathscr{H}}\Psi_n(\mathbf{R},\mathbf{r}) = E_n\Psi_n(\mathbf{R},\mathbf{r}). \tag{3.3}$$

In practice, only few trivial cases, like hydrogenoid atoms or H_2^+ molecule, could be analytically solved within a full quantum mechanical framework. Exact numerical solution covered for some more cases, but the number of "solvable systems" remains small. Several features contributes to this difficulty, but the most important is the two-body nature of the Coulomb interaction, which introduces high correlations between the degrees of freedom, making 3.3 a not-separable equation.

There are many approximations that could be used in order to conform 3.3 into a separable form that can be solved. I will present the approximation considered to be the cornerstone of the computational chemistry.

3.1.1 Adiabatic approximation

A straightforward simplification could be introduced, which arises from the observation that time scales associated with the motion of electrons and nuclei, are different. Namely nuclei move much slower than electrons: due to the small mass ratio between nuclear and electronic mass. In the hydrogen case, which represents the lower limit of this approximation, is less than 1/1836.

Based on this consideration, Born and Oppenheimer (1927) proposed a scheme for separating the nuclear and electronic variables, and showed that, under appropriate condition, no mixing of different electronic stationary states occurs due to interaction with the nuclei. In other words the electrons do not undergo transition between stationary states (adiabatic approximation). In this framework, the electrons could be thought as instantaneously following the motion of the nuclei, while remaining always in the same stationary state of the electronic Hamiltonian. This means that as the nuclei follow their dynamics, the electrons instantaneously adjust their wave functions according to the nuclear configuration (i.e. nuclear wave function).

In mathematical language, one could take as solution of 3.3

$$\Psi(\mathbf{R},\mathbf{r},t) = \sum_n \Theta_n(\mathbf{R},t)\Phi_n(\mathbf{R},\mathbf{r}), \tag{3.4}$$

where $\Theta_n(\mathbf{R},t)$ are wave functions describing the evolution of the nuclear subsystem, in each of the *adiabatic* electronic eigenstate $\Phi_n(\mathbf{R},\mathbf{r})$.

These satisfy the time-independent electronic Schrödinger equation

$$\hat{h}_e\Phi_n(\mathbf{R},\mathbf{r}) = E_n(\mathbf{R})\Phi_n(\mathbf{R},\mathbf{r}), \tag{3.5}$$

where the electronic Hamiltonian is:

$$\hat{h}_e = \hat{T} + \hat{U}_{ee} + \hat{V}_{ee} = \hat{\mathscr{H}} - \hat{T}_n - \hat{V}_{nn} \tag{3.6}$$

where $\hat{T}$ is the electronic kinetic operator, $\hat{U}_{ee}$ is electron-electron interaction, $\hat{U}_{ee}$, $\hat{V}_{ne}$ the electron-nuclear interaction, $\hat{T}_n$ and $\hat{V}_{nn}$ the inter-nuclear interaction potential operator.

By replacing the ansatz 3.4 into the Schrödinger equation, after some calculations, one ends up with mixed terms of the type $\int \Phi_q^*(\mathbf{R},\mathbf{r})\nabla_I\Phi_n(\mathbf{R},\mathbf{r})$.

Therefore, the reduction of the full wave function to an expression like

$$\Psi(\mathbf{R},\mathbf{r},t) = \Theta_n(\mathbf{R},t)\Phi_n(\mathbf{R},\mathbf{r}), \tag{3.7}$$

is still not completely correct, because even if a system is initially *prepared* in a *pure*, the off-diagonal terms will excite different electronic eigenstates along temporal evolution. These terms could be neglected, if the following condition is satisfied:

$$\frac{m}{M}\Big|\frac{\hbar\Omega_V}{E_q(\mathbf{R}) - E_n(\mathbf{R})}\Big| \ll 1 \tag{3.8}$$

where Ω_V is the maximum frequency of rotation of the electronic wave function due to the nuclear motion. Under these conditions, the Born-Oppenheimer approximation allows to pass from a N+M many-body problem, to the solution of the N electrons problem at fixed nuclear positions, and then to the solution of M nuclei problem, which move in a field generated by the N electrons (in absence of external fields).

Solving the electronic Schrödinger equation in a many-body is one key issue in physics, and many theories were developed to this aim. In the next section we will review one of the most important ones, which allows to treat relatively large systems at a reasonable computational cost.

3.2 Density Functional Theory

DFT is rigorous theory of the ground-state of many-particle system has gained in the last twenty years an important role in the computational science. The main idea beyond that approach exploits the one-to-one mapping between N particle density ($\rho(\mathbf{r})$) and the wave function of the N particle system.

This approach posseses two advantages:

i) the reduction of dimensionality of the problem from 3N variables into 3 (the spatial coordinates $\mathbf{r}$

ii) the use of a quantity (ρ) which is an observable easy to measure and visualize.

The Hohenberg-Kohn theorems

DFT theory is based on two theorems enunciated by Hohenberg and Kohn in 1964 [98]. The first one provides a justification for the use of density: if one consider a system with a Hamiltonian, which is characterized by an external potential V_{ext},

and the density $\rho(\mathbf{r})$, this theorem states: *the external potential V_{ext} is determined, within a trivial additive constant, by the electron density $\rho(\mathbf{r})$* .

Since the density determines the Hamiltonian, and also the number of electrons, it follows that ρ determines the ground-state wave function, and therefore all the observables of the system.

The second theorem provides a variational principle for the ground-state density: given any trial density $\tilde{\rho} > 0$ for which $\int \tilde{\rho}(\mathbf{r}) d\mathbf{r} = N$, it follows that $E[\tilde{\rho}] \geq E[\rho]$. From this result, one can get a variational equation to obtain the ground-state energy. Applying the Hohenberg-Kohn theorem to a system of N electrons in which the external potential is due to the nuclei, the energy in terms of the electronic density reads:

$$E[\rho] = T_e[\rho] + V_{ee}[\rho] + V_{en}[\rho] + V_{nn} = F[\rho] + \int d\mathbf{r}\rho(\mathbf{r})V_{ext} \qquad (3.9)$$

where $F[\rho] = T_e[\rho] + V_{ee}[\rho] = \langle \psi | T_e + V_{ee} | \psi \rangle$ is a universal functional independent from the external potential $V_{ext}[\rho] = V_{eN}[\rho] + V_{NN}$.

Applying to ρ the stationary principle

$$\delta \left\{ E[\rho] - \mu \left[\int \rho(\mathbf{r}) d\mathbf{r} - N \right] \right\} = 0, \qquad (3.10)$$

we obtain the Euler-Lagrange equation for the multiplier μ:

$$\mu = V_{ext}(\mathbf{r}) + \frac{\partial F[\rho]}{\partial \rho} \qquad (3.11)$$

The application of the variational principle requires in practice an explicit formulation of the functional F, which is not known in term of electron density.

3.2.1 The Kohn-Sham equations

The two theorems were also re-formulated into a suitable form to be computationally treatable. To this aim, Kohn and Sham propose the idea of mapping the system of N *interacting* particles into an equivalent one of non-interacting bodies, which is characterized by exactly the same ground-state density.

In that case, the density turns out to be expressed as a summation over single-particle contributions

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\varphi_i^{KS}(\mathbf{r})|^2, \qquad (3.12)$$

and the kinetic energy functional has a simple analytical expression:

$$T_0[\rho] = \sum_{i=1}^{N} \langle \varphi_i^{KS} | -\frac{1}{2}\nabla^2 | \varphi_i^{KS} \rangle. \qquad (3.13)$$

The functional $F[\rho]$ can be rewritten as:

$$F[\rho] = T_0[\rho] + V_H[\rho] + E_{xc}[\rho] \tag{3.14}$$

where $V_H = \frac{1}{2} \int d\mathbf{r'} \frac{\rho(\mathbf{r})\rho(\mathbf{r'})}{|\mathbf{r}-\mathbf{r'}|}$ is the classical part of the particle-particle interaction, and the 'exchange-correlation' functional E_{xc} is defined as:

$$E_{xc}[\rho] = T[\rho] - T_0[\rho] + V_{ee}[\rho] - V_H[\rho] \tag{3.15}$$

Formally E_{xc} can be written in terms of an exchange-correlation energy per particle ε_{xc}, which is itself functional of the total density:

$$E_{xc}[\rho] = \int d\mathbf{r} \rho(\mathbf{r}) \varepsilon_{xc}[\rho] \tag{3.16}$$

Equation 3.11, turns out to be:

$$\mu = V^{KS}(\mathbf{r}) + \frac{\partial T_0[\rho]}{\partial \rho} \tag{3.17}$$

with

$$V^{KS} = V_{ext}(\mathbf{r}) + \int d\mathbf{r'} \frac{\rho(\mathbf{r'})}{|\mathbf{r}-\mathbf{r'}|} + V_{xc}[\rho] \tag{3.18}$$

where we have defined the exchange correlation potential

$$\varepsilon_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho} \tag{3.19}$$

Eq. 3.17 says that we can solve the original problem by finding the ground-state energy for a system of non-interacting electrons in a effective potential. The single-particle orbitals describing these electrons solve the self-consistent KS equations:

$$\left[\frac{1}{2} \nabla^2 + V_{KS}(\mathbf{r}) \right] \varphi_i^{KS} = \varepsilon_i \varphi_i^{KS} \qquad i = 1, ..., N \tag{3.20}$$

The total energy of the system is not the sum of KS eigenvalues, but can be expressed as

$$E = \sum_{i=1}^{N} \varepsilon_i - \frac{1}{2} \int d\mathbf{r} d\mathbf{r'} \frac{\rho(\mathbf{r})\rho(\mathbf{r'})}{|\mathbf{r}-\mathbf{r'}|} + E_{xc}[\rho] - \int d\mathbf{r} V_{xc}(\mathbf{r}) \rho(\mathbf{r}) \tag{3.21}$$

With this method, the starting complexity of the problem is transferred to the problem of finding a suitable formulation of the exchange-correlation functional, which needs the introduction of other approximations.

3.3 Exchange-Correlation functionals

3.3.1 Local density approximation (LDA)

In this approximation, already proposed in the seminal paper of Kohn and Sham [99], and already introduced in the Thomas-Fermi-Dirac theory, the general inhomogeneous electronic system is assumed to be a locally homogeneous electron gas.

$$\varepsilon_{xc}^{LDA}[\rho] = \varepsilon_{xc}^{hom}(\rho(\mathbf{r})) \tag{3.22}$$

In that way, the otherwise *non-local* exchange-correlation functional becomes dependent on $\mathbf{r}$, i.e. *local*. By separating the two contributions to the E_{xc}, the exchange energy is exactly given by Dirac's expression:

$$\varepsilon_x^{LDA}[\rho] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}\rho^{\frac{1}{3}(\mathbf{r})} \tag{3.23}$$

Many good approximations for the correlation term are available, some of which solve analytically the problem in the high and low density limit [100, 101]. The most accurate results are given by quantum Monte Carlo simulations [102], which has been parameterized in the following ref. [103]. Suitable (approximate) analytical formulas have been derived by Vosko *et al.* (VWN correlation functional [104]) and by Perdew and Wang (PW) [105].

It is not within the aim of this thesis to outline good and bad aspects of LDA functional, but it is essential to mention that this functional is applicable to systems with slowly-varying densities, but cannot be formally justified for hughly inhomogeneous systems, such as to and molecules, anf therefore exhibits heavy deficiencies in describing biologically relevant systems [106, 107].

3.3.2 Gradient-Corrected approximations (GGA)

The progressive introduction of inhomogeneities in the electronic density could be done using an expansion of the density in terms of its gradient, and higher order derivatives. In general:

$$E_{XC}[\rho] = \int \rho(\mathbf{r})\varepsilon_{XC}[\rho(\mathbf{r})]F_{XC}[\rho(\mathbf{r}, \nabla\rho(\mathbf{r}, \nabla^2\rho(\mathbf{r}),...)]d\mathbf{r} \tag{3.24}$$

With the inclusion of these corrections GGAs functionals have been shown to give good results for all the main bond types (ionic, covalent, metallic and hydrogen bonds) [108], but still, as for LDA functionals, vdW interactions are poorly described [109].

One of the most popular GGA functional used in biochemistry is BLYP, which is derived from the combination of Becke's exchange functional B88 [110] and the correlation functional of Lee, Yang and Parr (LYP) [111].

3.3.3 Becke exchange functional

In this exchange functional the parameters were fitted to experimental molecular data, and the functional was formulated in a way to reproduce the exact asymptotic behavior of the exchange energy:

$$\varepsilon_x = \varepsilon_x^{LDA}\Big(1 - \frac{\beta}{2^{1/3}A_x}\frac{x^2}{1+6\beta x\sinh^{-1}(x)}\Big) \tag{3.25}$$

and

$$x = 2^{1/3}\frac{|\nabla\rho(\mathbf{r})|}{\rho(\mathbf{r})^{1/3}}, \tag{3.26}$$

x gives a measure of the inhomogeneity of the system, and the parameter $\beta = 0.0042$ was obtained by a fit on known Hartree-Fock (HF) data for rare gas.

3.3.4 Lee Yang Parr correlation functional: LYP

The Lee-Yang-Parr (LYP) functional for the correlation energy was derived from the Colle-Salvetti formula [111], and computes correlation energies from HF second order density matrices. Its expression reads:

$$E_c[\rho] = -a\int \frac{1}{1+d\rho^{-1/3}}\rho + b\rho^{-2/3}[C_F\rho^{5/3} - 2t_W + (\frac{1}{9}t_W + \frac{1}{18}\nabla^2\rho)]e^{-c\rho^{-1/3}}d\mathbf{r} \tag{3.27}$$

where $C_F = \frac{3}{10}(3\pi^2)^{\frac{2}{3}}$, $t_W = \frac{1}{8}\frac{|\nabla\rho|^2}{\rho} - \frac{1}{8}\nabla^2\rho$. a=0.04918, b=0.132, c=2533 and d=0.349 parameters are obtained by fitting the functional formula into HF results for the Helium atom.

3.3.5 Plane wave basis set

Following the Bloch's theorem [112] for periodic systems, a one-particle wavefunction can be written as Fourier's series:

$$\varphi^{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}}e^{i\mathbf{k}\cdot\mathbf{r}}\sum_{\mathbf{g}}c_j^{\mathbf{k}}(\mathbf{g})e^{i\mathbf{g}\cdot\mathbf{r}} \tag{3.28}$$

where V is the volume of the cell, $\mathbf{k}$ vectors belong to the first Brillouin zone, $\mathbf{g}$ is a reciprocal lattice vector, c is the first Fourier component of the plane-waves expansion, and the summation is extended to infinite lattice vectors. In the treatment of isolated clusters with a low symmetry, such as, organic molecules or the active sites of enzymes, the Γ-point approximation ($\mathbf{k}$=0) still guarantees a good accuracy, leading to a relevant reduction of the computational cost. The simulation of isolated clusters within periodic boundary conditions scheme needs some care, as self-interactions among replicas has to be canceled. In our calculations, we have used the procedure developed by Martyna and Tuckerman [113], which links the

expression of the electrostatic potential energy of a cluster to that of a infinitely periodic system, allowing thus to take advantage of the Ewald method (see 3.7.2) for isolated systems treated with a plane-wave orbital expansion. The algorithm is based on the assumption that the electrostatic potential Φ can be written as the sum of a short and long-range contributions:

$$\Phi(\mathbf{r}) = \phi^{long}(\mathbf{r}) + \phi^{short}(\mathbf{r}) \tag{3.29}$$

In a plane-wave expansion, the potential energy due to Φ can be expressed as

$$U = \frac{1}{2V}\sum_{\mathbf{g}}|\bar{\rho}(\mathbf{g})|^2\bar{\Phi}(-\mathbf{g}) = \frac{1}{2V}\sum_{\mathbf{g}}|\bar{\rho}(\mathbf{g})|^2[\bar{\phi}^{short}(-\mathbf{g}) + \bar{\phi}^{long}(-\mathbf{g})] \tag{3.30}$$

where $\bar{\rho}(\mathbf{g})$ and $\bar{\Phi}(-\mathbf{g})$ are the *finite* Fourier series of the density and the potential, respectively, the finite Fourier series of $f(\mathbf{r})$ being:

$$\bar{f}(\mathbf{r}) = \int_V d\mathbf{r}e^{-i\mathbf{g}\cdot\mathbf{r}}f(\mathbf{r}) \tag{3.31}$$

Requiring $\phi^{short}(\mathbf{r})$ to vanish exponentially quickly at large distances from the center of the cluster ($d \sim 2d_s$, where d_s is the typical system size), we can express this function as a Fourier transform plus an (exponentially vanishing) error term:

$$\bar{\phi}^{short}(\mathbf{g}) \qquad = \int_V d\mathbf{r}e^{-i\mathbf{g}\cdot\mathbf{r}}\phi^{short}(\mathbf{r}) \tag{3.32}$$

$$= \int_{all\,space} d\mathbf{r}e^{-i\mathbf{g}\cdot\mathbf{r}}\phi^{short}(\mathbf{r}) + \varepsilon(\mathbf{g}) \tag{3.33}$$

$$= \tilde{\phi}^{short}(\mathbf{g}) + \varepsilon(\mathbf{g}) \tag{3.34}$$

where $\tilde{\phi}^{short}(\mathbf{g})$ is the Fourier transform of $\phi^{short}(\mathbf{r})$. Thus, neglecting the error term:

$$\bar{\phi}(\mathbf{g}) \qquad = \tilde{\phi}^{short}(\mathbf{g}) + \bar{\phi}^{long}(\mathbf{g}) \tag{3.35}$$

$$= \tilde{\phi}^{short}(\mathbf{g}) + \bar{\phi}^{long}(\mathbf{g}) + \tilde{\phi}^{long}(\mathbf{g}) - \tilde{\phi}^{long}(\mathbf{g}) \tag{3.36}$$

$$= \tilde{\phi}(\mathbf{g}) + \tilde{\phi}^{screen}(\mathbf{g}) \tag{3.37}$$

The function

$$\tilde{\phi}^{screen}(\mathbf{g}) = \bar{\phi}^{long}(\mathbf{g}) - \tilde{\phi}^{long}(\mathbf{g}) \tag{3.38}$$

screens the interaction of the cluster with an infinite array of periodic images, realizing the link between isolated and replicated systems (its value is zero in this latter case). Thus, the potential energy can be written as:

$$U = \frac{1}{2V}\sum_{\mathbf{g}}|\bar{\rho}(\mathbf{g})|^2[\tilde{\Phi}(-\mathbf{g}) + \tilde{\phi}^{screen}(-\mathbf{g})] \tag{3.39}$$

In this expression $\tilde{\phi}^{long}(\mathbf{g})$ is supposed to be known, while $\bar{\phi}^{long}(\mathbf{g})$ can be evaluated efficiently over a grid using Fast Fourier Transform [114], which scales as $\mathscr{O}(N \ln N)$.

3.4 Pseudopotentials.

The greatest drawback in using a plane-wave basis-set comes from the impossibility, from a practical point of view, of describing core electrons within a reasonable computational expense. Indeed, the sharp spatial oscillations of their wave-functions near to the nuclei would require an extremely high number of plane-waves for an accurate characterization. On the other hand, the core levels are well separated in energy from valence electrons and, at a first level of approximation, do not play any role in the chemical properties of molecular systems. Thus, the core electron orbitals can be frozen in the KS equations and only the valence electrons are described explicitly. The core-valence electron interactions are implicitly included into the nuclear potential, which becomes an 'effective-potential' or 'pseudopotential'. Pseudopotentials are usually derived from all electron (AE) atomic calculations, and several recipes have been proposed to date. In the work presented in this thesis "norm-conserving" pseudopotentials derived from the Martins-Troullier (MT) method [115] have been used. These pseudopotentials have to satisfy the following conditions:

- The valence pseudo-wave-function should not contain any radial nodes.

- The valence AE and pseudopotential eigenvalues from the radial KS equations must be the same:

$$\varepsilon_\ell^{PP} = \varepsilon_\ell^{AE} \tag{3.40}$$

  where ℓ is the angular momentum.

- The pseudo and AE atomic radial wave-functions must be equal for r greater than a chosen cut-off distance r_{cut}.

These three conditions ensure that the pseudo-atom behaves like the real one in the region of interaction with other atoms while forming chemical bonds. Other conditions are the following:

- The integrated electron density within the cut-off radius for the two wave-functions must be the same. This requirement guarantees the transferability and the norm conserving rule of the MT pseudopotential.

- At $r = r_{cut}$, the pseudo wave-function and its first four derivatives should be continuous.

- The pseudopotentials should have zero curvature at the origin.

With these conditions, the general form for a pseudopotential wave-function is:

$$\varphi_\ell^{PP}(r) = \begin{cases} \varphi_\ell^{AE}(r) \ ; \ r > r_{cut} \\ r^\ell e^{p(r)}; \ r \leq r_{cut} \end{cases} \tag{3.41}$$

where $p(r) = c_0 + \sum_{i=1}^{6} c_i r^{2i}$, and the coefficients are obtained by imposing the first three conditions.

The functional form of the pseudopotential is

$$V_{pseudo} = V_{val}(r) + \sum_{m,l} |Y_{l,m}\rangle V_l(r) \langle Y_{l,m}| \tag{3.42}$$

where $|Y_{l,m}\rangle$ are spherical harmonics. The "semilocality" of this functional form (local in the radial coordinate, non local in the angular ones), implies an increase in the computational cost. This difficulty can be overcome by using the method of Kleinman-Bylander [116], which implies addition and subtraction of an "ad-hoc" radial function $V_L(r)$ to the pseudopotential, leading to a new functional form, where the local and non-local parts can be completely separated.

3.5   Car-Parrinello MD

Now that we have presented the method used in this work, to solve 3.6, with the help of suitable approximations, we face the prolem of how to solve the time-dependent Schröedinger equation, because we are interested in simulating the time evolution of biological systems (namely enzyme in this work) which undergo bonds breaking and formation: that requires a *first principle* approach.

Recalling 3.7, one can write the time-dependent adiabatic Schröedinger equation for the nuclear wave-function

$$i\hbar \frac{\partial \Theta_m(R,t)}{\partial t} = \left(-\sum_{I=1}^{P} \frac{\hbar^2}{2M_I} \nabla_I^2 + \varepsilon(\mathbf{R}) \right) \Theta_m(\mathbf{R},t), \tag{3.43}$$

with

$$\varepsilon_n(\mathbf{R}) = \varepsilon_n^o(\mathbf{R}) + \sum_{I=1}^{P} \frac{\hbar^2}{2M_I} \langle \Psi_q | \nabla_I^2 | \Psi_q \rangle, \tag{3.44}$$

which represents the diagonal correction to the electronic energy levels, due to the dependence of the electronic wave function on the nuclear coordinates [1]

Using Ehrenfest's theorem, mean values of position and momentum operator could be obtained:

$$i\hbar \frac{d\langle \mathbf{R} \rangle}{dt} = \langle [\hat{\mathscr{H}}, \mathbf{R}] \rangle = i\hbar \frac{\mathbf{P}}{M} \tag{3.45}$$

---

[1]This correction is proportional to m/M, and therefore feasible: when explicitly neglected, it receives the name of *Born Oppenheimer* approximation, when included one speaks of *adiabatic* approximation.

and

$$ i\hbar\frac{d\langle\mathbf{P}\rangle}{dt} = \langle[\hat{\mathscr{H}},\mathbf{P}]\rangle = -i\hbar\langle\nabla\varepsilon_n(\mathbf{R})\rangle, \tag{3.46} $$

which combined give rise to the following Newtonian equation of motion:

$$ M\frac{d^2\langle\mathbf{R}\rangle}{dt^2} = -\langle\nabla\varepsilon(\mathbf{R})\rangle. \tag{3.47} $$

Using Hellmann-Feynman theorem [117, 118], it could be possible to write the second term in 3.47, in the *classical nuclei approximation* (which means assuming to have obtained the mean value of the position operator in some way):

$$ M_I\frac{d^2\mathbf{R}_I(t)}{dt^2} = -\left\langle\Psi_n(\mathbf{R})\left|\frac{\partial h_e(\mathbf{R})}{\partial\mathbf{R}_I}\right|\Psi_n(\mathbf{R})\right\rangle - \frac{V_{nn}(\mathbf{R})}{\partial\mathbf{R}_I}, \tag{3.48} $$

where V_{nn} is the nuclear-nuclear interaction.

A first principle self-consistent MD approach can be obtained by solving the electronic structure problem for the ground-state (for example within the Khon-Sham approach to DFT).

In this way one can obtain the electronic ground-state for each MD steps, but this requires diagonalization of the electronic Hamiltonian at each step, which takes a lot of computational time.

3.5.1 the Car-Parrinello Lagrangian

In 1985 Car and Parrinello developed a new scheme, based on the extended Lagrangian formalism and avoiding the optimization of the electronic wave function of every step by introducing a second order *fictitious* dynamics on the electrons. These latter are kept sufficiently close to the adiabatic surface, allowing for an increase of the time step by a factor ~ 10 with respect to Ehrenfest dynamics. The method is based on the observation that $\langle\Psi_0|\mathscr{H}|\Psi_0\rangle$ can be viewed not only as a function of $\{\mathbf{R}_I\}$, but also as a *functional* of the wave function Ψ_0, and thus of the set of one-electron orbitals $\{\psi_i\}$ used to build it. In this case, the force acting on these orbitals can be obtained from a functional derivative of a suitable Lagrangian containing $\langle\Psi_0|\mathscr{H}|\Psi_0\rangle$, like in classical mechanics for the nuclear motion. The Lagrangian $\mathscr{L}$ proposed by Car and Parrinello has the form

$$ \mathscr{L} = \sum_{i=1}^{M}\frac{1}{2}M_I\dot{\mathbf{R}}_I^2 + \sum_{i=1}^{N}\frac{1}{2}\mu_i\left\langle\dot{\psi_i}\middle|\dot{\psi_i}\right\rangle - \langle\Psi_0|\mathscr{H}_e|\Psi_0\rangle + constraints \tag{3.49} $$

where the first term is the kinetic energy of nuclei, and $\mu_i = \mu$ are the "fictitious masses" assigned to orbitals; the second term represents the fictitious kinetic energy associated to them (the sum is on the occupied orbital only). The (holonomic) constraints act in general on both the orbitals (e.g. to guarantee orthonormality) and on the nuclei (e.g. if one would perform molecular dynamics with geometric

restraints). The dynamics is described by the Euler-Lagrange equations associated to $\mathscr{L}$

$$\frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{\mathbf{R}}_I} = M_I \ddot{\mathbf{R}}_I = \frac{\partial \mathscr{L}}{\partial \mathbf{R}_I} = -\frac{\partial}{\partial \mathbf{R}_I}\langle \Psi_0|\mathscr{H}_e|\Psi_0\rangle + \frac{\partial}{\partial \mathbf{R}_I}\{constraints\} \quad (3.50)$$

$$\frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{\psi}_i^*} = \mu_i \ddot{\psi}_i^* = \frac{\partial \mathscr{L}}{\partial \psi_i^*} = -\frac{\partial}{\partial \psi_i^*}\langle \Psi_0|\mathscr{H}_e|\Psi_0\rangle + \frac{\partial}{\partial \psi_i^*}\{constraints\} \quad (3.51)$$

Note that if $|\mu \ddot{\psi}_i| \to 0$ eq. 3.51 reduces to a stationary problem, and the electronic system will stay on the Born-Oppenheimer surface (no forces acting on orbitals), corresponding to the true equilibrium dynamics. Higher is the fictitious kinetic energy $T_e = \sum_{i=1}^N \frac{1}{2}\mu_i \langle \dot{\psi}_i|\dot{\psi}_i\rangle$, more the electrons will be far from the minimum energy configuration. In particular, a ground-state wave function optimized at time t_0 will stay close to its ground-state if it is kept at sufficiently low temperature. The only quantity one can change to ensure this condition is μ, often called "adiabaticity parameter" [119]. If μ and τ are chosen consistently the energy flow between electronic and nuclear subsystems is slow enough to cause no drift in T_e [2], thus conserving the "physical" total energy E_{phys}:

$$E_{phys} = E_{tot} - T_e = T_I + V_e = \sum_{i=1}^M \frac{1}{2}M_I \dot{\mathbf{R}}_I^2 + \langle \Psi_0|\mathscr{H}_e|\Psi_0\rangle \quad (3.52)$$

The choice of a reasonable fictitious mass μ meets two opposite requirements. In fact, considering a simple harmonic model for the electronic system around the Born Oppenheimer (BO) surface, (discrete) excitations frequencies are given by

$$\omega_{ij}^e = \sqrt{\frac{2(\varepsilon_i^* - \varepsilon_j)}{\mu}} \quad (3.53)$$

where ε^* and ε are energy levels of unoccupied and occupied orbitals, respectively. If ω_{max}^n is the maximum vibrational frequency of the nuclear system, in order to perform adiabatic dynamics it must be $\omega_{gap}^e >> \omega_{max}^n$. As the only tunable parameter is μ, one could decrease it arbitrarily to increase the frequency of the gap ω_{min}^e. However, decreasing μ stretches the entire spectrum $\{\omega_{ij}^e\}$ and in particular increases ω_{max}^e, which is inversely proportional to the maximum time step. Typical values of μ are in the range $500 - 1500$ a.u., which allow for a time step of about $5 - 10$ a.u. $(0.12 - 0.24$ fs$)$. For calculations discussed here we used $\mu = 600$ a.u. and a time step of 5 a.u.

---

[2] T_e actually performs two-frequency *bound* oscillations around a constant value. The first frequency is associated to the drag exerted by the nuclei, and it is in anti-phase with V_e oscillations, while the second is a small-amplitude high-frequency oscillation intrinsic to the fictitious electronic dynamics. Note that having a non vanishing masses, also the electrons dampen nuclear motion, causing a renormalization of the nuclear masses which can be important in the case of light atoms.

3.6 Hybrid Models: QM/MM molecular dynamics

The biological systems like those addressed in this study, namely enzyme in action, require special computational techniques. They are proteins, usually composed by a thousands of atoms, and they perform chemical reactions, which happen in a limited part of the enzyme, called active site (usually smaller than a hundred of atoms). Treating the whole system quantum mechanically, by using *first principle* approaches, is computationally not-affordablei. In addition this may not necessarily cause one is interested only in the catalytic region (a hundred of atoms). The approximation of the enzyme to only a models of its catalytic site in vacuo is rather drastic, and one looses electrostatic contributions of the whole system and a correct description of the internal motions of the protein, which are crucial for a full description of the catalytic mechanism. For that issue, hybrid models which allow to treat different parts of the system at different level of theory were developed [120, 121, 122, 123]. According to these approaches, the Hamiltonian for a hybrid system could be written as:

$$H = H_{QM} + H_{MM} + H_{QM/MM} \qquad (3.54)$$

where H_{QM} and H_{MM} are the Hamiltonian for the quantum and classical systems respectively, and the remaining term $H_{QM/MM}$, which incorporates the essence of hybrid methods, describes the interactions between the two subsystems.

In this work we have benefit the use of the hybrid method developed by U. Roethlisberger and co-workers [123], in which the QM part is treated at DFT/BLYP level of theory (as described in 3.2) and dynamics is performed following Car-Parrinello idea, whereas classical part is described using AMBER force-field [124].

An important issue when using hybrid methods, is the way in which the quantum and the classical subsystems are divided: it is possible to group degrees of freedom as bonded and non-bonded interactions, dependently on how the QM/MM boundary cuts or not chemical bonds of molecule.

3.6.1 Bonded interactions

The first problem that one has to face when a QM atom is chemically bonded to a MM one, is the remaining unsaturated valence of the terminal QM atom. Among the possible proposed solutions [125, 126, 127, 128, 129]: here we have simply added an additional (dummy) hydrogen atom to saturate the dangling bond.

Stretch, angle and dihedral interactions involving MM and QM atoms are kept in account by the classical MM force field used.

3.6.2 Non-bonded interactions

The natural choice for the non bonded interaction Hamiltonian is:

$$H_{non-bonded} = \sum_{i \in MM} q_i \int dr \frac{\rho(r)}{|r - r_i|} + \sum_{i \in MM} \sum_{j \in QM} v_{vdW}(r_{ij}), \qquad (3.55)$$

The steric non-bonded interactions due to the Pauli repulsion and the dispersion interactions are kept into account by retaining the van der Waals interaction as described by the MM force field.

The electrostatic interaction between the QM density and the point MM charges representing the main environmental effect on the QM system, presents some issues: (i) the electron spill-out phenomenon and (ii) a very high computational cost.

In the QM/MM scheme followed here [123], these problems are overcome by treating in a different way short and long range interactions.

3.6.3   Short-range electrostatic interactions

The so-called *electron spill-out* problem is a consequence of the non-physical attraction of the electron cloud, caused by positive MM charges, located near the QM/MM boundary. These positive charges are not surrounded by its own electron density, and the Pauli repulsion is therefore absent: as a consequence the QM electron density is over polarized.

This effect is particularly pronounced in a PW basis set approach, in which the electrons are fully free to de-localize.

In the solution proposed in [123], the Coulomb potential is replaced by

$$H_{el} = \sum_{j \in MM} q_j \int dr \rho(r) \frac{r_{cj}^4 - r^4}{r_{cj}^5 - r^5}, \qquad (3.56)$$

where r_{cj} is the covalent radius of the atom j. In that way, for large r value, the 1/r behavior of the Coulomb potential is still present, while it tends to finite number for small r.

This choice of the smoothing function has been tested to produce accurate results for the structural properties of a quantum water molecule in a box of classical water molecules without any *ad hoc* re-parametrization of the force field, and an accurate description of hydrogen bond distances and energetics can be obtained.

3.6.4   Long-range electrostatic interaction

The explicit calculation of the long-range electrostatic term is too expensive in a plane-wave based approach. In fact, the calculation of about $N_r \cdot N_{MM}$ integrals, where N_r is the number of real-space grid points ($\approx 10^3$) and N_{MM} is the number of classical atoms ($\approx 10^5$) would be required. This problem is solved by lowering the degree of accuracy of the calculation as the MM atoms get further from the QM region. The MM region is partitioned into three regions by providing two cut-off radii r_1 and r_2.

- Interaction of the QM charge distribution with all MM atoms within r_1 is explicitly calculated.

- For MM atoms with $r_1 < r < r_2$, the Coulomb interaction is approximated to interaction between the MM atomic point charges and the D-RESP charges [130] of the QM system (see below for D-RESP charges derivation). Thus, the electrostatic Hamiltonian reads:

$$\mathcal{H}^{el}_{QM/MM} = \sum_{i,j}^{r_1 \leq r \leq r_2} \frac{q_i Q_j^{RESP}}{|R_i - R_j|} \tag{3.57}$$

- for $r > r_2$ The MM charges interact with a multi-polar expansion (up to the quadrupolar term) of the QM charge distribution. In this case, the Hamiltonian takes the form:

$$\begin{aligned}
\mathcal{H}^{el}_{QM/MM} &= C \sum q_i \frac{1}{|r - R_i|} + \sum_{\alpha} D^{\alpha} \sum q_i \frac{(R_i^{\alpha} - \bar{r}^{\alpha})}{|r - R_i|^3} + \\
&+ \frac{1}{2} \sum_{\alpha\beta} Q^{\alpha\beta} \sum q_i \frac{(R_i^{\alpha} - \bar{r}^{\alpha})(R_i^{\beta} - \bar{r}^{\beta})}{|r - R_i|^5} + ...
\end{aligned} \tag{3.58}$$

where $\bar{r}$ is the origin of the multi-polar expansion (e.g., the geometrical center of the quantum system) and C, D and Q are the total charge, the dipole and the quadrupole quantum charge distributions, respectively.

The only free parameters in this approach are the cut-off radii r_1 and r_2. This approach allows a fully Hamiltonian description of the electrostatic interactions. Thus, energy conservation during dynamics is achieved if the potential energy and the forces are consistently computed. The correctness of the implementation can be directly verified by monitoring the conservation of the energy during MD simulations.

3.7 Classical molecular dynamics

All the techniques described above, albeit accurate, are computationally expensive and allow to treat relatively small problems. More than this, not all the questions addressed in computational biophysics involve bond breaking or formation (for example one could be interested in study the large scale motion of a protein, or movement of some particular domain). For all these cases, an accurate description as obtained by *ab-initio* calculations is out of scope, and sometime even a waste of computational resources. In the so called *force field* methods, the electronic energy for a given nuclear configuration is not calculated solving the Schröedinger equation, using some approximate methods, but it is given as a parametric function of nuclear coordinates, which are fitted to experimental values, or higher level

computational data. In this framework electrons are not explicitly considered, and the problem becomes a classical one: the equation of motion are therefore obtained using a standard Hamiltonian (or Lagrangian) formalism, and trajectories are obtained solving these equations of motion using some *ad hoc* algorithm [131].

In this work we have used not only in the classical MD calculations, but in mixed QM/MM as well, the AMBER force field [132], whose functional analytical form is:

$$U = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\vartheta (\vartheta - \vartheta_{eq})^2 + \qquad (3.59)$$
$$+ \sum_{dihedrals} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i<j} [4\varepsilon_{ij} (\frac{\sigma_{ij}}{r_{ij}})^{12} - (\frac{\sigma_{ij}}{r_{ij}})^6 + \frac{q_i q_j}{\varepsilon r_{ij}}]$$

Atom bond stretching and angle bending are represented as harmonic terms, while dihedrals or torsionals are described by a sinusoidal term. Non-bonded interactions comprise two terms, the first is a Lennard-Jones (LJ) 6-12 which describes atom-atom repulsion and dispersion interactions, the second is the Coulomb electrostatic term. In eq. 3.59, r and ϑ are respectively the bond length and valence angle; ϕ is the dihedral or torsion angle and r_{ij} is the distance between atoms i and j. Parameters include the bond force constant and equilibrium distance, K_r and r_{eq}, respectively; the valence angle force constant and equilibrium angle, K_ϑ, and ϑ_{eq}, respectively; the dihedral force constant, multiplicity and phase angle, V_n, n, and γ, respectively. Non-bonded parameters between atoms i and j include the partial atomic charges, q_i, along with the LJ well-depth, ε_{ij}, and σ_{ij}, the (finite) distance at which the inter–particle potential is zero. These terms are also referred to as the interaction or external parameters. Typically, ε_{ii} and σ_{ii} are obtained for individual atom types and then combined to yield ε_{ij} and σ_{ij} for the interacting atoms via combining rules. The dielectric constant ε is typically set to 1 (corresponding to the permittivity of vacuum) in calculations that incorporate explicit solvent representations [133]. Van der Waals and electrostatic interactions are calculated between atoms belonging to different molecules or for atoms in the same molecules separated by at least three bonds.

3.7.1   Long-range interactions in classical MD

In simulations of biological systems it is highly desirable to avoid the calculation of all non-bonded pair interactions, since the computational cost would be proportional to the square of the number of atoms. These interactions primarily dictates the dynamics of biomolecules, and cannot be merely truncated beyond a given cutoff [131] when long-ranged. An interaction can be roughly defined long-range if the associated energy falls off slower than r^{-d}, where d is the dimensionality of the system. Thus, Coulomb ($\sim \phi^{-1}$) and dipole-dipole ($\sim \phi^{-3}$) should be considered long-range when dealing with three dimensional systems. Indeed, cutting the interaction on a sphere of radius r_c makes it is easy to see that the tail correction, proportional to $\int_{r_c}^{\infty} \phi(r) 4\pi r^2$ diverges if the potential function goes to zero with $r \rightarrow \infty$

faster than r^{-3}. [3] Fortunately efficient methods to handle the evaluation of long-range interactions were developed. Among these, the most used are reaction field methods and lattice methods [134]. The first class assumes that beyond a given distance the interaction can be treated using an "average field" approach, derived from macroscopic electrostatic. The second class of methods, much more widely used, is represented by the Ewald summation algorithm [135], which efficiently handles long-range interactions in periodic systems (and thus requires Periodic Boundary Conditions - PBC - to be used for the system being simulated).

3.7.2 Ewald sum

Within PBC each particle interacts with all the other N-1 particles into the simulation box and with all the N particle images in an infinite 3D array of periodic cells. The electrostatic potential energy U^{Coul} of the infinite system, therefore, takes the form:

$$U^{Coul} = \frac{1}{2} \sum_{\mathbf{n}=0}' \sum_{ij}^{N} \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{n}L|} \qquad (3.60)$$

where L is the length of the periodic box (assumed cubic for simplicity), N is the total number of atoms, and $\mathbf{n}$ are the direct lattice vectors. The prime on the summation indicates that the sum is over all the images $\mathbf{n}$, and $i \neq j$ if $\mathbf{n} = 0$. The convergence of this series is extremely slow, and thus a very large number of images is required to achieve a reliable estimate of U^{Coul}. A simple recipe to cope with this problem consists in splitting the electrostatic energy in two parts, a short-range and a long-range term. This means to find some function β such that the Coulomb potential can be rewritten ($x \equiv |\mathbf{r}_{ij} + \mathbf{n}L|$):

$$\frac{1}{x} = \frac{\beta(x)}{x} + \frac{1 - \beta(x)}{x} \qquad (3.61)$$

If β is properly chosen, the first term will be negligible beyond a given cutoff x_c, and the latter will be a slowly varying function of x, that can be expressed as Fourier Series with a low number of $\mathbf{k}$ points. Ewald [135] made for $\beta(x)$ the natural choice of the complementary error function $\text{erfc}(x) = (2/\sqrt{(\pi)}) \int_x^\infty \exp(-u^2)du$, such that $1 - \beta(x) = 1 - \text{erfc}(x) = \text{erf}(x)$. Thus, in eq. 3.61 the term $\text{erfc}(x)$ is evaluated cutting the interactions at a given cut-off, while the long-range term is Fourier transformed in the reciprocal space, where it becomes short-ranged and can be thus accurately estimated by summing over a limited number of reciprocal vectors (both the sums in real and reciprocal space becomes in fact exponentially converging). The parameter α tunes the relative weights of real and reciprocal sums, although the final result is independent of it. An optimal choice makes the Ewald

---

[3] The LJ interactions, instead, can be cut at a given cutoff (typically 10-12 Å), since they decay very rapidly. Methods have been developed that treat the region beyond the cutoff radius r_c as being homogeneous (i.e. with LJ parameters averaged over all the LJ atom types) [131]. These interactions are thus effectively treated as short-range interaction, with the addition of a constant correction term.

sum to converge as $N^{3/2}$, which can be further improved to $N \ln N$ with the use of Particle-Mesh methods (as the Particle-Mesh Ewald, PME [136] or the Particle-Particle Particle-Mesh, PPPM [137]), making advantage of the Fast Fourier Transform (FFT) [114] algorithm.

3.8   Performing MD in a statistical ensemble

Solving the equation of motion presented above, independently on the level of theory used (classical MD or *ab initio*), will generate trajectories which map a micro-canonical ensemble[4], as the total energy is strictly conserved during time evolution. Experiments on the systems I have addressed here (catalytic enzymes), but generally for every biophysical experiments, are performed at constant pressure and temperature (NPT ensemble), or constant volume and temperature (NVT ensemble). This means that, from the statistical mechanics point of view, it is necessary to use special techniques which allow one to sample statistical relevant canonical ensembles.

3.8.1   Nosé thermostat

A way to achieve constant temperature simulations, within the framework of MD, was introduced about twenty years ago by Nosé [138] and reformulated by Hoover [139], who modify Newton equation of motion by adding two non physical variables, thus introducing the following non-Hamiltonian dynamical system:

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i}, \tag{3.62}$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i - \frac{p_\eta}{Q}\mathbf{p}_i, \tag{3.63}$$

$$\dot{\eta} = \frac{p_\eta}{Q}, \tag{3.64}$$

$$\dot{\mathbf{p}}_\eta = \sum_{i=1}^{N} \frac{\mathbf{p}_i{}^2}{m_i} - LkT. \tag{3.65}$$

where $\{\mathbf{r}_i\}, \{\mathbf{p}_i\}$ are coordinates and momenta of the N particles with masses m_i, the forces $\mathbf{F}_i$ are derived from the N-particle potential and L is a parameter to be determined. The two nonphysical variables η and p_η in eq. 3.62 regulate the fluctuations in the total kinetic energy of the physical variables, and can be thus regarded as an effective "thermostat" for the physical system. The parameter Q controls the strength of the coupling to the thermostat: high values result into a low coupling and vice versa. These equations of motion conserve the following quantity:

$$C = H(\mathbf{p}, \mathbf{r}) + \frac{p_\eta^2}{2Q} + LkT\eta \tag{3.66}$$

---

[4]assuming that ergodic hypotheses are verified

where $H(\mathbf{p},\mathbf{r})$ is the physical Hamiltonian. Assuming that C is the only constant of motion, the microcanonical partition function at temperature T can be written as:

$$\Omega_T(N,V,C) \quad = \int d^N\mathbf{p}\,d^N\mathbf{r}\,dp_\eta\,d\eta\,e^{3N\eta}\,\delta(H(\mathbf{p},\mathbf{r}) + \frac{p_\eta^2}{2Q} + LkT\eta - C) \quad (3.67)$$

$$= \frac{e^{3NC/LkT}}{LkT}\int dp_\eta e^{-3Np_\eta^2/2QLkT}\int d^N\mathbf{p}\,d^N\mathbf{r}\,e^{-3NH(\mathbf{p},\mathbf{r})/LkT} \quad (3.68)$$

that, with $L = 3N$, is proportional to the canonical partition function for the physical system. Thus, the Nosé "extended" scheme gives conformations of the physical systems that belong to the canonical ensemble. In eq. 3.67, the metric $e^{3N\eta}$ has been used as $d^N\mathbf{p}\,d^N\mathbf{r}\,dp_\eta\,d\eta\,e^{3N\eta}$ is the phase-space conserved measure for this set of equation of motion.

Nosé-Hoover chains

When there is more than one constant of motion a single Nosé-Hoover thermostat is unable to get the correct canonical sampling for the physical system [131]. To allow the simulation of more general systems Martyna *et al.* [140] devised a scheme in which additional dynamical variables are inserted in the Lagrangian. This corresponds to couple the first thermostat to another one, and so on, generating a whole chain of thermostats (here the name Nosé-Hoover chains). In fact, as the momentum of the extended variable entering in the first Nosé-Hoover thermostat is distributed canonically, it can also be thermostated. The chains take into account for additional conservation laws, and generates the correct canonical distribution. The equations of motion for a system coupled to N_c thermostats are:

$$\dot{\mathbf{r}}_i = \frac{\dot{\mathbf{p}}_i}{m_i}, \tag{3.69}$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i - \frac{p_{\eta_1}}{Q_1}\mathbf{p}_i, \tag{3.70}$$

$$\dot{\eta}_k = \frac{p_{\eta_k}}{Q_k}, \quad k = 1,...,N_c, \tag{3.71}$$

$$\dot{\mathbf{p}}_{\eta_1} = (\sum_{i=1}^{N}\frac{\mathbf{p}_i^2}{m_i} - LkT) - \frac{p_{\eta_2}}{Q_2}p_{\eta_1}, \tag{3.72}$$

$$\dot{\mathbf{p}}_{\eta_k} = (\frac{\mathbf{p}_{\eta_{k-1}}^2}{Q_{k-1}} - kT) - \frac{p_{\eta_{k+1}}}{Q_{k+1}}p_{\eta_k}, \tag{3.73}$$

$$\dot{\mathbf{p}}_{\eta_{N_c}} = (\frac{\mathbf{p}_{\eta_{N_c-1}}^2}{Q_{N_c-1}} - kT). \tag{3.74}$$

The dynamics generated by these equations conserves the following "energy":

$$H_{NHC} = H(\mathbf{p},\mathbf{r}) + \sum_{k=1}^{N_c}\frac{p_{\eta_k}^2}{2Q_k} + LkT\eta_1 + \sum_{k=2}^{N_c}kT\eta_k \tag{3.75}$$

3.8.2   Isobaric simulations: Langevin piston

As for the thermostat many methods [134] have been developed for the aim of performing simulations at constant pressure. Here will be presented only the one used in this work (in classical MD performed using the NAMD code [141]). The equation of motion of each atom are modified by the introduction of an additional degree of freedom, the piston σ, which undergoes a Langevin-like dynamics, in analogy with the motion of each particle

$$\dot{\mathbf{r}}_i \qquad = \frac{\mathbf{p}_i}{m_i} + \frac{1}{3}\frac{\dot{V}}{V}\mathbf{r}_i, \tag{3.76}$$

$$\dot{\mathbf{p}}_i \qquad = \mathbf{f}_i - \frac{1}{3}\frac{\dot{V}}{V}\mathbf{p}, \tag{3.77}$$

$$\ddot{V} \quad = \frac{1}{W}[P(t) - P_{ext}] - \gamma\dot{V} + R(t), \tag{3.78}$$

$$\tag{3.79}$$

where γ is the collision frequency and $R(t)$ is a random force taken from a Gaussian distribution with zero mean and variance

$$\langle R(0)R(t)\rangle = \frac{2\gamma k_b T \delta(t)}{W} \tag{3.80}$$

where k_b is Boltzmann's constant.

3.9   Free energy calculations

The biochemical events addressed with computational approaches usually involve the crossing of relatively high energy barriers: in particular, one aim of this work is the characterization of enzymatic reaction, for which we have experimentally measured quantities (k_{cat}) of the order of tenths of kcal/mol.

   The time required for crossing such high energy barriers is not accessible with a normal, canonical, molecular dynamics simulations, and methods for enhanincing the sampling of the free energy surface (or the phase-space) are necessary. For that, a number of powerful methods have been developed, in recent years, to characterize the thermodynamics of many rare events [142, 143, 144, 145, 146, 147, 148, 149, 150, 151].

   Forcing the system to undergo the requested transformation, and then reconstracting the free energy profile, implies two major problems: (i) the simplification of the phase space into few (or even only one) relevant paramenters, called *reaction coordinates*, along which to enforce the transformation, and (ii) to achieve the definition of efficient schemes for sampling the reduced phase space.

   In principle reduction of the complexity and intrinsic multidimensionalty of a system to the few reaction coordinates could seem a rather drastic simplification, but for some case (in particular for some enzymatic reactions), the low-dimensional approximations are reasonable. In particular in this work, the hydrolysis reaction performed by the enzyme are approximated using a monodimensional reaction coordinate, and profile of (Helmotz) free energy are here briefly presented.

3.9.1 Constraint MD

The *Constraint MD* method has been developed by Ciccotti *et al.* [152, 153] and has been used to a variety of molecular systems up to date.

In this scheme, the reaction free energy is calculated by rigidly constraining the chosen reaction coordinate. The force on the constraint ($\mathfrak{f}$) is time-averaged upon successive variations of the reaction coordinate value $\mathscr{R}$, and then integrated along $\mathscr{R}$. The free energy, then, reads:

$$\Delta G(\mathscr{R}) = \int_{\mathscr{R}_0}^{\mathscr{R}} \bar{\mathfrak{f}}(\mathscr{R}') d\mathscr{R}' \tag{3.81}$$

For the simple case of an interatomic constraint, like in the case of the simulations of this work, the force on the constraint is equal to the corresponding Lagrange multiplier:

$$\bar{\mathfrak{f}}(\mathscr{R}) = \langle \lambda(\mathscr{R}) \rangle = -\left\langle \frac{\partial V}{\partial R_{AB}} - \frac{2k_B T}{R_{AB}} \right\rangle_{R_{AB}=\mathscr{R}} \tag{3.82}$$

where the brackets stand for an ensemble average over the trajectory, and the reaction coordinate $\mathscr{R}$ is defined by the distance between atoms A and B.

The validity of this thermodynamic integration scheme is independent on the nature of the interatomic forces, and it has been successfully extended to a variety of chemical systems. Practically, the value of $\mathscr{R}$ is increased or shortened in a small amount starting from the final configuration of the simulation at the former $\mathscr{R}$ value. The amount of these increments determines the accuracy of the free energy estimation for e.g. a chemical transformation: nonetheless, the shorter they are, the more time-consuming the simulation will be.

Part II

Computational studies on B2 MβL CphA

Chapter 4

Protonation state and substrate binding of CphA

4.1 Summary

The zinc enzymes metallo β-lactamase counteract the beneficial action of β-lactam antibiotics against bacterial infections. These enzymes hydrolyze the β-lactam ring. To understand structure/function relationships on a representative member of this class, the B2 MβL CphA, we have investigated the H-bond pattern at the Zn site and substrate binding mode by molecular simulation methods. Extensive QM calculations at the DFT-BLYP level on 11 models of the protein active site, along with MD simulations of the protein in aqueous solution, allow us to propose two plausible protonation states for the unbound enzyme, which are probably in equilibrium. Docking procedures along with MD simulations and QM calculations suggest that in the complex between the enzyme and its substrate (biapenem), the latter is stable in only one of the two protonation states, in addition it exhibits two different binding modes, of which one agrees with previous proposals. In both cases, the substrate is polarized as in aqueous solution. We conclude that addressing mechanistic issues on this class of enzymes requires a careful procedure to assign protonation states and substrate docking modes.

4.2 CphA in the free state.

We attempted to identify the most plausible protomers of B2 MβL active site by evaluating the relative stability of different protomers representing the **CphA** active site, optimized at the DFT-BLYP level of theory (Tab.4.1).

57

Model	His118	His196	4^{th}ligand	RMSD (Å)	ΔE (kcal/mol)
A	ε	ε	WAT	0.15	0
A2	δ	δ	WAT	0.29	15.15
A3	ε	δ	WAT	0.20	7.10
A4	ε	δ	WAT	0.37	1.70
B	ε and δ	ε	OH	0.14	0.67
B2	δ	δ	OH	0.37	39.99
B3	ε	ε	OH	0.16	36.13
B4	ε	δ	OH	0.39	50.97
B5	ε	δ	OH	0.21	52.77
B6	ε	ε	OH	0.39	27.80
B7	ε and δ	ε	OH	0.25	2.38

Table 4.1:   Different protomers used in the QM calculations.   The first three columns report the protonation of His118, His196 and the ligand that replaces crystallographic carbonate anion [154]. Structure (RMSD) and relative energies (ΔE) of different protomers representing the **CphA** active site are reported in the last two columns.

The Zn(II) ion was assumed to be tetrahedral, based on the spectroscopic evidence [155]. These models included: i) The metal and its protein ligands Asp120 (assumed to be ionized), Cys221 and His263; the latter binds Zn(II) on Nε and reasonably H-bonds with Gln68, This residue was not included in models, since its inclusion was expected not to significantly alter the results of calculations: H-bond Gln68@CO... HNε@His263 would be present in all models considered, and would stabilize them in the same extent. ii) Either a water molecule (Waz) or a hydroxide ion (Oz) acting as fourth ligand; this ligand replaces the carbonate (or bicarbonate) of the X-ray structure.

iii) The three residues of the "Zn1" binding site: Asn116, His118 (protonated either in Nδ, H-bonding to water Wat2, or in Nε, H-bonding to OH-Tyr117. This residue was not included in calculation. According to the crystal structure [154], Tyr117@HO...HNε@His118 is most likely to be present, compared to Tyr117@OH...Nε@His118 (putative HO...N angle of about 220°, Fig. SI): for that reason most of our models were built with His118 protonated on Nε. Explicit inclusion of Tyr117 would be more effective on stabilization of models with His protonated on Nε than on those protonated on Nδ. Furthermore, A2 and B2, the only two models with His118 protonated on Nδ, turned out to be less stable compared to all other models. and His196. The latter faces water in the crystal, and it was consequently assumed to be protonated either in Nε (thus forming and H-bond with Waz or Oz,) or on Nδ, forming a H-bond with a third water molecule (Wat3); iv) Arg121 was included for its positive charge, which is expected to play an important role in electrostatics of the active site cavity [93, 94]. v) Two additional water molecules (Wat2 and Wat3), which H-bond to His118 and Asp120, and to Asn116 and His196, respectively.

The most stable models in terms of energetic and structural properties (RMSD relative to the X-ray structure) turned out to be **A** and **B** in which the fourth Zn ligand is a water or a hydroxide group, respectively (Tab. 4.1 and in Fig. 4.1).

The energies of the two models differed by less than 1 kcal/mol and their structural parameters were fairly similar to those of the X-ray structure (RMSD= 0.14 Å for **A** and 0.13 Å for **B**); in particular, the Zn(II) ion exhibits a tetrahedral coordination sphere. Two other models turned out to have low energy: A4 and B7 (Tab. 4.1).

Model A4 differs from model A only in the protonation of His196. We performed short test MD simulation of this protomer observing that the $\chi 1$ and $\chi 2$ of His196 adjust, exposing the NδH group to the solvent, in a similar geometry as those of His196 in model A. Therefore we decided to use the protonation state of A. Model B7 features the same protonation state of B, differing only in an additional crystallographic water molecule. (Tab. 4.1) We exclude all the other models considered by using an energetic criterion and/or a structural one (Tab. 4.1). In **A**, His118 and His196 are protonated in Nε (Fig. 4.1): Wat2 is involved in a network of H-bonds with Oδ2@Asp120, Nδ@His118 as acceptors, and one hydrogen of Waz as donor. In **B** His118 is diprotonated, and Wat2 forms a H-bond network with O@Oz and Oδ2@Asp120 as acceptors, and NδH@His118 as donor (Fig. 4.1). In both models Wat3 is H-bonded to the Nδ of His196 and the amide group of Asn116. Overall,

a comparison between these two models indicates that the difference is a relayed proton transfer between the metal bound Waz and His118. Thus, both protomers could coexist, being almost isoenergetic and they could be also in equilibrium with each other.

The stability of the protein in the protonation states **A** and **B** was then investigated by MD simulations of the entire proteins in aqueous solutions. The overall folds were well conserved for 10 ns, as shown by RMSD value of the backbone (MD-averaged value 1.1 Å, Fig. 4.2).

The RMSF calculated over Cα atoms (Fig. 4.3a) pointed to a mobility of regions L1, helixα3/L5 and L9, located near the edge of active site: these values are in good agreement with the experimental B-factors [154], except for L1, which has low experimental values.

The geometrical parameters of the Zn(II) site were preserved during the dynamics in both models. In the **CphA-A** model a second water molecule was approaching the metal site with an exchange time of about 0.3 ns. The H-bonds between NεH@His118 and O@Tyr117, NH@His118 and Oδ@Asn116, and His263 Nδ and Gln68 as well as the Cys221-Arg121 salt bridge, were well conserved in both models.

Novel information is obtained on the structural role of Lys224, which in the X-ray structure was involved in a salt bridge with the exogenous carbonate anion. During the dynamics, it is H-bonded to Tyr297, Tyr242, Leu232 and His196 Nδ (Fig. 4.2). The structure of this H-bond network prevents the terminal positive charge of the Lys residue to be solvent exposed. In summary, our calculations pointed out that both **CphA-A** and **CphA-B** protomers may be present in resting-state of **CphA**.
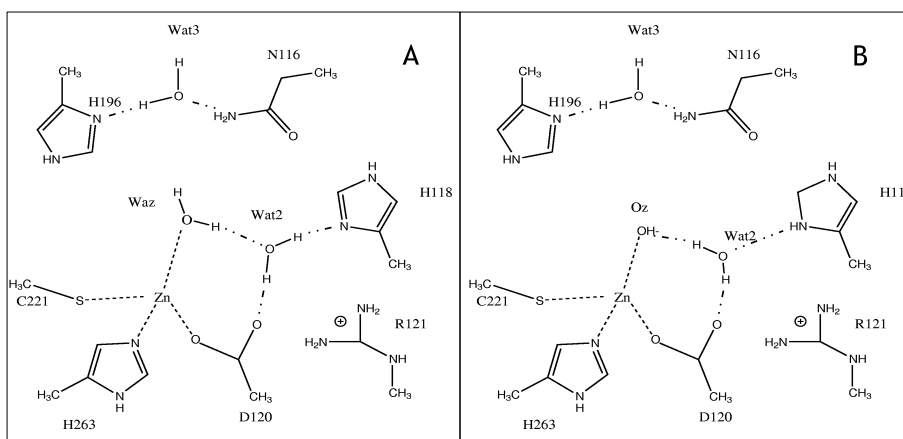

Figure 4.1: Active site protomers of **CphA** emerging from our DFT-BLYP calculations. The 4th Zn ligand is a water molecule (Waz) in **A**, and an OH group (Oz) in **B**.
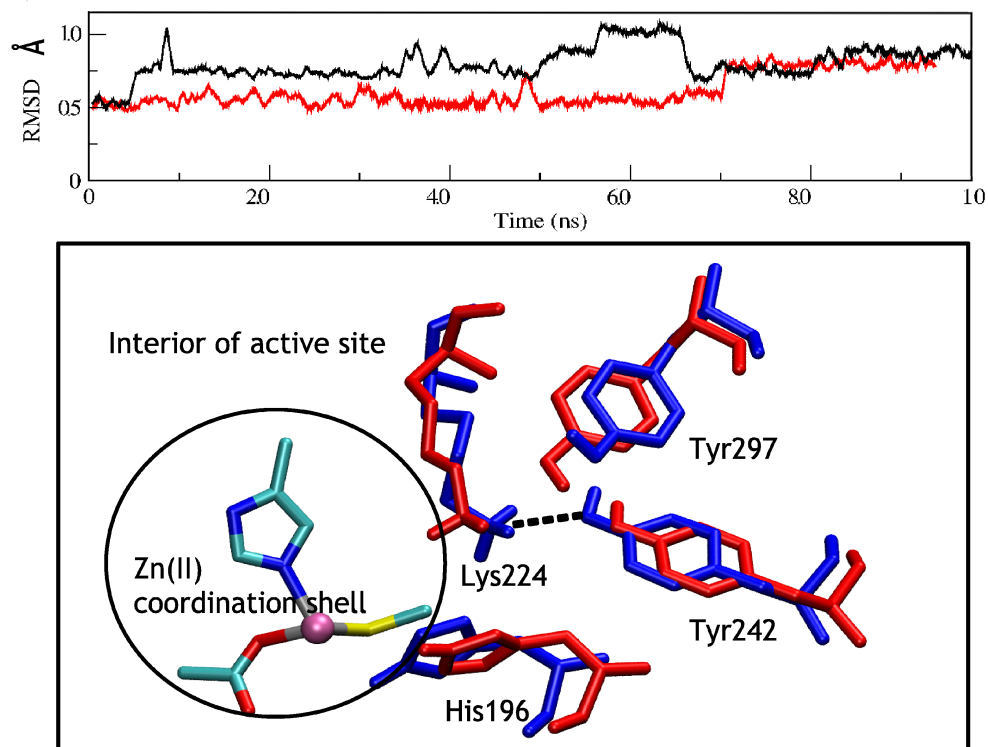
Figure 4.2: MD simulations of **CphA** in the free state. Top: the RMSD (Å) of **CphA-A** (red) and **CphA-B** (black).
Bottom: initial (red) and final (blue) representative structures from MD simulations of **CphA-A**: in the absence of the carbonate anion, Lys224 forms salt bridges with two tyrosines (Tyr242 and Tyr 297) and it is buried in the active site cavity. The MD structure of **CphA-B** turned out to be very similar to **CphA-A**.
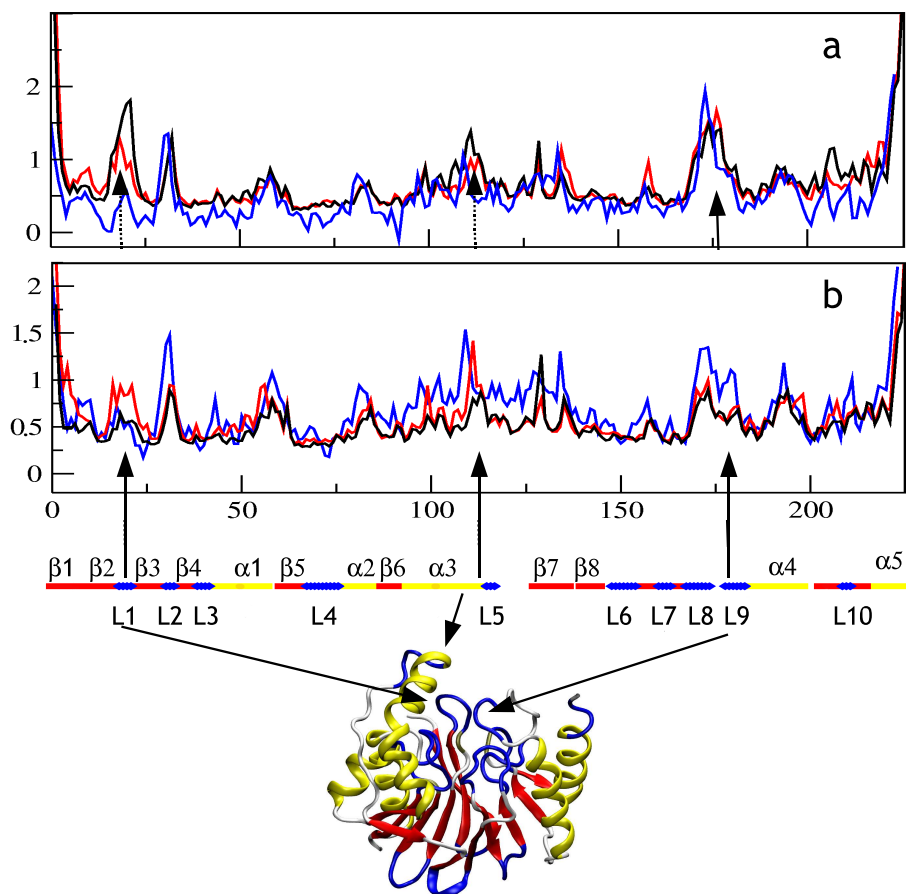
Figure 4.3: Panel (a): calculated and experimental B-factors (blue) for **CphA-A** (red) and **CphA-B** (black). Panel (b): the B-factors of CphA in complex with **Bia** derivative [154] are compared with those of our stable MD complexes, namely those of **ES-1A**(red) and **ES-2A** (black).
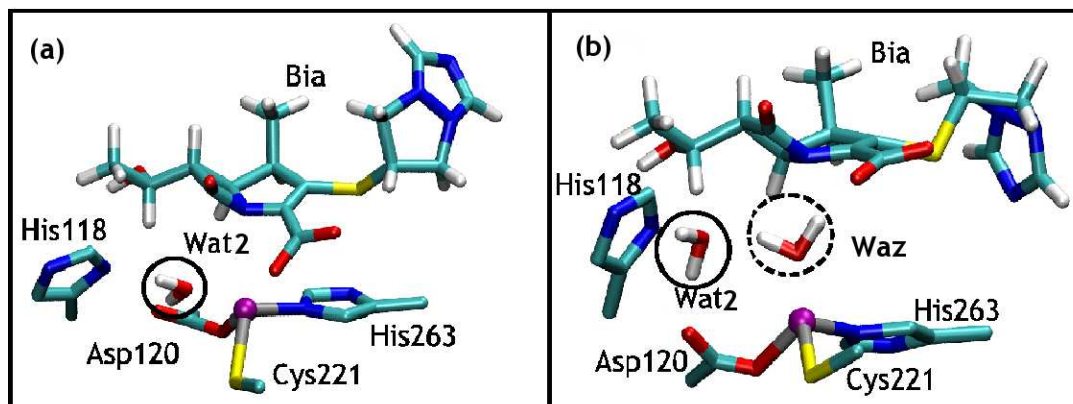
Figure 4.4: **ES-1A** (a) and **ES-2A** (b) MD average structures. The putative water molecules (Wat2) acting as nucleophile in the enzymatic reaction are highlighted with circles; the second zinc coordinated water (Waz) is indicated with dashed circle in (b).

4.3 CphA-Bia Michaelis complex.

In this section we investigate the stability of the complex between **CphA** and its substrate (**Bia**) in both protonation states by performing MD simulations of the adducts. Two docking models were considered for each protonation state of the protein residues of the zinc environment. These models differ also for the Zn coordination: in models **ES-1** the Zn is bound to the substrate, whilst in models **ES-2** a water molecule or an hydroxide are directly coordinated with the metal ion (Fig. 4.5). The latter is based on the hypothesis that B2 MβL use this moiety to perform the catalytic reaction, as found in all B1 MβLs (COOREGGI REFS). Specifically: in models **ES-1A** and **ES-1B** the carboxylic moiety of **Bia** was coordinated to the metal ion, as in the crystallographic structure of **CphA** in complex with a **Bia** derivative [154], and with a salt bridge between Lys224 and O(12)@**Bia**. The possible nucleophile was a water molecule or an hydroxide (Wat2 or OH-2) respectively, placed in the pocket between His118 and Asp120. Instead, models **ES-2A** and **ES-2B** were based on the protomers **A** and **B** in which **Bia** was docked into the active site without assuming direct binding of the carboxylate moiety to the zinc ion (Fig. 4.4, 4.5). In that model, Lys224 forms a salt bridge with O(13)@**Bia**.

After 10 ns of classical MD simulation of **ES-1A** and **ES-2A** the RMSD of the final MD structures were close to those of the free states, ranging from 0.9 Å to 1.1 Å ($\pm$ 0.1 Å). The enzyme folds were fully maintained along the dynamics and the structure of **Bia** was well conserved (RMSD 0.15 Å $\pm$ 0.01 Å). Wat2 formed stable hydrogen bonds with Nδ@His118 and with Oδ1@Asp120. In model **ES-2A** Waz acted always as fourth ligand to the metal ion, H-bonding with carboxylate group of **Bia**.

The H-bond pattern in the enzyme was similar to that of the free state, except for Lys224, that was interacting with O13@**Bia** and O12@**Bia**, other **Bia-CphA** interactions include O(14)@**Bia**-Nε@His196, and NH@Asn233-O(14)@**Bia** (Fig. 4.5), while O(12)@**Bia** interacts with Asn233 backbone amide in **ES-1A**, but not in **ES-2A**; His196 is mobile in both adducts. The OH group of **Bia** was trapped between Thr119 and Thr157 sidechains, as in the X-ray structure of the adduct with the **Bia** derivative.

CphA-**Bia** hydrophobic contacts involved Val67, that together with Tyr70 forms the loop L1, and C(10) methyl@**Bia** and also with the C(17) and C(18) of the 3,5-diaza-1-azoniobicyclo group.

Therefore, **Bia** was well anchored to the enzyme cavity and the distance between Wat2 and the C(7) (the carbon that undergoes the nucleophilic attack) of the β-lactam ring is 3.3 ± 0.1 Å and 3.5 ± 0.1 Å for **ES-1A** and **ES-2A**. Consistently, the RMSF of the Cα (Fig. 4.3) was similar to that of the free state, except for L1 and L9, which were smaller, possibly because of the formation of interactions involving Asn233, Val67 and **Bia**.

Thus, both MD structures could be considered as productive Michaelis complexes.

In contrast, the MD simulations of **ES-1B** and **ES-2B** showed that **Bia** rearranged dramatically or moved almost completely out of the active site. This may be caused, at least in part, by the electrostatic repulsion between negatively charged OH (Oz or OH-2) and carboxylate group of **Bia**. The latter pushed **Bia** outside the active pocket and caused the breaking of the hydrogen bonds between Lys224 and the O(12) of the carboxylic group of **Bia**, and the one between the His196 and the other oxygen atom of the same group. As **Bia** moved outside of the active site, we discarded both docking models in the **B** protonation state.

4.4   Discussion of results

The study of the structural and dynamic features of MβLs can provide useful information for the design of new effective antibiotics or lactamase inhibitors, which are likely to escape the selection of pathogens resistant to present agents. A key issue in this respect is the knowledge of the protonation state of enzymatic active sites, which is a prerequisite for investigating the reaction mechanism and for computer-aided design of inhibitors. Here, we have addressed this issue in the context of **CphA**, the only B2 MβLs for which the X-ray structure is presently available. The stability in terms of structure and energetics of 11 plausible protomers of the active site in the free state (Tab 4.1) was evaluated by QM and MD simulations. Two protomers were the most likely ones. They were practically isoenergetic, reproduced the structural features of the active site in QM calculations and conserved the protein fold during 10 ns of dynamics. The protomer **A** was characterized by the presence of a Zn-coordinated water molecule and His118 protonated in Nε, and the protomer **B** was characterized by an hydroxide Zn-coordinated and a diproto-
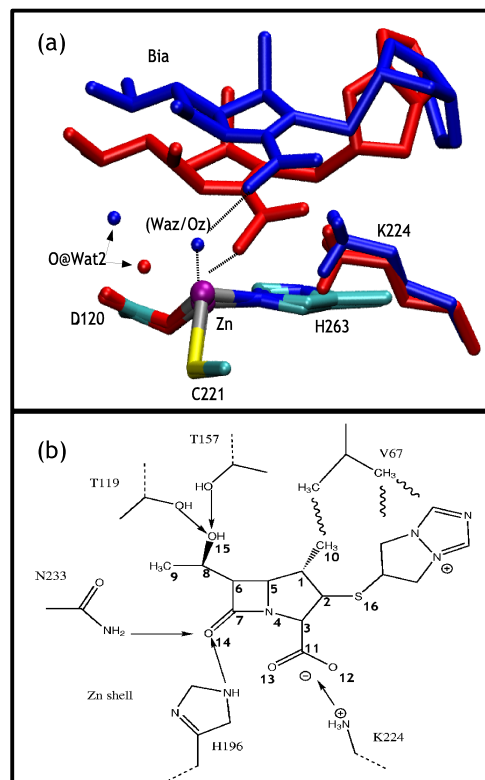
Figure 4.5: (a) Main differences in docking pose of the antibiotic in **ES-1A**, **ES-2A**, **ES-1B** and **ES-2B**. Red the orientation of **Bia** in models **ES-1**, in which the carboxylic group binds directly to the Zn. Blue the orientation of the antibiotics for models **ES-2**. The oxygen of the fourth metal ligand (either a water molecule (Waz) in **ES-2A** or a hydroxide (Oz) in **ES-2B**), which binds the carboxylic group of **Bia** in models 2, and those of the possibles nucleophiles (Wat2 in models **ES-1A**, **ES-2A** and **ES-2B**, OH-2 for **ES-1B**) are also shown. (b) The substrate/enzyme stable interactions, emerging from MD simulations, are shown: H-bonds and salt bridges are represented by arrows, while the wavy lines indicate the hydrophobic contacts. During the dynamics both **ES-1A**, **ES-2A** maintained all these interactions, while **ES-1B** and **ES-2B** turned out to be unstable. Lys224 interacts with O(12) in model **ES-1A**, while with O(13) in **ES-2A**.

nated His118 (Fig. 4.1). Asp120 turns out to be deprotonated in both cases. Both models contained a water molecule within ≈ 4 Å from the zinc and not bound to this ion. The MD simulations of the protein in both protonation states further confirmed that **CphA** possesses a compact fold, revealing, however, some degrees of mobility of the L1 loop, helixα3/L5 and L9 regions, structures located at the entrance of the active site.

In B1 enzymes, L3 loop is the analogous of L1 in B2, but it is longer, being formed by 7-8 residues compared to only two residues, Tyr60 and Val67, in the latter subclass. NMR measurements [156] and classical MD calculations on B1 [157, 158], showed that this loop, located at the entrance of the active site, is highly mobile, and became stabilized by the binding of a substrate or an inhibitor.

In **CphA**, L1 edges the active site from the N-terminal domain, and a decrease in mobility emerges from calculations, as a consequence of interaction with the carbapenem -SR groups upon substrate binding. These effects were not evidenced by X-ray experiments and are unexpected findings of MD calculations, which suggest a possible role of L1 for the mechanism of binding carbapenem molecules.

We conclude that an equilibrium between the two protonation states should be considered for a detailed description of the structural properties of the enzymes in resting state.

The binding of the substrate **Bia** was investigated by docking **Bia** in **CphA**'s active site. Two different poses were considered. The first one was obtained starting from the X-ray structure of the adduct of **CphA** with a **Bia** derivative [154], in which the nucleophile (water molecule or an hydroxide) was located in a pocket between His118 and Asp120 (models **ES-1A** and **ES-1B**). In particular, we have used the experimental information, showing that in the structures of CphA in complex with **Bia** derivative the substrate carboxylate group interacts directly with the zinc and with Lys224. Docking was then performed assuming that this interaction characterizes also the binding of the substrate before the nucleophilic attack.

The second pose was obtained by docking the **Bia** in the active site, in the structures obtained from **CphA-A** and **CphA-B**, preserving the main protein ligand interactions as in previous pose, but assuming here that the carboxylic O13@**Bia**-Zn bound is not present before the nucleophilic attack (models **ES-2A** and **ES-2B**). This mode is built under the proposal that two solvent molecules should be present in Zn environment to complete the catalytic reaction, as found in all B1 MβLs [159, 24, 160, 161, 162, 163, 164, 165, 166, 157, 167, 168].

The dynamics of **ES-1B** and **ES-2B** caused either a dramatic rearrangement of **Bia** or a departure of **Bia** from the active site, pointing to the instability of these complexes. This instability may be caused, at least in part, by the electrostatic repulsion between the Zn-bound hydroxide and **Bia** carboxylate.

In contrast, **ES-1A** and **ES-2A** turned out to be very stable during the dynamics. Residues His196, Lys224 and Asn233 were crucial for **Bia** binding by means of salt bridges or H-bonds with carboxylate and carbonyl groups.

Thus, our calculations suggest that **Bia** may bind in two ways to the protein, and that the reaction mechanism of the CphA could be different from the one re-

cently proposed by QM/MM calculations [97], in which the ligand binds directly to the Zn.

In addition, our findings establish that the hydrolysis mechanism of members of subclass B2 is different with respect to B1 MβLs, where a Zn(II) bound hydroxide is believed to promote the cleavage of the β-lactam ring [25,27-33].

Chapter 5

Insights from the reaction mechanism of CphA enzyme

5.1 Summary

The bacterial metallo β-lactamases (MβLs) are rapidly evolving, because the evolutionary pressure of β-lactam antibiotics. These enzymes accommodate one or two zinc ions, in its 'Zn1' and 'Zn2' sites, responsible for substrate hydrolysis. Recent X-ray studies on a member of the B2 MβLs, CphA from *Aeromonas hydrophila*, have prompted a variety of mechanistic studies of this important subclass, specific for carbapenems.

Hybrid QM/MM simulations reported here, support a novel mechanistic view, consistent with experimental data and pointing to a common functional role of the zinc ion in the 'Zn2' binding site in B1 and B2 MβLs. This single metal is able to promote the activation of the nucleophile and the engagement of a catalytic water for the completion of the hydrolytic cycle. Based on a variety of computational studies along with recent experimental evidences, we conclude that the conserved functional motif can be specifically targeted for the design of improved antibiotics as well common inhibitors.

5.2 The Michaelis complexes

Based on classical molecular dynamics (MD) simulations described on 4, we suggested that **ES-1** [14, 169, 97] and **ES-2** (Chart 5.1) are plausible Michaelis complexes [169].

They differ essentially by one non-proteic ligand coordinated to the metal ion, which is either the substrate (**ES-1**), as proposed specifically for this enzyme [14, 97], or a water molecule (**ES-2**), as already proposed across the B1 members [28, 27, 22, 23, 26].

The nucleophilic water would be activated, in both cases, by an H-bond acceptor, suggested to be either His118 [14] or Asp120 [97].
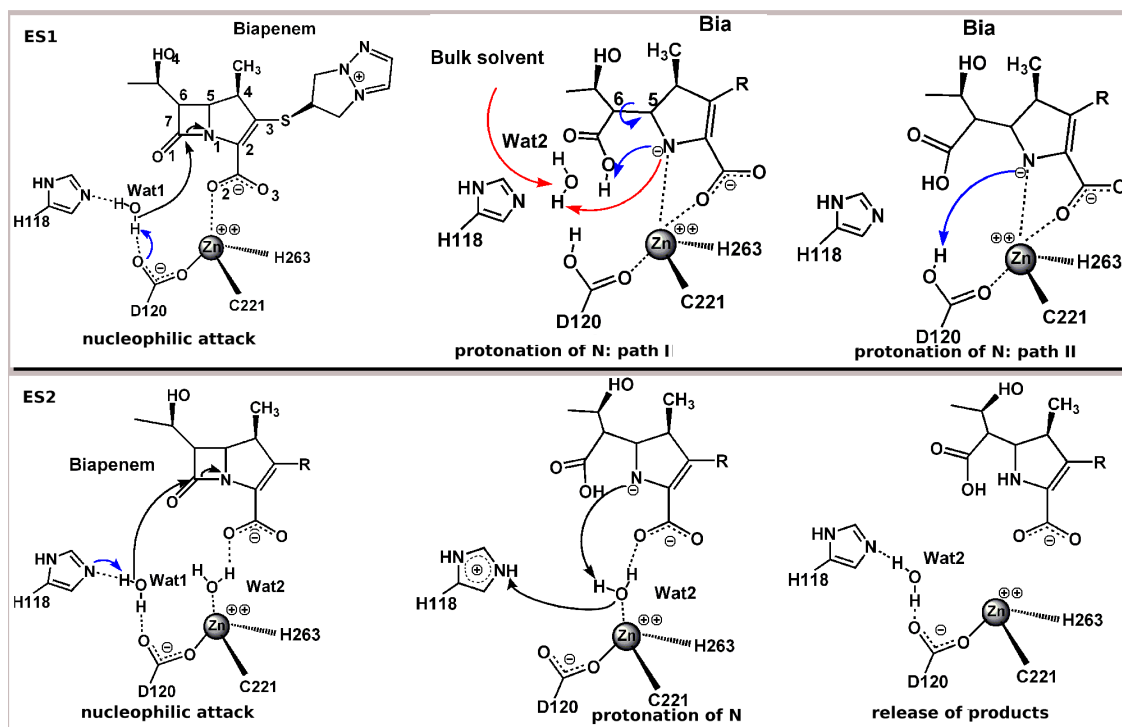
Figure 5.1: Schematic representation of the postulated reaction mechanism of CphA, a B2 MβL. **Top panel**: proposals reported so far [14]. In the Michaelis complex **ES-1**, **Wat1** would be activated by being deprotonated by a general base, either His118 or Asp120. Subsequently, the negatively charged N(1) atom of the substrate would be protonated. In **path I** an internal geometry rearrangement of **Int** around the C(5)-C(6) bond, or the coming of a second water (**Wat2**) from the bulk solvent (the two possibilities are depicted in blue and red respectively), could bring a proton donor in a suitable position for its transfer to N(1)@**Bia**. In **path II** such proton would come from the general base [97]: in the figure we show Asp120 as general base and proton donor. **Bottom panel**: possible catalytic mechanism emerging from this work. In the Michaelis complex **ES-2**, two water molecules are simultaneously present in the Zn catalytic pocket. As in **ES-1**, **Wat1** would be activated by a generalized base (His118 has been shown in this picture) and performs the nucleophilic attack. However, the second water (**Wat2**), which is bound to the Zn(II) ion and H-bonds to **Bia** carboxylate, could simultaneously act as proton donor in the protonation of the negatively charged β-lactam nitrogen.

Here, we used Car-Parrinello [170] DFT-BLYP [171, 172] based QM/MM [173, 174] simulations to investigate the reaction mechanism of CphA with **ES-1** and **ES-2** conformations as initial reactant state. The free energy has been calculated as function of a selected reaction coordinate, using the thermodynamics integration method [152].

5.3 Reaction mechanism based on ES-1

The Michaelis complex. Substrate and Zn-protein interactions were maintained during the QM/MM simulations. In particular, the substrate forms interactions with Val67 and Lys224 (Fig. SI 3), consistently with mutagenesis studies [14, 175], which point to an important role of these two residues.

However, substrate binding to the Zn(II) ion, which is a key facet of this mechanism [14, 97] was readily disrupted (see O(2)@**Bia**-Zn distance Tab. SI 1). This might be a consequence of the salt bridge formed between the substrate and Lys224 (Fig. SI 3). As a result, **Wat1** could enter the coordination shell of Zn(II), producing a slightly distorted tetrahedral geometry of the metal environment (Tab. SI 2).

Key factors for the alterations of Zn(II) polyhedron, as seen during QM/MM simulations, may include temperature effects and/or by the electrostatic and mechanical contribution of the enzyme frame. In fact the simple energy minimization of the QM part only (i.e. in the gas phase), using exactly the same exchange-correlation functional (BLYP [171, 172]) and plane wave basis set, turned out to give similar results to those reported by Xu *et al.* [97], which energy-minimized a similar model using B3LYP [171, 172, 176] and 6-31++G** basis set: in particular, **Bia**'s carboxylate remained bound to the metal ion. These results explain why in the previous report this Michaelis complex appeared to be stable, whereas it is actually unstable. Thus, several effects, including the mobility of the Lys224, the presence of the protein frame and the accurate description of the chemical bonding between Zn and donor atoms (as obtained in a QM framework), are needed for a proper modeling of the Michaelis complex.

5.3.1 Simulation of the enzymatic reaction.

The distance between the nucleophilic water oxygen O@**Wat1** and the C(7) β-lactam carbon ($d_{O-C(7)}$) was selected as reaction coordinate (RC, see also Methods section). This choice is similar to those used in others studies [22, 23, 97] and, to a first approximation, is the natural one to describe the direct nucleophilic attack.

$d_{O-C(7)}$ was progressively decreased in steps of 0.2 Å until the transition state (**TS1**) was identified at 1.7 Å (namely, where the force on the constraint changed sign). The coordination polyhedron of **ES-1** was maintained until **TS1** was reached; then the polyhedron was subsequently modified due to the simultaneous occurrence of several events (Fig. 5.2): (i) **Bia**'s carboxylate oxygen and the β-lactam nitro-

gen N(1) entered in the ligand shell of Zn; (ii) Asp120 and **Wat1** detached from the metal ion; (iii) the catalytic **Wat1** was deprotonated by Asp120, becoming an activated hydroxide; (iv) the β-lactam C-N bond weakened, as shown by the increased distance and fluctuation of the C-N bond (the distance increased from 1.5 $\pm$ 0.05 Å to 1.7 $\pm$ 0.1 Å).

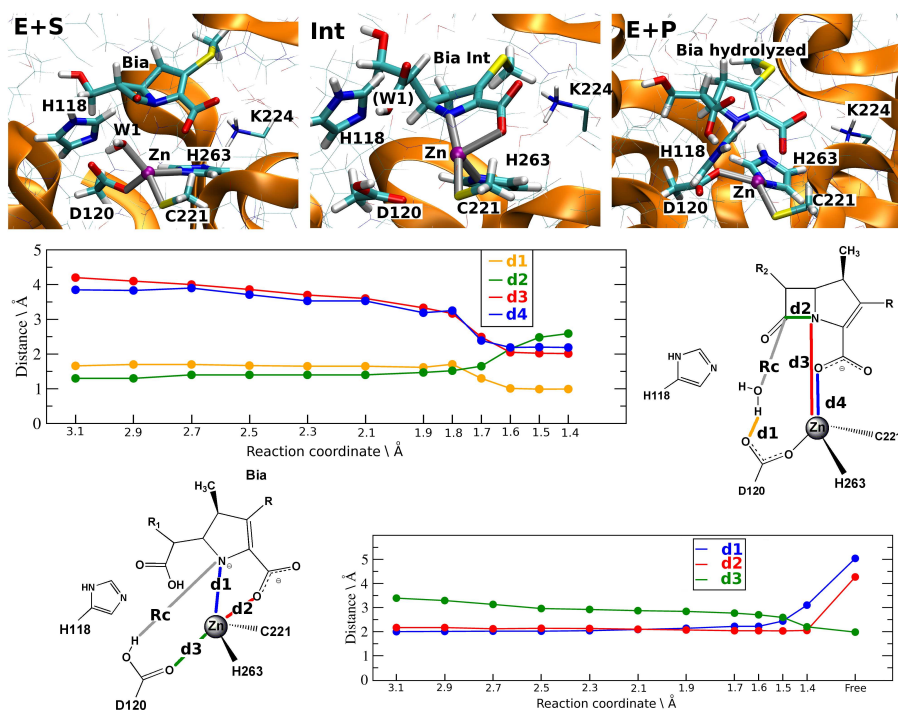Figure 5.2: Investigation of **ES-1**. **Top**: representative structures of the adduct (called **E+S** in the picture), intermediate (**Int**) and products (**E+P**) along the reaction path starting from **ES-1**. QM atoms along with the key residue Lys224 are shown in bold and thin licorice respectively. **Middle**: evolution of some key distances as function of RC, during the nucleophilic attack. In particular: d1 = d(H@**Wat1**, d2 = d(N(1)@**Bia**, d3 = d(N(1)@**Bia**-C(7)@**Bia**), d4 = d(O(2)@**Bia**-Zn). **Bottom**: evolution of key distances during the protonation nucleophilic attack, at various RC values (d1 = d(N(1)@**Bia**-Zn), d2 = d(O(2)@**Bia**-Zn) and d3 = d(N(1)@**Bia**-Zn)).

The calculated activation free energy for this step is $\Delta F^{\ddagger} \sim 15 \pm 3$ kcal/mol, and this value is not dissimilar from the experimentally measured data ($\Delta G^{\ddagger} 14$ kcal/mol) [14] and the computational estimation by Xu *et al.* ($\Delta F^{\ddagger} \sim 14 \pm$) [97].

By further decreasing RC, we identified a stable intermediate of the reaction (**Int**), and calculated that $\Delta F^{(\mathbf{Int})} \sim 9 \pm 2$ kcal/mol. From **TS1** to **Int**, the C-N distance further increased (Tab. SI 1) until the bond was cleaved. In the intermediate

the substrate directly coordinated with the metal ion: N(1)@**Bia** and O(2)@**Bia** formed a direct coordination bond with Zn(II) (Tab. SI 1). The remaining ligands were Cys221 and His263, and the coordination sphere was characterized by a highly distorted tetrahedral geometry (see Fig. 5.2, structure of **Int**, and Tab. SI 1). For the rest of the active site, the same H-bond network of **ES-1** was maintained, apart from the interactions formed by atoms involved in the nucleophilic attack.

N(1)@**Bia** and O(2)@**Bia** were located in a similar position as that occupied by the same atoms in the crystal structure of the complex **Bia** derivative/CphA [14]. From the **Int** state, two different scenarios have been proposed for the second step of the reaction speculated: (i) a rotation of C(5)-C(6) bonds along with the entry of a water molecule (**Wat2**) in the catalytic pocket (**path I**, Fig. 5.1) [14], or (ii) a proton transfer from Asp120@H to N(1)@**Bia** (**path II**, Fig. 5.1) [97].

We explored both these proposals by molecular simulation methods starting from the **Int** structure.

As for (i), we noticed that the large internal structural rearrangement of **Int**, and/or possible re-solvation of the active site, required time-scales well beyond first principle QM/MM calculations. Therefore some hints may be obtained by classical MD In the time-scale investigated (5 ns) our calculations showed no relevant rearrangements or re-solvation within the active site pocket. Although these events may occur in longer time-scales, our calculations provided no support for the mechanism of Garau *et al.* [14].

As for (ii), we performed constrained QM/MM, in which the distance between H@Asp120 and N(1)@**Bia**, selected as our reaction coordinate for this step, was progressively decreased (Fig. 5.2 and Tab. SI 2). The resulting free energy barrier is $\Delta F^{(\mathbf{Int})\ddagger} \sim 15 \pm 2$ kcal, giving a barrier of $\sim 24 \pm 3$ kcal/mol for the overall reaction. This value highly overestimates the experimental finding, even more if one considers the well-known deficiencies of DFT-BLYP scheme in calculating the energy of transition states and intermediates [177, 178]. Thus, our calculations do not seem to support the hypothesis proposed in ref. [97].

We conclude that the Michaelis complex **ES-1** may not be the correct one to describe the complete hydrolysis in CphA. Furthermore, the two-steps mechanism proposed by Xu *et al.* [97] is ruled out due to the high free energy barrier of the reaction. The mechanism which undergoes internal rearrangements, as proposed by Garau *et al.* [14], is not supported (but not ruled out) by the present calculations.

5.4   Reaction mechanism based on ES-2

The Michaelis complex. In contrast to **ES-1**, during the QM/MM simulations, the Zn(II) coordination polyhedron was fully maintained and conserved a regular tetrahedral geometry (Tab. SI 3). Remarkably, the ligand-Zn distances agree with the characteristic values obtained for a variety of Zn proteins [179]. In particular, **Wat2** remained within a distance of 2.1 Å from the Zn ion, keeping also its H-bond to **Bia**'s carboxylate. The ligand/protein interactions were also maintained (Tab. SI

3) and were similar to **ES-1**, except that His196 H-bonded not only to O(1)@**Bia** but, at times also to O(2)@**Bia**. Thus, the existence of Val67 and Lys224-substrate interactions are consistent, also in this case, with molecular biology data [175], which show the role of these two residues for the function of the enzyme.

Importantly, the β-lactam was kept in an suitable orientation for the nucleophilic attack by a H-bond between O(2)@**Bia** and **Wat2**.

5.4.1 Simulation of the enzymatic reaction.

The mechanism was investigated using as reaction coordinate the distance between O@**Wat1** and C(7)@**Bia** (starting value $d_{O@-C(7)}$=3.1 Å). All the bonds which characterized the Michaelis complex were maintained during the simulations performed at different values of RC, before reaching the TS. A change of the constraint force sign was observed, at $d_{O-C(7)}$=1.7 Å, as in **ES-1**, leading to TS formation. Conversely to the reaction starting from **ES-1**, the complete cleavage of the β-lactam moiety occurred in a single step, characterized by the following events in the sub-ps timescale (see Fig. 5.3): (i) deprotonation of the nucleophile by His118, which became di-protonated (this event was simultaneous to the nucleophilic attack); (ii) protonation of the nitrogen N(1) by **Wat2**, which became transiently a metal-bound hydroxide group, (iii) proton transfer from **Wat1** to the metal bound hydroxide, restoring the water molecule **Wat2** (Fig. 5.3); replacement of **Wat2** in **Wat1**'s position in **ES-2**; (iv) finally, a proton transfer from His118 in its charged state to the transient form of **Bia**. This led to the product of the enzymatic reaction (Fig. 5.3).

The calculated activation free energy for the entire concerted process was $\Delta F^{(\mathbf{Int})\ddagger} \sim$ 15 ± 3 kcal/mol, in good agreement with experimental findings [14].

The agreement with the experimental data allows us to discuss the available information on the functional role of key residues in the active site [175], in light of this novel mechanism. His118 has been shown to be an important residue for the reaction, as the k_{cat} of H118A decreases by three orders of magnitude relative to wild-type (wt) [175]. Consistently in our mechanism His118 acts as general base and it is therefore crucial for the function.

His196 is also believed to be a key residue, because it may create an oxyanion hole [14]. Somehow consistently with this hypothesis, our averaged D-RESP charges of O(1)@**Bia**, in both Michaelis complexes and the transition state ($-0.53\pm$ 0.04) are more negative than in simulation of the substrate in water ($-0.36\pm$ 0.04). The negative potential is created by a specific H-bond between His196 and O(1)@**Bia**.

A role of His196 has been more recently established by the decrease of k_{cat} by two order of magnitudes in the H196A mutant [175]. However, we have to point out that our calculations were not able to distinguish the effect of the residue on the transition state relative to the Michaelis complex, possibly because of the well known approximations in the calculations of the atomic charges.

The role of these two His residues for the catalysis has been recently further
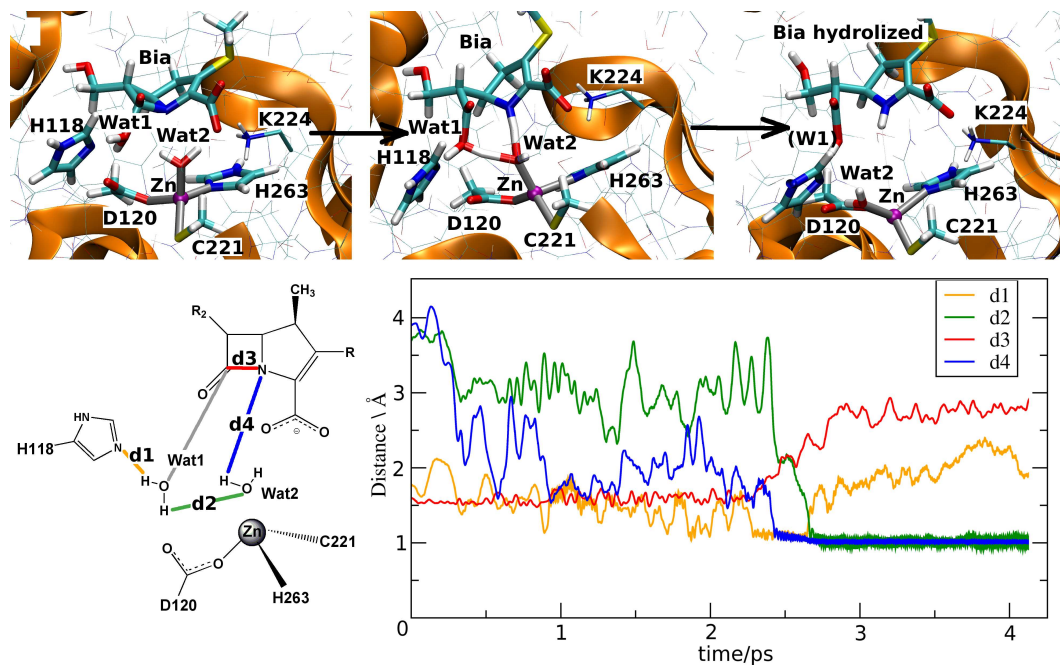
Figure 5.3: Investigation of **ES-2**. Here we show the time evolution of **ES-2** at $d_{O-C(7)}$=1.7 Å (**TS2**): in the top part of the figure, the movement of QM atoms (in bold licorice) along with Lys224 (thin licorice) is shown (snapshots were taken from the QM/MM constrained run at value of $d_{O-C(7)}$=1.7 Å). In the bottom part, key distances vs time evolution at **TS2** are plotted.

confirmed by another experimental study: a B1 enzyme (the so-called BCII from *Bacillus cereus* [180]) has been engineered so as to be similar to a mononuclear MβL, in which only one metal ion binds to the 'Zn2' site ('Zn2' MβLs), as found in the B2 subclass (Fig. SI 1). This was achieved by two mutations, H118S and H196S, and in the presence of only one equivalent of Zn^{2+}. The resulting impaired catalytic activities of these mutants suggests that residues His118 and His196, from the 'Zn1' site, play a key role for catalysis in 'Zn2' MβLs.

Two residues playing an important role for the *affinity* of the substrate have been identified by molecular biology data. The first is Lys224: in the K224Q mutant, the K_M measured for the carbapenem, was 10 times greater than in the wild type enzyme [175]. A structural basis for a role of the residue for binding is obviously provided by our model of the Michaelis complex. Indeed Lys224 forms a strong salt bridge with the protein (Fig. SI 3), as found for the B1 subclass ([181]. In addition, Val67 form hydrophobic interactions with the methyl group at position 4 of the substrate. At the speculative level, we may suggest that in the case of Ile, these might be affected, whilst with Asp they will be reduced.

We conclude that our mechanism provides a molecular basis for the role of all the residues which have shown to affect k_{cat} or K_M.

Part III

Directed evolution studies of B1 MβL BcII

Chapter 6

Directed evolution studies of B1 MβL BcII

6.1 Summary

Protein evolution is crucial for organismal adaptation and fitness. This process takes place by shaping a given three dimensional fold for its particular biochemical function within the metabolic requirements and constraints of the environment. The complex interplay between sequence, structure, functionality and stability that gives rise to a particular phenotype has hampered the identification of traits acquired through evolution. This is further complicated by the fact that mutations are pleiotropic, and interactions between mutations are not always understood. Antibiotic resistance mediated by β-lactamases represents an evolutionary paradigm in which organismal fitness depends on the catalytic efficiency of a single enzyme.

Here we have dissected the structural and mechanistic features acquired by an optimized metallo-β-lactamase obtained by directed evolution. We have solved the crystal structure of the evolved lactamase, and performed a kinetic, biophysical, mechanistic and in vivo study of this enzyme and related mutants, which allowed us to trace a general picture of molecular evolution of metallo-β-lactamases directly linked to organismal fitness, measured in terms of bacterial survival. We show that antibiotic resistance mediated by this enzyme is driven by two mutations with sign epistasis, which correlates to an in vitro synergistic effect of these mutations on the enzyme activity. One mutation stabilizes a catalytically relevant intermediate by fine tuning the position of one metal ion in the active site, while the other acts by augmenting the enzyme's flexibility. We found that enzyme evolution (and the associated antibiotic resistance) occurred at the expense of protein stability, revealing that metallo-β-lactamase have not exhausted their robustness threshold. Our results demonstrate that flexibility is an essential trait that can be acquired during evolution on stable protein scaffolds. The fact that metallo-β-lactamases can evolve by optimizing the reaction mechanism suggests that these enzymes possess a large evolutionary potential, in contrast with serine-β-lactamases. These results

suggest that these metalloenzymes have been recruited more recently by bacteria to supplement the catalytic profile of serine-β-lactamases. Directed evolution aided by a thorough characterization of the selected proteins can be successfully employed to predict future evolutionary events and to design inhibitors with an evolutionary perspective.

6.2   Introduction

Protein evolution is crucial for organismal adaptation and fitness. This process takes place by shaping a given three dimensional fold for its particular biochemical function within the metabolic requirements and constraints of the environment. The evolution of a protein relies on a subtle interplay between sequence, structure, functionality and stability, that is manifested in a particular phenotype. many studies rely on statistical description of adaptive molecular changes, but recent efforts have aimed at identifying changes in sequence and structure that affect fitness within a defined protein fold. The identification of such structural traits acquired through evolution allows both the reconstruction of ancestral genes and the prediction of future evolutionary events. The latter perspective could be immensely valuable for anticipating the molecular features that enhance bacterial resistance to antibiotics, and therefore designing strategies against this clinical threat. The main mechanism of bacterial resistance to β-lactam antibiotics is the expression of β-lactamases, a family of enzymes whose evolution is continuously challenged by the indiscriminate clinical use of antibiotics. MβLs and serine β-lactamases (sβLs) are not evolutionary related, their mechanisms of action are entirely different, and MβLs display a broader substrate spectrum, with catalytic efficiencies that are in general far from being diffusion limited. Since MβLs may have not yet achieved optimal fitness, the investigation of their future evolutionary potential is both appealing and extremely important clinically.

Directed evolution on the *Bacillus cereus* MβL, BcII, by DNA shuffling and selection by cephalexin (Fig. 1.10) resistance (a poor substrate of BcII) elicited an evolved variant, M5, displaying an enhanced cephalexinase activity. Surprisingly, this enhancement did not take place at the expense of impaired hydrolytic activity toward other substrate, thus generating an enzyme with broader substrate spectrum. The evolved enzyme M5 contains four mutations (N70S, V112A, L205S and G262S), all of them located outside the active site. We solved the X-ray structure of M5 (Table 6.1), which shows that none of these mutations entails changes in the global fold (0.4 Å RMSD on all non-hydrogen atoms *vs* WT BcII).

The active site of BcII can be described as a shallow groove flanked by loops L3 and L10, hosting two Zn(II) ions located at its floor (Fig. 6.1).

Zn1 is bound to 3 His residues (His116, His118 and His196), while Zn2 is coordinates to Asp120, Cys221 and His263. Despite all mutations are remote, Zn2 has moved closer to Zn1 by 0.7-1.0 Å in the mutant 6.2, while preserving the same ligand set. In the evolved variant M5, the two closest mutations to the active site are

Substrates	G262S			N70S			N70S/G262S			M5			BcII WT		
	k_{cat}	K_M	k_{cat}/K_M	k_{cat}	K_M	k_{cat}/K_M	k_{cat}	K_M	k_{cat}/K_M	k_{cat}	K_M	k_{cat}/K_M	k_{cat}	K_M	k_{cat}/K_M
Benzylpenicillin	320	150	2.13×10^6	543	¿3200	1.69×10^5	636	172	3.69×10^6	570	574	9.93×10^5	1020	662	1.54×10^6
Imipenem	40.1	227	1.76×10^5	ND	¿3000	1×10^5	93.2	63.1	1.47×10^6	129	721	1.79×10^5	279	687	4.06×10^5
Nitrocefin	680	50.2	1.36×10^7	1.73	14.5	1.19×10^6	460	30.2	1.52×10^7	81.3	27.8	2.92×10^6	30.1	9.82	3.07×10^6
Cefotaxime	87.3	60.2	1.45×10^6	115	113	1.01×10^6	182	32.6	5.58×10^6	44.9	30.3	1.49×10^6	80.5	43.3	1.86×10^6
Cephalexin	70.4	217	3.24×10^5	2.05	331	6.19×10^3	221	146	1.51×10^6	41.2	209	1.97×10^5	3.32	125	2.65×10^4

Table 6.1: Kinetic parameters for the hydrolysis of different β-lactam substrate by BcII wild type and BcII mutants: M5, G262S, N70S and N70S/G262S. units $k_{cat} (\text{sec}^{-1})$, K_M (μM) and k_{cat}/K_M (sec^{-1} M^{-1}).
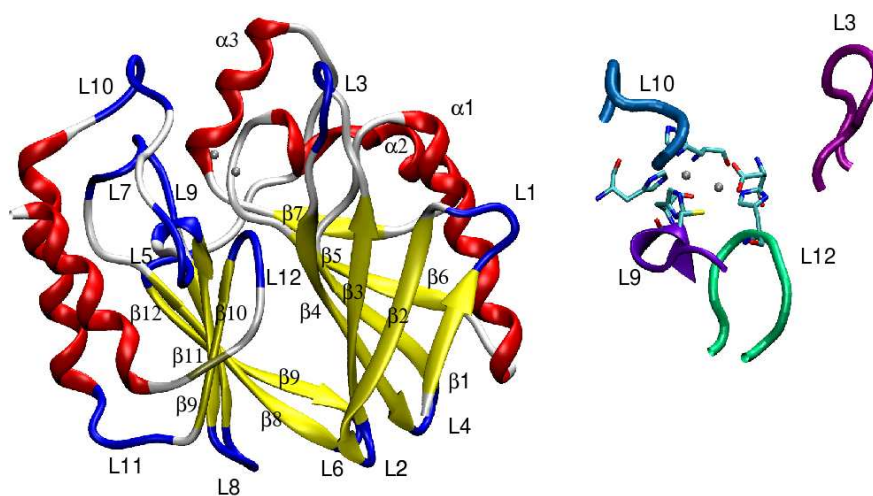

Figure 6.1: BcII secondary structure, and close up of the active site and of the loops which surrounded it.

N70S and G262S, both located below the active site floor. The hydroxyl oxygen of Ser262 is 3.1 Å from the Cys221 sulfur, and 3.48 Å from the Cys221 backbone nitrogen atom, resembling two interactions present in the plasmidic MβL IMP-1, in which a Ser residue is naturally found in this position.
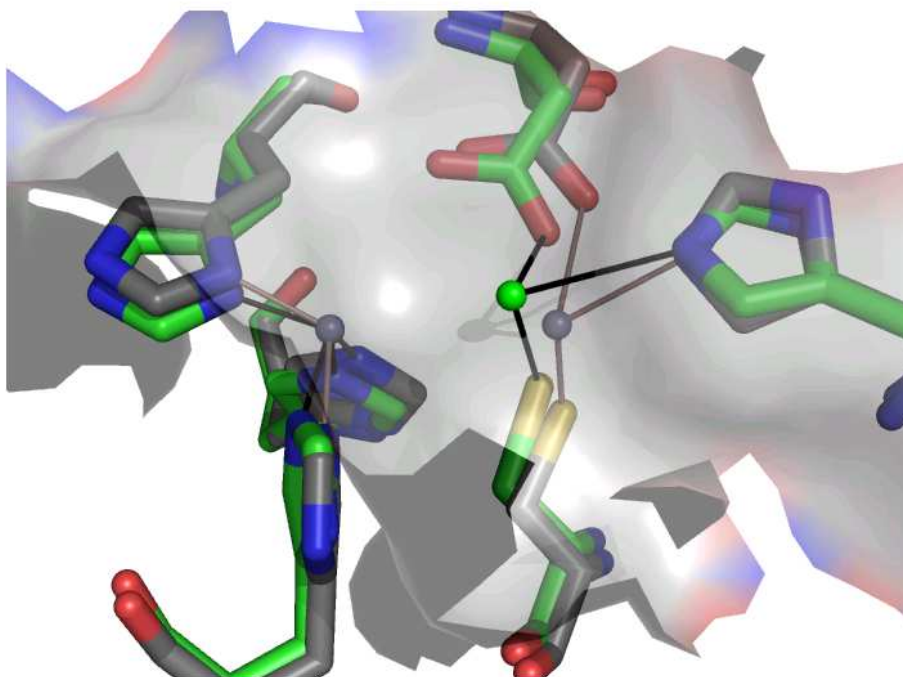

Figure 6.2: Comparison of the active sites of BcII WT (gray) and M5 mutant (green). The Zn2 ion (green) of the M5 mutant is more exposed that the WT one.

Ser262 thus behaves as a second shell ligand of "Zn2" and at the same time connects L10 and L12. Ser70 is located in a β-strand two residues upstream from Pro68, which defines the N-terminus of a conserved mobile loop (L3). The absence of the Asn70 sidechain results in the removal of two H-bonds with Pro261 and Gly264, which flank the "Zn2" ligand His263 and the mutated residue 262 (located in loop L12). These are the only H-bonds connecting loops L3 and L12. Another notable change in the interloop connections is that Arg121 (from loop L7) adopts a different conformation in M5, resulting in the loss of a H-bond of its guanidinium group with the backbone oxygen of residue 262 (loop L12).

6.3   Experimental characterization of mutants

[1]

In order to examine the influence of each of these mutations in the active site structure, mutants N70S, G262S and N70S/G262S were obtained, and the metal site was probed by substitution of the native Zn(II) ion by Co(II). The electronic spectra of the Co(II) derivatives (Fig. 6.3) show that the ligand field bands between 450 and 660 nm corresponding to the "Zn1" site are identical in all variants, revealing that this metal site is unaltered by either of these mutations. However, a Cys221-Co(II) charge transfer band at 343 nm (a reported of the "Zn2" site) is perturbed only when Ser262 is present, pointing to the key role of the S262-C221 H-bond.
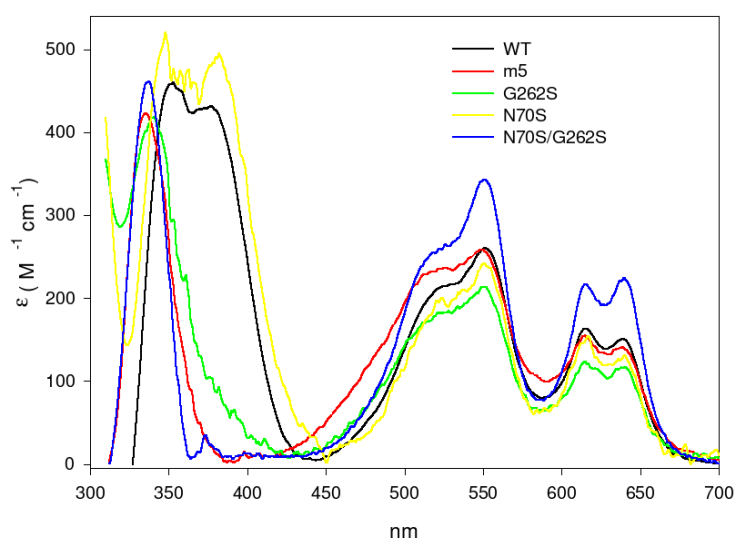
Figure 6.3: Electronic spectra of Co(II)-substituted BcII, M5, G262S, N70S and N70S/G262S in 10 mM Hepes (pH 7.5) 0.2 M NaCl. The enzyme concentrations were 250, 220, 180, 120 and 250 μM respectively.

These experiments reveal that (1) G262S is the only mutation that has direct effect on the metal site structure (via the latter H-bond), and (2) the Zn2 site is perturbed, while the "Zn1" site does now show any change upon this mutation. The catalytic performances of the single mutants N70S and G262S and the double mutant N70S/G262S were assayed with five different β-lactam substrates (Fig. 6.4).

The G262S mutation selectively enhances the catalytic efficiency toward ni-

---

[1]Most of the work on this research project was carried out experimentally. Therefore, it is reported in this thesis (in italics).
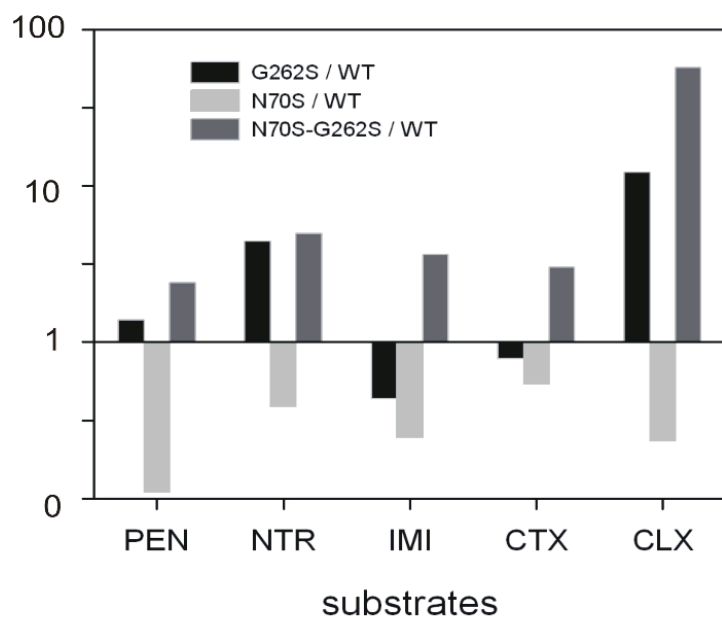
Figure 6.4: Relative enzyme activities of the G262S, N70S and N70S/G262S mutants compared to BcII WT against benzylpenicillin (PEN), nitrocefin (NTR), imipenem (IMI), cefotaxime (CTX) and cephalexin (CLX). Relative activity is the ratio of the mutant's k_{cat}/K_M value to that of WT. Y axis is in log scale.

trocefin and cephalexin (the substrate used in the directed evolution experiment), while the catalytic efficiency of the single mutant N70S is compromised for all substrates tested. However, the catalytic efficiency of the double mutant was enhanced for all substrates, with a notable 56-fold increase in activity towards cephalexin. This observation reveals a synergistic effect between these two mutations, suggesting that these two site are coupled. G262S, the only mutation that has a direct impact on the structure of the metal site, induces the greatest activity enhancement on cephalexin hydrolysis, driven by a 20-fold increase on k_{cat} (Tab. 6.1). This suggests that evolution may have acted by stabilizing the enzyme-transition state complex. This hypothesis was tested by studying the hydrolysis of nitrocefin, a chromophoric cephalosporin that is a useful probe of the catalytic mechanism of MβLs. Whereas nitrocefin hydrolysis by wt BcII under single turnover conditions proceeds through a direct conversion of substrate to product, the same reaction catalyzed by G262S BcII reveals the accumulation of an intermediate, as indicated by the rise of an absorption band with a maximum at 665 nm that decays within 10 msec (Fig. 6.5).
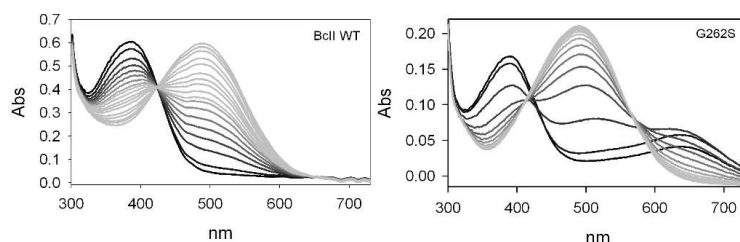

Figure 6.5: Relative enzyme activities of the G262S, N70S and N70S/G262S mutants compared to BcII WT against benzylpenicillin (PEN), nitrocefin (NTR), imipenem (IMI), cefotaxime (CTX) and cephalexin (CLX). Relative activity is the ratio of the mutant's k_{cat}/K_M value to that of WT. Y axis is in log scale.

This spectral feature is identical to that reported for nitrocefin hydrolysis by the MβL CcrA from B. fragilis, and attributed to an anionic intermediate in which a negatively charged nitrogen is bound to the metal ion at the "Zn2" site. In terms of steady-state parameters, the G262S mutation has increased k_{cat} from 30 to 680 s^{-1}, that is even larger than the one displayed by CcrA (250 s^{-1}). This confirms that the directed evolution has led to stabilization of a catalytically relevant intermediate, thus acting on the rate-limiting step of the reaction. The G262S mutation, resulting from a single base mutation, has already been observed in enzymes coded by transferable genetic elements in opportunistic and pathogenic bacteria: IMP-1 and its putative ancestor, IMP-6, differ only by this mutation, which expands the enzyme's substrate spectrum. This relatively recent event in the natural evolution in MβLs may have also acted by accelerating the rate-limiting step of the reaction, confirming the hypothesis that MβLs may not yet have reached their optimal

performance. Recent mutagenesis studies on BcII have shown that shifting the position of the "Zn2" site into a more buried location (and thereby also increasing the "Zn1"-"Zn2" distance) results in a severe impairment of the lactamase activity. "Zn2" is much more exposed in M5 than in WT BcII (Fig. 6.2), and its active site display the shortest metal-metal distance among all native and mutant enzymes from this family (3.16-3.39 Å). Thus, MβL-mediated resistance can still evolve further by improving the rate-limiting step through changes in the active site, and specifically the location of "Zn2". The data discussed above provide a structural explanation for the effect of the G262S mutation on catalysis, but do not account for the selection of the N70S mutation, nor explain the broader substrate spectrum displayed by the double N70S/G262S mutant. The N70S mutation is deleterious for the activity of WT BcII, but beneficial when acting on G262S BcII (Fig. 6.4), suggesting that it acts as a compensatory mutation. The crystal structure of M5 shows that the principal effect of removing the Asn70 side chain is the elimination of the two H-bonds connecting loops L3 and L12 (Fig. 6.1).

The spectroscopic data reveal that N70S does not alter the metal site structure, and therefore support the idea that this mutation acts exclusively on the loop structure and dynamics. Thus, the restoring effect of the N70S mutation is to act by enhancing loop flexibility. Compensatory mutations in evolution frequently act to restore protein stability. Guanidine hydrochloride-induced protein unfolding was therefore analyzed for the different enzymes. BcII is an extremely stable protein (ΔG_μ = 18.4 kcal/mol, see Fig. 6.6), and the G262S substitution has a large impact on its stability, reducing it by 13.9 kcal/mol. The N70S mutation causes a reduction of 5.6 kcal/mol, and the double mutant N70S/G262S is even less stable than either of the single mutants (ΔG_μ = 3.3 kcal/mol). It is clear that although N70S serves to restore a broad substrate profile and enhance activity for all substrate tested, it does not compensate for the loss of protein stability.

In SβLs, mutations that enhance enzyme activity and destabilize the protein are compensated by stabilizing mutations (e.g. M182T in TEM-1). Thus, the fitness landscape of these enzymes is determined by a trade-off between the activity and stability features introduced by the mutations. MβLs present a different picture: the compensatory mutation (N70S) does not stabilize, in fact it further destabilize, the protein and instead facilitates greater flexibility in the loops flanking the substrate binding cavity. The exceptionally high stability of the MβL fold compared with SβLs (ΔG_μ for the TEM-1 is 10 kcal/mol) is such that there may be no need yet to compensate the destabilizing effects brought upon the introduced mutations. This suggests that robustness does not play a central role in the fitness landscape of MβL evolution. Instead, the impact of the restoring mutation on loop flexibility suggests that a coupled network of motions would be operative in MβLs-mediated catalysis, that could be of importance for the evolution of these enzymes given their broad substrate spectrum. Recently, Park and coworkers succeeded in eliciting MβL activity in the scaffold of glyoxalase II. This variant is still 4000-fold less active that a native MβL. At the light of the present results, loop engineering aimed to increase the flexibility may be exploited to further enhance the lactamase
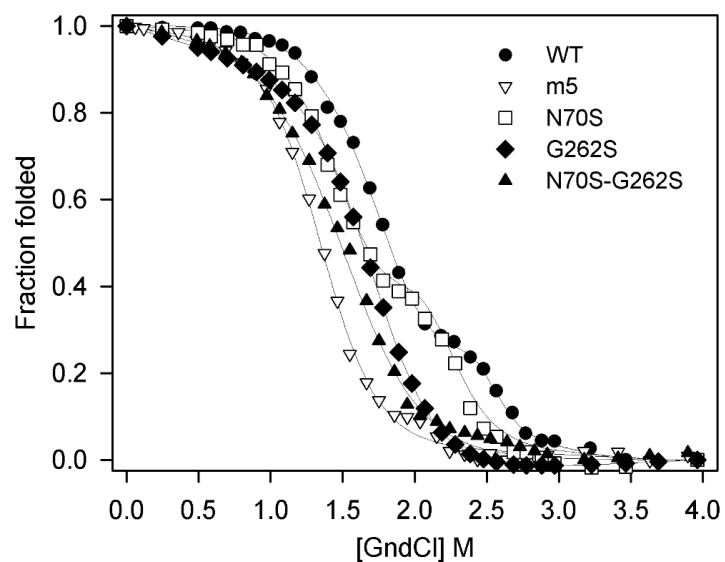
Figure 6.6: Equilibrium denaturation curves of wild type BcII, m5, G262, N70S and N70S/G262S in 50 mM Hepes (pH 7.5) 0.2 M NaCl at different GdnCl concentrations monitored by fluorescence spectroscopy. The theoretical curves are drawn by usgin a three state model (WT and N70S) and two-state model (M5, G262S and N70S-G262S).

activity in that case. Protein stability is not necessarily associated with evolutionary fitness: indeed, optimum protein stability may fall within a narrow range with highly stable protein lacking the required flexibility to meet their functional needs. The MβL studied here, BcII, is a remarkably stable protein with scope for enhancing its activity at the expense of thermodynamic stability. Thus, in this case, fitness takes place at the expenses of loosing thermodynamic stability, with a clear epistatic effect caused by the N70S mutation by augmenting the flexibility of the active site loops. MβLs thus follow a different evolutionary paradigm compared to SβLs, and clearly represent an alarming clinical threat.

MβLs are a relatively recent response to antibiotic usage, and there is limited information on mutations that increase resistance. Nevertheless, one of the mutations generated in this study has already been found in nature (G262S in IMP-1), demonstrating that studies such as this are capable of predicting the future course of bacterial evolution of resistance. In the present case, the effect of this mutation on the rate-limiting step of the reaction allows us to propose a strategy for designing a mechanism-based inhibitors with an evolutionary perspective. Anionic transition state analogues targeted to bind the "Zn2" site are expected to be good inhibitors, instead of ligands aimed to fit into an oxyanion hole (a common strategy for designing hydrolase inhibitors, that has failed for MβLs).

6.4   Discussion and conclusions

G262S mutant simulation results [2] A were fully consistent with experimental evidence of the S262-C221 H-bond in the mutants; in addition , they pointed to a larger flexibility of loop L3 in the mutants relative to WT (Fig. 6.7). Loop L10 instead, is less flexible in G262S BcII that in WT, and it shows the largest mobility in the double mutant. The double mutant's active site is wider than the other two. The simulations of WT and mutant Michaelis complexes showed that both loop are overall less flexible that the corresponding free states. They close over the substrate, getting close to each other. Such conformational change involves the formation of an electrostatic interaction (weaker in WT relative to the mutants) between the carboxylate moiety of the substrate and the conserved lys224. loop L10 and l3 are more flexible in the double mutants than in the other two 6.7). These simulations together with the structural, functional and mechanistic data previously discussed, allow us to postulate the folowing scenario for MβL evolution. The G262S mutation is responsible for the enhanced cephalexinase activity through a two-fold effect: the H-bond between Ser262 and Cys221 not only allows loop L10 to adopt a less open conformation upon cephalexin binding, but it also leads to a better positioning of the "Zn2" ion. "Zn2" has been suggested to play an essential role in substrate binding and orientation (in concert with Lys224), and in favoring C-N bond cleavage through the stabilization of a negative charge on the β-lactam nitrogen. While enhancing cephalexin hydrolysis, this relative tightening of L10

---

[2]Detailed results of the MD simulations performed on this subject are reportred in Appendix A

with respect to L12, does not favor penicillin, cefotaxime and imipenem hydrolysis
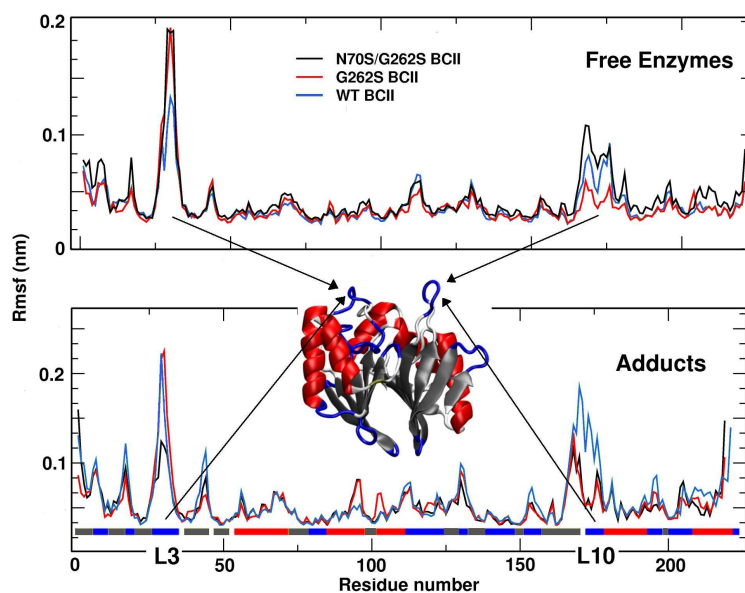(Fig. 6.1).

Figure 6.7: Root Mean Square fluctuation of the Cα from MD simulations of the
free enzymes and adducts.

Instead, the N70S mutation eliminates two H-bonds connecting L3 and L12,
giving rise to a larger active site groove and thus compensating for the effect of
G262S and rendering an enzyme with a broader substrate spectrum (and even fur-
ther enhanced activity towards cephalexin).

Part IV

Conclusions

Chapter 7

Conclusions

MβLs counteract the beneficial action of β-lactam antibiotics against bacterial infections by enzymatically hydrolyzing their β-lactam rings.

In this thesis I have addressed important issues regarding the mechanism of resistance, by computational methods.

Most efforts have been devoted to dissecting analogies and differences among the reaction mechanisms in two different subclasses of these enzymes. This information may be of relevance to design ligands able to inhibit more than one MβL. I have focused on one subclass, the B2 MβL, and compared my results with those obtained for the B1 subclass, for which some mechanistic insights had already been obtained in our lab. My calculations were based on the CphA B2 MβL from *Aeromonas hydrophila*, for which an X-ray structure has been solved [14].

A key issue and prerequisite to investigate the reaction is the exact knowledge of the active site residues protonation state [83].

We have used a two-steps strategy to address this issue: first we have evaluated structural determinants and energetics of active site models by DFT calculations, and then we have constructed the Michaelis complex.

The stability of eleven protomers was analyzed considering their energetics and RMSD values relative to the X-ray. Our calculations suggested that two protomers (**A** and **B** Fig. 7.1) may be present for the unbound enzyme, which could be reasonably assumed to be in equilibrium. Protomer **A** was characterized by the presence of a Zn-coordinated water molecule. One of the second shell ligands, His118, was protonated in Nε. Protomer **B** was characterized by a hydroxide bound to Zn. His118 was diprotonated.

Docking procedures along with MD simulations were performed to provide a structural model of the Michaelis complex with Biapenem substrate. Two binding modes were obtained (**ES-1** and **ES-2**). The nucleophile molecule was located in a pocket between His118 and Asp120. The carboxylate group of the antibiotic interacts directly with the zinc in one binding mode, and with a metal-bound water in the other one. The Michaelis complex featured only protonation state (**A**). Thus, the presence of the substrate affected the protonation states of the enzyme. Of

these, **ES-1** is that proposed experimentally based on the X-ray structure [14] (see Fig. 7.2).

Hybrid QM/MM simulations, based on the two binding modes, reported in this work, support a novel mechanistic view, different from the one proposed by Garau and partially addressed in the other computational work on CphA [97]. Based on our calculations, we suggest that the reaction proceeds through a previously unrecognized single-step concerted mechanism. In our proposed pathway a water molecule performs the nucleophilic attack to the β-lactam carbonyl and a second Zn-bound water protonates the β-lactam nitrogen. The reaction was characterized by a free energy barrier of about 15 kcal/mol, in agreement with experimental data [14]. Our mechanism provides a structural role for all of the residues which have been shown to be important for the function of B2 MβLs. These include i) His118, which orients the catalytic water and acts as generalized base; ii) His196 which creates an oxyanion hole, as proposed by Garau *et al.* [14]; iii) Val67 and Lys224, which affect substrate binding. Furthermore the calculations point to the key roles of the metal ion in the "Zn2" site and of two water molecules for the hydrolysis, in agreements with recent experimental [182, 183, 184] and theoretical [23] findings.

We now come to our comparison with hydrolysis mechanism of the others subclasses. We suggest that the "Zn2" site, conserved in both B1 and B2 MβLs, is able to indirectly contribute to substrate binding by a water molecule which H-bonds the antibiotic's carboxylate group (Fig. 7.3). Moreover, the metal in the "Zn2" site is able to activate a water molecule as a proton donor for the completion of the antibiotics hydrolysis. In addition, both B1 and B2 MβLs are suggested to proceeds through a single step mechaism. At a speculative level, we suggest that some aspects of this mechanism might be shared also by the B3 MβL members
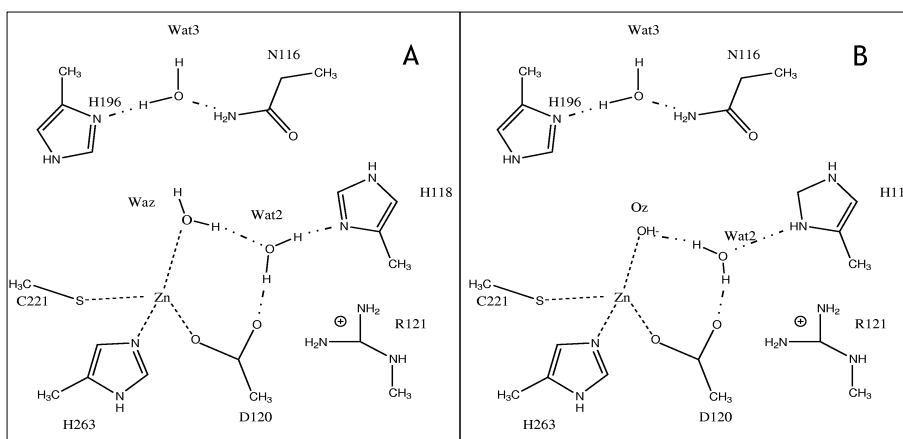

Figure 7.1: Active site protomers of **CphA** emerging from our DFT-BLYP calculations. The 4th Zn ligand is a water molecule (Waz) in **A**, and an OH group (Oz) in **B**.
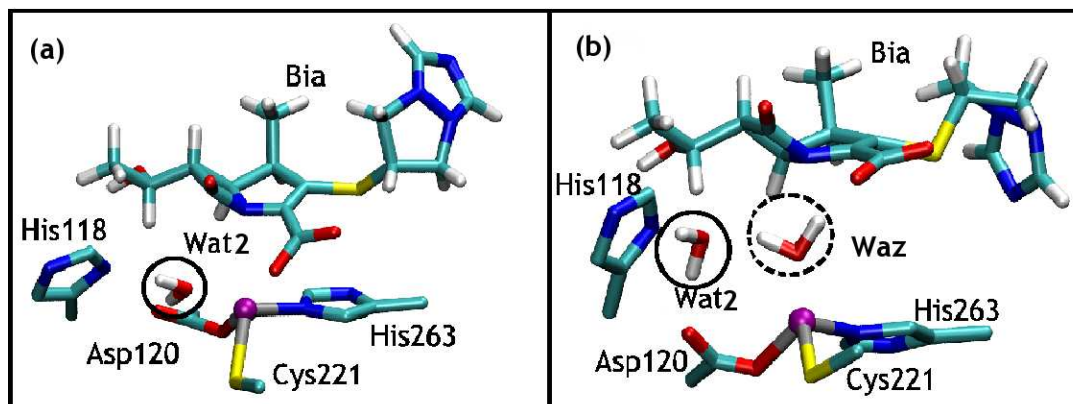
Figure 7.2: **ES-1A** (a) and **ES-2A** (b) MD average structures. The putative water molecules (Wat2) acting as nucleophile in the enzymatic reaction are highlighted with circles; the second zinc coordinated water (Waz) is indicated with dashed circle in (b).

[10], for the high similarity in the Zn content in the active site, along with the composition of Zn ligands. We hope that our results may help to design of novel inhibitors capable to counteract the action of more than one MβLs subclass.

In the second part of the thesis, I have focused on the exploration of possible pathway of evolution of a B1 MβL, the BcII from *Bacillus cereus*. Directed evolution on these bacteria yielded optimized variants able to confer more resistance than native enzyme. These variants harbor four mutations remote from the active site that contribute to broaden its substrate repertoire.

Our MD calculations on these mutants showed that G262S modifies the Zn. This of course may affect the geometry of the second shell ligands. QM/MM calculations could better address this issue. N70S-G262S cause an increase of size and flexibility of the active sive cavity. This provides a rationale for the larger spectrum of substrate hydrolized by the enzyme. It is also our hope that new ligand inhibitors for B1 BcII, may take into account the structural role of these mutants [185].

Published or submitted articles

"Protonation state and substrate binding to B2 metallo-beta-lactamase CphA from Aeromonas hydrofila", Fabio Simona, and Alessandra Magistrato, Mariano D. Vera, and Gianpiero Garau, and Alejandro J. Vila, and Paolo Carloni, Proteins 2007, 69, 3.

" A common catalytic architecture for zinc in B1 and B2 metallo β-lactamases: insights from the reaction mechanism of CphA enzyme", Fabio Simona, Alesandra Magistrato, Matteo Dal Peraro, Alejandro J. Vila and Paolo Carloni, JACS,
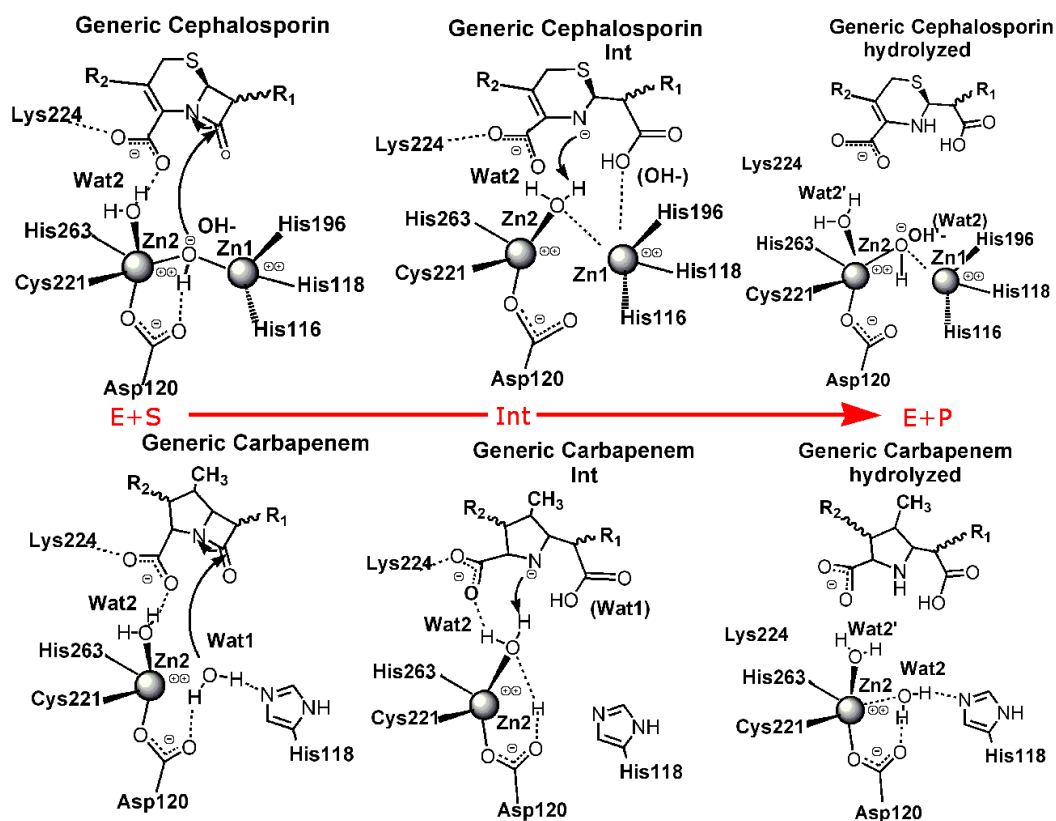
Figure 7.3: Comparison between the reaction mechanism proposed for B1 (on the top panel) and that proposed for B2 MβL (on the bottom part), in which the key role of the 'Zn2' binding site in both subclasses emerges clearly. The metal loaded in this site is involved, almost simultaneously, either in nucleophilic attack of the β-lactam ring and in the protonation of the β-lactam nitrogen.

submitted.

" Adaptive protein evolution grants organismal fitness by improving catalysis and flexibility " Pablo E. Tomatis, Stella M. Fabiane, Fabio Simona, Paolo Carloni Brian J. Sutton and Alejandro J. Vila, PNAS *accepted.*

Part V

Appendix

Appendix A

Details of MD simulation of WT, G262, N70S-G262S BcII mutants.

A.1 Methods and general comments

The structure of di-Zn(II) BcII from Bacillus cereus (PDB entry code 1BC2,) was selected as the reference starting structure for all our calculations. Loop L3, which is not present in the corresponding coordinates file, was reconstructed and modeled as already reported . The starting structures for G262S BcII and N70S-G262S BcII were obtained by superimposing the backbone atoms of the serine residues over those of the corresponding wild type residues. The resulting $\chi 1$ and $\chi 2$ angles were verified to fall in the most favorable regions. Each system was immersed in a box (85.6 Å x 80.3 Å x 85.9 Å) of water molecules (approximately 17,000), with 4 chloride anions for neutralizing the excess charge. The force field parameters for the protein frame, the counterions and for water are those of the AMBER PARM98 citeAMBER7 and TIP3P , respectively. For the coordination sphere of the Zn(II) ions, we followed the same scheme as for CphA. For the antibiotics, the gaff force field was used , as for the charges. Electrostatic interactions were computed using the Particle Mesh Ewald (PME) algorithm . A cutoff of 12 Å was used for van der Waals and for short range component of electrostatic interactions. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm . An integration time step of 1.5 fs was used. Loop L3 and the sidechains of mutations were first energy minimized using a steepest descent algorithm. The systems were initially relaxed with 5,000 steps of optimization of the solvent, while the protein was held fixed, followed by 5000 steps of optimization of the whole system. The systems were slowly heated to 300 K performing runs at constant volume and constant temperature using Langevin Dynamics . Constant temperature and pressure (1 atm, 300 K) MD runs were finally performed . The pressure coupling was accomplished with Langevin piston and the temperature with the Nosé-Hoover method . 10 ns of trajectory were produced for every systems, with the aim of NAMD programs .Subsequently, Cephalexin (Cfa) was docked inside the last MD snap-
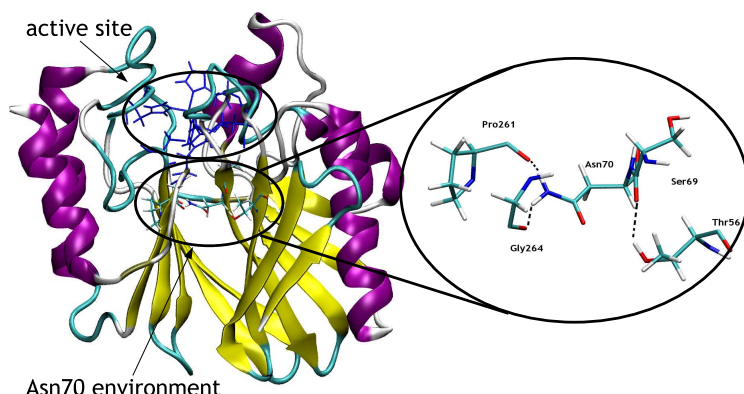
101

Figure A.1: Close up of the Asn70 environment: all its Hbond are outlined

shots of the three structures. The three Michaelis complexes underwent the same computational protocol as the free states.

It turned out that the geometries of active site residues in all the simulations, were stable and well conserved (Tab. A.1 and A.2), as well as the fold.

A.2   Free enzymes

A.2.1   Wild type BcII

N70 kept the two H-bonds (NH$_2$ group of sidechain) with CO@P261 (d^{av}_{O-H} = 1.92 ± 0.17 Å, Θ_{D-H-A} 162 ± 8 deg), and CO@G264 (d^{av}_{O-H} = 1.82 ± 0.08 Å, Θ_{D-H-A} 163 ± 8 deg), during all the simulation, while H-bond with Thr56 was lost and it was replaced by another between OH@Th56 with CO@Ser69 (d^{av}_{O-H} = 2.12 ± 0.3 Å, Θ_{D-H-A} 153 ± 10 deg) (Fig. A.1).

A.2.2   G262S

The starting structure of the OH@S262 (Fig. A.3), was characterized by the H-bond with Oδ_2@D120 (H...O distance of ≈ 2 Å, angle between acceptor-H-donor ≈ 156 degree).

This contact broke at 4.9 ns, and was replaced by the OH@S262-S@C221 interaction, that remained until end of this simulation (7.5 ns) (d_{H-S} 2.3 ± 0.1 Å, angles Θ_{OH-S} 143.9^o ± 8.6).
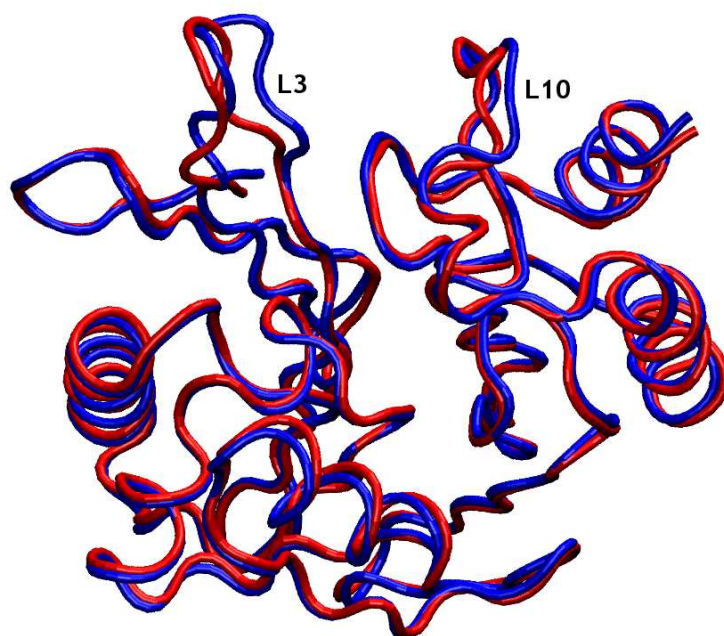
Figure A.2: Comparison between initial (red) and final (blue) Cα trace in WT simulation.
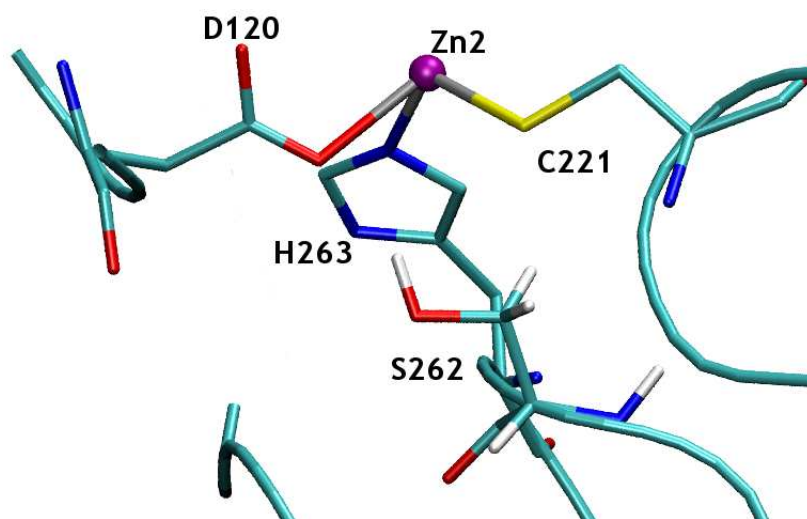

Figure A.3: Starting structure of BC2 G262S, with the contact OH@S262 Oδ2@D120
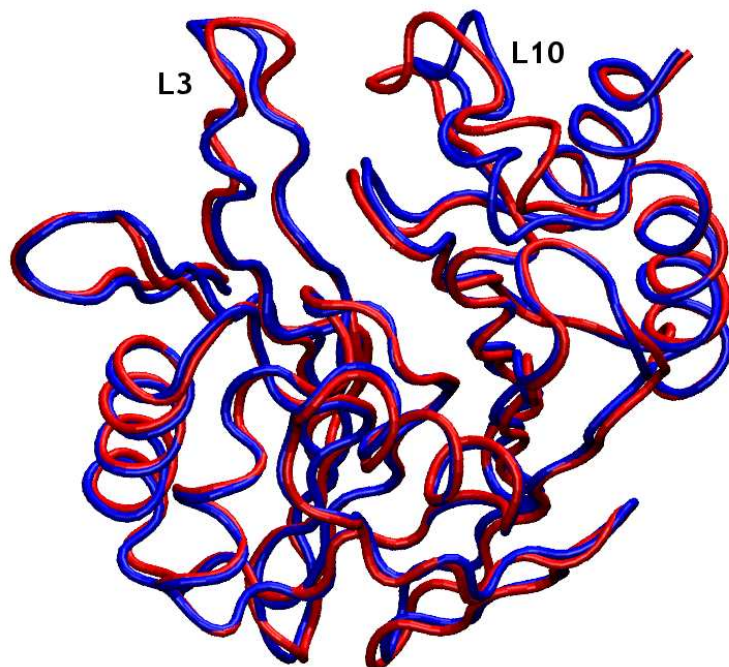
Figure A.4: Starting structure (only backbone trace, in blue) and final (red) one.

Loop L12 is located near the base of loop L3 (6.1). The two loops interacted through an H-bond between O@Pro68 and NδH@His263, and this interaction was maintained until 6 ns, when it was broken in favor of formation of a new interaction with NH@Gly264.

Asn70, in the WT BcII was connected to Loop L12 with two interloop H-bonds of the NH$_2$ group, one with Pro202 and a second with Gly205. Moreover, the CO group of Asn70 was involved in a intragroup interaction with Thr56, a residue located in β3-strand, placed in close proximity of loop L3. The interaction between Pro202 and Thr56 was lost, while the one with Gly205 was maintained: Asn70 mediated interaction between the two loops, L3 and L12.

All these events led to a movement of L3 (roto-translation) which put it toward L10 (Fig. A.5).

A.2.3   N70S-G262S

The starting structure, showed in Fig. A.6, was characterized by the presence of interaction between L3-L12, achieved through H-bond between OH@Ser70 and Pro261 (Fig. A.6)

The contact between OH@S262 and S@C221 was present for about 80% of the total simulated time (average d$_{S-OH}$ 2.19 Å $\pm$0.19 Å, average angle Θ_{OH-S} 154deg $\pm$ 2deg), and in the remaining time it was replaced by a second H-bond,
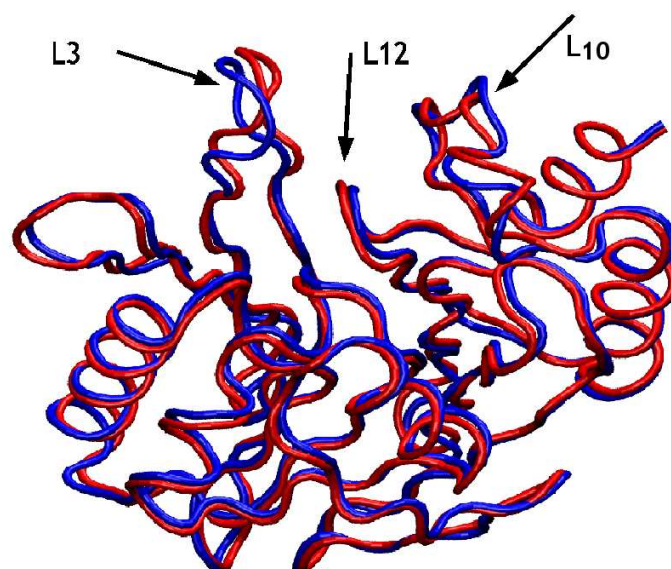
Figure A.5: Initial and final structure (only backbone is showed) of G262S: as could be seen only the three loops L3, L10 and to a less degree L12, were significantly different.
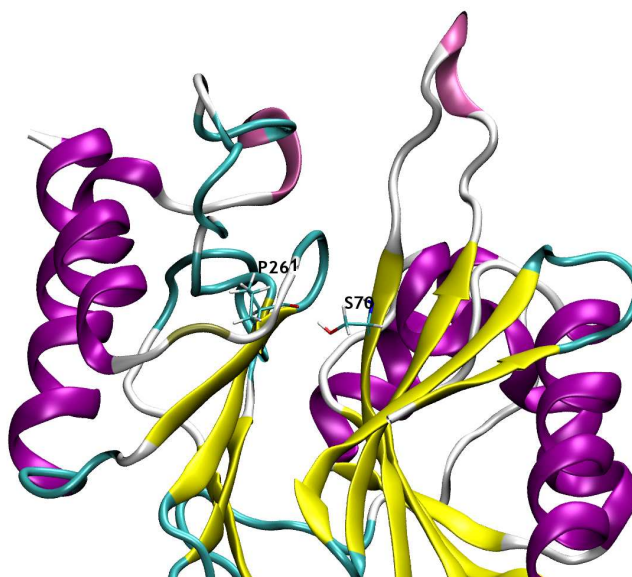

Figure A.6: Starting structure of G262S G70S: the OH group of serine Hbonds with CO backbone of Pro261.
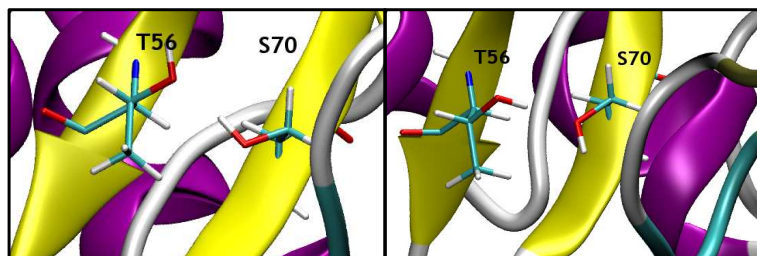
Figure A.7: Intra-structure interaction between Ser70 and Thr56, both located at the base of loop L3

namely between OH@C221 with Oδ2@D120.

Asn70-Pro261 interaction was observed only in the first half of the total simulation, and was lost in the second half, while another interaction between Ser70 and Thr56 (Fig. A.7) was formed: a crucial interaction connecting L3/L12 was lost, while a second new interaction *intra*structure was formed.

This compensate the conformational changes induced by the G262, and on the contrary resulted in an increased degree of mobility of loop L3, which in fact is the opposite effects of what observed in the case of G262S single mutant).

A.3 Simulation of Cfa complexes

A.3.1 WT

In order to evaluate the stability of the antibiotics in the catalytic pocket of the enzyme, the distance between the N1@Cfa from Zn2 was used. The distance was stale during all the MD simulation (d_{N1-Zn2} 3.4 $\pm$ 0.1 Å).

Eventual changes in the orientation of the plane of β-lactam ring in the catalytic pocket were evaluated monitoring the angle Θ between Zn2-N1@Cfa-C6@Cfa ($\Theta_{Zn2-N1-C6}$ 111 $\pm$ 6 deg) and another one defined ad Θ Zn2-N1@Cfa-C8@Cfa ($\Theta_{Zn2-N1-C8}$ 91 $\pm$ 6 deg), and both turned out to be stable during the simulated time.

The rotation of the β-lactam plane respect the line Zn1-Zn2 was followed ddmeasuring the dihedral evolution angle defined by the atoms Zn1-Zn2-N1@Cfa-C6@Cfa (χ = -94 $\pm$ 5).

Zn2-O2@Cfa was stable (d_{aver}=1.95 Å $\pm$ 0.08Å), the salt bridge NH@Ala119 and CO@Cfa present for 91 % of the simulation ($d_{averH-O}$=1.99 Å $\pm$ 0.15 Å, Θ_{D_A}=157.3 $\pm$ 11.7): the breaking of this bond was connected to the decrease of χ mentioned before (the plane of β-lactam rotates for a moment respect to Zn1-Zn2 line). Hydrophobic interactions with C2@Cfa and Methyl@Cfa (R1 substitute group): at first Methyl@Cfa interacts only with the backbone of Val67, and Phe62 was far from C2@Cfa.
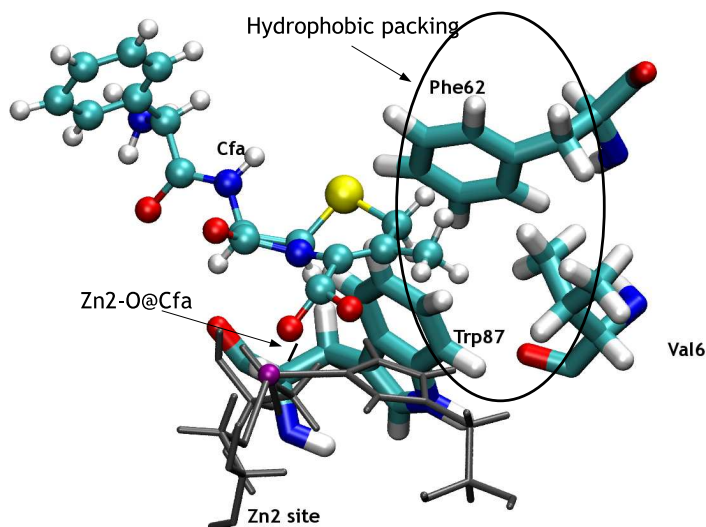
Figure A.8: Close up of the ending structure of simulation. Hydrophobic interaction were outlined in the circle.

There were little differences in the hydrophobic interaction pattern from N70S-G262S mutant, mainly contacts involved with Val67 (Fig. A.8).

Different tridimensional position of the loop L3, could be connected to the Hbond between $N\delta1H@H263$ and CO@Val67 (this helped to keep Val67 close to Cfa) (present 76 % of time, and when present $d_{averH-O}$=2.08 Å $\pm$ 0.18 Å, Θ_{D_A}=150.2 $\pm$ 10.5).

A.3.2 G262S

Also in this simulation it turned out that the Cfa was well trapped in the active site, and any significant fluctuation in the four parameter chosen above were observed during all the dynamic ($d_{N1-Zn2} = 3.52 \pm 0.15$, $\Theta_{Zn2-N1-C6} = 111 \pm 6$ deg, $\Theta_{Zn2-N1-C8} = 90 \pm 6$ deg, $\chi = -93 \pm 5$)

Zn2-O2@Cfa was stable (d_{aver}=1.94 Å $\pm$ 0.08Å), as well as the salt bridge between NH_3@Lys224-O3@Cfa ($d_{averN-O}$=2.85 Å $\pm$ 0.15 Å).

H-bond between NH@Ala119 C8@Cfa was always present during the simulation ($d_{averH-O}$=2.07 Å $\pm$ 0.22 Å, Θ_{D_A}=154.2 $\pm$ 12.5).

Hydrophobic interactions of Methyl@Cfa (R1 substitute) and of C2@Cfa: they were quite similar of those of N70s-G262S mutant, but here Phe62 lost its hydrophobic contact with Cfa at the end of simulation (Fig. A.9 and Fig.A.10).

Other key interprotein interactions were:

S@C68-OH@Ser262 contact, present always in this simulation, while Asn70 interact with the Thr56 and Pro261 (Hbond present 100% : $d_{averH-O}$=2.02 $\pm$ 0.28,
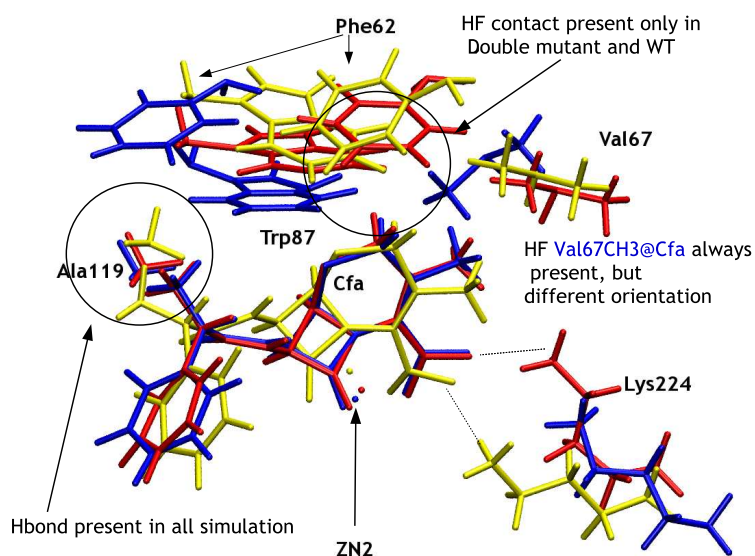
Figure A.9: Docking pattern compared between three simulations: in blue the structure of single mutant, in red the one of the N70S-G262S mutant and in yellow those of WT (only sidechains of protein residues), the different position of residues involved in the hydrophobic packing could be seen.
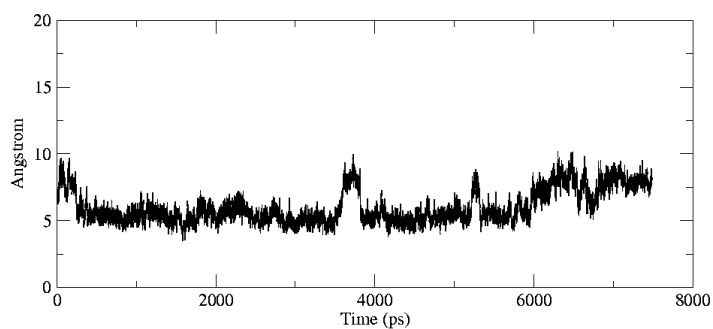
Figure A.10: Distance between one of the C of the phenyl ring of Phe62, and C8@Cfa. the distance decreased with time, and at the end Phe was not involved in hydrophobic packing of the Cfa.

Figure A.11: The three distances of loops, compared in the three simulation of Cfa adducts: as could be seen in that case the effect is the opposite of what seen in free enzymes simulations, where the maximum opening of loops L3 and L10 was observed for Double mutant (as a consequence of its increased mobility due to mutation in position 70). In that case the increased mobility allow the loops (in particular L3) to fit better and docking in a more efficent way the Cfa.
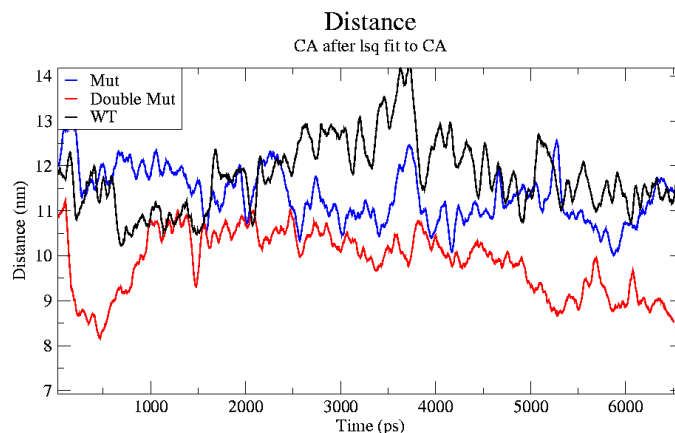
Θ_{D_A}=152.57 $\pm$ 13.3 for the one with Pro, and d$_{averH-O}$=2.05 $\pm$ 0.28, Θ_{D_A}=161.64 $\pm$ 11.2 for the one with Thr).

In that simulation O@Pro68 H-bonds always with NδH@H263 (d$_{averH-O}$=1.90 $\pm$ 0.16, Θ_{D_A}=151.72 $\pm$ 12.6).

All these effects are connected, and caused a different orientation of loop L3 (see Fig. A.11).

A.3.3   N70S-G262S

d$_{N1-Zn2}$ was stale during all the MD simulation (d$_{N1-Zn2}$ = 3.5 $\pm$ 0.15 Å), as well as the other selected parameter ($\Theta_{Zn2-N1-C6}$ 111 $\pm$ 6 deg, $\Theta_{Zn2-N1-C8}$ 88 $\pm$ 6 deg, χ = -91 $\pm$ 5)

Other interactions were the following:

i) Zn2-O2@Cfa (d$_{av}$ = 1.94 $\pm$ 0.01 Å),

ii) the salt bridge between NH$_3$@Lys224...O3@Cfa (d$_{avN-O}$ = 2.8 $\pm$ 0.2),

iii) H-bonds between NH@Ala119 and C8@Cfa (d$_{averH-O}$=2.1 $\pm$ 0.3, Θ_{D_A}=148 $\pm$ 14).

iv) Hydrophobic interaction of Methyl@Cfa (R1 substitute in Cfa) with Gly232 (average distance between C2@Cfa and Cα@Gly232 4.16 Å $\pm$ 0.45) and Val67 (average distance between C@Cfa and Cβ@Val=4.69 $\pm$0.045). Another interaction was the one between phenyl group @Phe62 and C2@Cfa (av. dist. 4.05 $\pm$
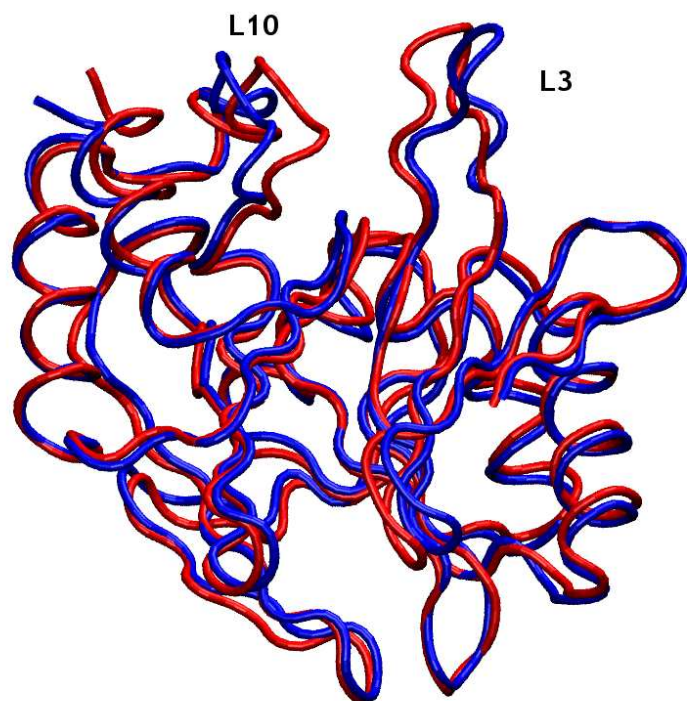
Figure A.12: Initial (blue) and final (red) backbone trace.

0.67) and with Trp87. All interactions formed at the beginning of simulation, and were stable during the dynamic.

The HO@Ser262-S@Ser221 contact, present from the beginning of the simulation, was lost in the last 2 ns, in favor of a H-bond with $O\delta2$@Asp84, while Ser70 didn't form any stable interaction, but many at different times, with the three residues, and also with Ser69 and with its own C=O backbone atom.

	force field	Bc2 WT	Bc2 G262S	BC2 G262S N70S
ZN1 ZN2	.3.85	3.22 ± 0.02	3.18 ± 0.02	3.15 ± 0.02
ND@H88-Z1-NE@H49	101.22	95.5 ± 1.8	97.31 ± 2.8	84.9 ± 1
NE@H86-Z1-ND@H88	101.22	98.12 ± 1.8	101.09 ± 1.6	97.02 ± 0.8
NE@H86-Z1-NE@H49	107.95	96.22 ± 3.7	94.60 ± 1.1	96.22 ± 0.7
O2@D90-Z2-S@C68	107.23	102.408 ± 1.0	100.0 ± 1.8 *	103.3 ± 0.8
O2@D90-Z2-NE@H21	73.96	67.88 ± 0.4	68.78 ± 0.6	68.65 ± 0.74
S@C68Z2-NE@H21	107.23	114.5 ± 0.8	117.19 ± 0.7	117.21 ± 2.1

Table A.1: Structural parameter for the active sites, simulation of free enzyme

	force field	Bc2 WT-Cfa	Bc2 G262S-Cfa	BC2 G262S N70S-Cfa
ZN1 ZN2	.3.85	3.14 ± 0.09	3.15 ± 0.10	3.22 ± 0.10
ND@H88-Z1-NE@H49	101.22	88.14 ± 4.8	88.93 ± 5.3	97.82 ± 5.5
NE@H86-Z1-ND@H88	101.22	97.3 ± 4.3	95.84 ± 4.3	98.52 ± 5.0
NE@H86-Z1-NE@H49	107.95	96.63 ± 4.3	97.18 ± 4.5	97.03 ± 4.4
O2@D90-Z2-S@C68	107.23	103.93 ± 3.1	104.24 ± 3.2 *	105.54 ± 3.4
O2@D90-Z2-NE@H21	73.96	67.18 ± 2.7	68.00 ± 2.8	67.06 ± 3.6
S@C68Z2-NE@H21	107.23	103.93 ± 3.1	104.24 ± 3.2	114.7 ± 3.3

Table A.2: Structural parameter for the active sites, simulation of adducts

Bibliography

[1] C Foucault and P Brouqui. How to fight antimicrobial resistance. *FEMS Immunol Med Microbiol*, 49(2):173–183, Mar 2007.

[2] D M Livermore. The need for new antibiotics. *Clin Microbiol Infect*, 10 Suppl 4:1–9, Nov 2004.

[3] A R Sapkota, L Y Lefferts, S McKenzie, and P Walker. What do we feed to food-production animals? a review of animal feed ingredients and their potential impacts on human health. *Environ Health Perspect*, 115(5):663–670, May 2007.

[4] L Soulsby. Antimicrobials and animal health: a fascinating nexus. *J Antimicrob Chemother*, 60 Suppl 1:77–78, Aug 2007.

[5] J M Siegford, W Powers, and H G Grimes-Casey. Environmental aspects of ethical animal production. *Poult Sci*, 87(2):380–386, Feb 2008.

[6] L Tollefson and M A Miller. Antibiotic use in food animals: controlling the human health impact. *J AOAC Int*, 83(2):245–254, Mar-Apr 2000.

[7] J F Fisher, S O Meroueh, and S Mobashery. Bacterial resistance to beta-lactam antibiotics: compelling opportunism, compelling opportunity. *Chem Rev*, 105(2):395–424, Feb 2005.

[8] C. Walsh. *Antibiotics. Action, origins, resistance*. ASM press, Washington, D.C., Washington, D.C., 1 edition, 2003.

[9] M W Crowder, J Spencer, and A J Vila. Metallo-beta-lactamases: novel weaponry for antibiotic resistance in bacteria. *Acc Chem Res*, 39(10):721–728, Oct 2006.

[10] C Bebrone. Metallo-beta-lactamases (classification, activity, genetic organization, structure, zinc coordination) and their superfamily. *Biochem Pharmacol*, Jun 2007.

[11] T R Walsh, M A Toleman, L Poirel, and P Nordmann. Metallo-beta-lactamases: the quiet before the storm? *Clin Microbiol Rev*, 18(2):306–325, Apr 2005.

[12] B H Normark and S Normark. Evolution and spread of antibiotic resistance. *J Intern Med*, 252(2):91–106, Aug 2002.

[13] S M Fabiane, M K Sohi, T Wan, D J Payne, J H Bateson, T Mitchell, and B J Sutton. Crystal structure of the zinc-dependent beta-lactamase from bacillus cereus at 1.9 a resolution: binuclear active site with features of a mononuclear enzyme. *Biochemistry*, 37(36):12404–12411, Sep 1998.

[14] G Garau, C Bebrone, C Anne, M Galleni, J M Frère, and O Dideberg. A metallo-beta-lactamase enzyme in action: crystal structures of the monozinc carbapenemase cpha and its complex with biapenem. *J Mol Biol*, 345(4):785–795, Jan 2005.

[15] J H Ullah, T R Walsh, I A Taylor, D C Emery, C S Verma, S J Gamblin, and J Spencer. The crystal structure of the l1 metallo-beta-lactamase from stenotrophomonas maltophilia at 1.7 a resolution. *J Mol Biol*, 284(1):125–136, Nov 1998.

[16] P M Shah and R D Isaacs. Ertapenem, the first of a new group of carbapenems. *J Antimicrob Chemother*, 52(4):538–542, Oct 2003.

[17] F Walsh. Doripenem: A new carbapenem antibiotic a review of comparative antimicrobial and bactericidal activities. *Ther Clin Risk Manag*, 3(5):789–794, Oct 2007.

[18] S R Norrby. Carbapenems. *Med Clin North Am*, 79(4):745–759, Jul 1995.

[19] K F Barker. Antibiotic resistance: a current perspective. *Br J Clin Pharmacol*, 48(2):109–124, Aug 1999.

[20] V Carnevale, S Raugei, C Micheletti, and P Carloni. Convergent dynamics in the protease enzymatic superfamily. *J Am Chem Soc*, 128(30):9766–9772, Aug 2006.

[21] M Cascella, C Micheletti, U Rothlisberger, and P Carloni. Evolutionarily conserved functional mechanics across pepsin-like and retroviral aspartic proteases. *J Am Chem Soc*, 127(11):3734–3742, Mar 2005.

[22] M Dal Peraro, L I Llarrull, U Rothlisberger, A J Vila, and P Carloni. Water-assisted reaction mechanism of monozinc beta-lactamases. *J Am Chem Soc*, 126(39):12661–12668, Oct 2004.

[23] M Dal Peraro, A J Vila, P Carloni, and M L Klein. Role of zinc content on the catalytic efficiency of b1 metallo beta-lactamases. *J Am Chem Soc*, 129(10):2808–2816, Mar 2007.

[24] N Díaz, D Suárez, and K M Merz. Molecular dynamics simulations of the mononuclear zinc-beta-lactamase from bacillus cereus complexed with

benzylpenicillin and a quantum chemical study of the reaction mechanism. *J Am Chem Soc*, 123(40):9867–9879, Oct 2001.

[25] P Oelschlaeger, R D Schmid, and J Pleiss. Insight into the mechanism of the imp-1 metallo-beta-lactamase by molecular dynamics simulations. *Protein Eng*, 16(5):341–350, May 2003.

[26] H Park, E N Brothers, and K M Merz. Hybrid qm/mm and dft investigations of the catalytic mechanism and inhibition of the dinuclear zinc metallo-beta-lactamase ccra from bacteroides fragilis. *J Am Chem Soc*, 127(12):4232–4241, Mar 2005.

[27] D Suárez, N Díaz, and K M Merz. Molecular dynamics simulations of the dinuclear zinc-beta-lactamase from bacteroides fragilis complexed with imipenem. *J Comput Chem*, 23(16):1587–1600, Dec 2002.

[28] D Suárez and K M Merz. Molecular dynamics simulations of the mononuclear zinc-beta-lactamase from bacillus cereus. *J Am Chem Soc*, 123(16):3759–3770, Apr 2001.

[29] P E Tomatis, R M Rasia, L Segovia, and A J Vila. Mimicking natural evolution in metallo-beta-lactamases through second-shell ligand mutations. *Proc Natl Acad Sci U S A*, 102(39):13761–13766, Sep 2005.

[30] G Taubes. The bacteria fight back. *Science*, 321(5887):356–361, Jul 2008.

[31] E Sauvage, F Kerff, M Terrak, J A Ayala, and P Charlier. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol Rev*, 32(2):234–258, Mar 2008.

[32] P Born, E Breukink, and W Vollmer. In vitro synthesis of cross-linked murein and its attachment to sacculi by pbp1a from escherichia coli. *J Biol Chem*, 281(37):26985–26993, Sep 2006.

[33] Kevin D Young David L Popham. Role of penicillin-binding proteins in bacterial cell morphogenesis. *Current Opinion in Microbiology*, 6(6):594–599, 2003.

[34] D J Tipper and J L Strominger. Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-d-alanyl-d-alanine. *Proc Natl Acad Sci U S A*, 54(4):1133–1141, Oct 1965.

[35] E P Abraham and E Chain. An enzyme from bacteria able to destroy penicillin. 1940. *Rev Infect Dis*, 10(4):677–678, Jul-Aug 1988.

[36] B A Rasmussen, Y Gluzman, and F P Tally. Cloning and sequencing of the class b beta-lactamase gene (ccra) from bacteroides fragilis tal3636. *Antimicrob Agents Chemother*, 34(8):1590–1592, Aug 1990.

[37] M Watanabe, S Iyobe, M Inoue, and S Mitsuhashi. Transferable imipenem resistance in pseudomonas aeruginosa. *Antimicrob Agents Chemother*, 35(1):147–151, Jan 1991.

[38] O Massidda, G M Rossolini, and G Satta. The aeromonas hydrophila cpha gene: molecular heterogeneity among class b metallo-beta-lactamases. *J Bacteriol*, 173(15):4611–4617, Aug 1991.

[39] E Osano, Y Arakawa, R Wacharotayankun, M Ohta, T Horii, H Ito, F Yoshimura, and N Kato. Molecular characterization of an enterobacterial metallo beta-lactamase found in a clinical isolate of serratia marcescens that shows imipenem resistance. *Antimicrob Agents Chemother*, 38(1):71–78, Jan 1994.

[40] G M Rossolini, N Franceschini, M L Riccio, P S Mercuri, M Perilli, M Galleni, J M Frere, and G Amicosante. Characterization and sequence of the chryseobacterium (flavobacterium) meningosepticum carbapenemase: a new molecular class b beta-lactamase showing a broad substrate profile. *Biochem J*, 332 (Pt 1):145–152, May 1998.

[41] G M Rossolini, M A Condemi, F Pantanella, J D Docquier, G Amicosante, and M C Thaller. Metallo-beta-lactamase producers in environmental microbiota: new molecular class b enzyme in janthinobacterium lividum. *Antimicrob Agents Chemother*, 45(3):837–844, Mar 2001.

[42] M J Saavedra, L Peixe, J C Sousa, I Henriques, A Alves, and A Correia. Sfh-i, a subclass b2 metallo-beta-lactamase from a serratia fonticola environmental isolate. *Antimicrob Agents Chemother*, 47(7):2330–2333, Jul 2003.

[43] A M Simm, C S Higgins, S T Pullan, M B Avison, P Niumsup, O Erdozain, P M Bennett, and T R Walsh. A novel metallo-beta-lactamase, mbl1b, produced by the environmental bacterium caulobacter crescentus. *FEBS Lett*, 509(3):350–354, Dec 2001.

[44] M Stoczko, J M Frère, G M Rossolini, and J D Docquier. Postgenomic scan of metallo-beta-lactamase homologues in rhizobacteria: identification and characterization of bjp-1, a subclass b3 ortholog from bradyrhizobium japonicum. *Antimicrob Agents Chemother*, 50(6):1973–1981, Jun 2006.

[45] G Garau, I García-Sáez, C Bebrone, C Anne, P Mercuri, M Galleni, J M Frère, and O Dideberg. Update of the standard numbering scheme for class b beta-lactamases. *Antimicrob Agents Chemother*, 48(7):2347–2349, Jul 2004.

[46] M Hussain, A Carlino, M J Madonna, and J O Lampen. Cloning and sequencing of the metallothioprotein beta-lactamase ii gene of bacillus cereus 569/h in escherichia coli. *J Bacteriol*, 164(1):223–229, Oct 1985.

[47] S Bellais, D Girlich, A Karim, and P Nordmann. Ebr-1, a novel ambler subclass b1 beta-lactamase from empedobacter brevis. *Antimicrob Agents Chemother*, 46(10):3223–3227, Oct 2002.

[48] N Laraki, N Franceschini, G M Rossolini, P Santucci, C Meunier, E de Pauw, G Amicosante, J M Frère, and M Galleni. Biochemical characterization of the pseudomonas aeruginosa 101/1477 metallo-beta-lactamase imp-1 produced by escherichia coli. *Antimicrob Agents Chemother*, 43(4):902–906, Apr 1999.

[49] T R Walsh, S Gamblin, D C Emery, A P MacGowan, and P M Bennett. Enzyme kinetics and biochemical analysis of imis, the metallo-beta-lactamase from aeromonas sobria 163a. *J Antimicrob Chemother*, 37(3):423–431, Mar 1996.

[50] T R Walsh, L Hall, S J Assinder, W W Nichols, S J Cartwright, A P MacGowan, and P M Bennett. Sequence analysis of the l1 metallo-beta-lactamase from xanthomonas maltophilia. *Biochim Biophys Acta*, 1218(2):199–201, Jun 1994.

[51] L Boschi, P S Mercuri, M L Riccio, G Amicosante, M Galleni, J M Frère, and G M Rossolini. The legionella (fluoribacter) gormanii metallo-beta-lactamase: a new member of the highly divergent lineage of molecular-subclass b3 beta-lactamases. *Antimicrob Agents Chemother*, 44(6):1538–1543, Jun 2000.

[52] A Felici, G Amicosante, A Oratore, R Strom, P Ledent, B Joris, L Fanuel, and J M Frère. An overview of the kinetic parameters of class b beta-lactamases. *Biochem J*, 291 (Pt 1):151–155, Apr 1993.

[53] B Segatore, O Massidda, G Satta, D Setacci, and G Amicosante. High specificity of cpha-encoded metallo-beta-lactamase from aeromonas hydrophila ae036 for carbapenems and its contribution to beta-lactam resistance. *Antimicrob Agents Chemother*, 37(6):1324–1328, Jun 1993.

[54] M Vanhove, M Zakhem, B Devreese, N Franceschini, C Anne, C Bebrone, G Amicosante, G M Rossolini, J Van Beeumen, J M Frère, and M Galleni. Role of cys221 and asn116 in the zinc-binding sites of the aeromonas hydrophila metallo-beta-lactamase. *Cell Mol Life Sci*, 60(11):2501–2509, Nov 2003.

[55] C Bebrone, C Anne, K De Vriendt, B Devreese, G M Rossolini, J Van Beeumen, J M Frère, and M Galleni. Dramatic broadening of the substrate profile of the aeromonas hydrophila cpha metallo-beta-lactamase by site-directed mutagenesis. *J Biol Chem*, 280(31):28195–28202, Aug 2005.

[56] G Garau, A M Di Guilmi, and B G Hall. Structure-based phylogeny of the metallo-beta-lactamases. *Antimicrob Agents Chemother*, 49(7):2778–2784, Jul 2005.

[57] B G Hall and M Barlow. Revised ambler classification of beta-lactamases. *J Antimicrob Chemother*, 55(6):1050–1051, Jun 2005.

[58] A Carfi, S Pares, E Duée, M Galleni, C Duez, J M Frère, and O Dideberg. The 3-d structure of a zinc metallo-beta-lactamase from bacillus cereus reveals a new type of protein fold. *EMBO J*, 14(20):4914–4921, Oct 1995.

[59] N O Concha, B A Rasmussen, K Bush, and O Herzberg. Crystal structure of the wide-spectrum binuclear zinc beta-lactamase from bacteroides fragilis. *Structure*, 4(7):823–836, Jul 1996.

[60] A Carfi, E Duée, M Galleni, J M Frère, and O Dideberg. 1.85 a resolution structure of the zinc (ii) beta-lactamase from bacillus cereus. *Acta Crystallogr D Biol Crystallogr*, 54(Pt 3):313–323, May 1998.

[61] N O Concha, C A Janson, P Rowling, S Pearson, C A Cheever, B P Clarke, C Lewis, M Galleni, J M Frère, D J Payne, J H Bateson, and S S Abdel-Meguid. Crystal structure of the imp-1 metallo beta-lactamase from pseudomonas aeruginosa and its complex with a mercaptocarboxylate inhibitor: binding determinants of a potent, broad-spectrum inhibitor. *Biochemistry*, 39(15):4288–4298, Apr 2000.

[62] T A Murphy, L E Catto, S E Halford, A T Hadfield, W Minor, T R Walsh, and J Spencer. Crystal structure of pseudomonas aeruginosa spm-1 provides insights into variable zinc affinity of metallo-beta-lactamases. *J Mol Biol*, 357(3):890–903, Mar 2006.

[63] I García-Sáez, P S Mercuri, C Papamicael, R Kahn, J M Frère, M Galleni, G M Rossolini, and O Dideberg. Three-dimensional structure of fez-1, a monomeric subclass b3 metallo-beta-lactamase from fluoribacter gormanii, in native form and in complex with d-captopril. *J Mol Biol*, 325(4):651–660, Jan 2003.

[64] C Moali, C Anne, J Lamotte-Brasseur, S Groslambert, B Devreese, J Van Beeumen, M Galleni, and J M Frère. Analysis of the importance of the metallo-beta-lactamase active site loop in substrate binding and catalysis. *Chem Biol*, 10(4):319–329, Apr 2003.

[65] J Morán-Barrio, J M González, M N Lisa, A L Costello, M D Peraro, P Carloni, B Bennett, D L Tierney, A S Limansky, A M Viale, and A J Vila. The metallo-beta-lactamase gob is a mono-zn(ii) enzyme with a novel active site. *J Biol Chem*, 282(25):18286–18293, Jun 2007.

[66] M Hernandez Valladares, A Felici, G Weber, H W Adolph, M Zeppezauer, G M Rossolini, G Amicosante, J M Frère, and M Galleni. Zn(ii) dependence of the aeromonas hydrophila ae036 metallo-beta-lactamase activity and stability. *Biochemistry*, 36(38):11534–11541, Sep 1997.

[67] U Heinz, R Bauer, S Wommer, W Meyer-Klaucke, C Papamichaels, J Bateson, and H W Adolph. Coordination geometries of metal ions in d- or l-captopril-inhibited metallo-beta-lactamases. *J Biol Chem*, 278(23):20659–20666, Jun 2003.

[68] P A Crawford, K W Yang, N Sharma, B Bennett, and M W Crowder. Spectroscopic studies on cobalt(ii)-substituted metallo-beta-lactamase imis from aeromonas veronii bv. sobria. *Biochemistry*, 44(13):5168–5176, Apr 2005.

[69] A L Costello, N P Sharma, K W Yang, M W Crowder, and D L Tierney. X-ray absorption spectroscopy of the zinc-binding sites in the class b2 metallo-beta-lactamase imis from aeromonas veronii bv. sobria. *Biochemistry*, 45(45):13650–13658, Nov 2006.

[70] D de Seny, U Heinz, S Wommer, M Kiefer, W Meyer-Klaucke, M Galleni, J M Frere, R Bauer, and H W Adolph. Metal ion binding and coordination geometry for wild type and mutants of metallo-beta -lactamase from bacillus cereus 569/h/9 (bcii): a combined thermodynamic, kinetic, and spectroscopic approach. *J Biol Chem*, 276(48):45065–45078, Nov 2001.

[71] S Wommer, S Rival, U Heinz, M Galleni, J M Frere, N Franceschini, G Amicosante, B Rasmussen, R Bauer, and H W Adolph. Substrate-activated zinc binding of metallo-beta -lactamases: physiological importance of mononuclear enzymes. *J Biol Chem*, 277(27):24142–24147, Jul 2002.

[72] S Bounaga, A P Laws, M Galleni, and M I Page. The mechanism of catalysis and the inhibition of the bacillus cereus zinc-dependent beta-lactamase. *Biochem J*, 331 (Pt 3):703–711, May 1998.

[73] R Paul-Soto, M Hernandez-Valladares, M Galleni, R Bauer, M Zeppezauer, J M Frère, and H W Adolph. Mono- and binuclear zn-beta-lactamase from bacteroides fragilis: catalytic and structural roles of the zinc ions. *FEBS Lett*, 438(1-2):137–140, Oct 1998.

[74] W Fast, Z Wang, and S J Benkovic. Familial mutations and zinc stoichiometry determine the rate-limiting step of nitrocefin hydrolysis by metallo-beta-lactamase from bacteroides fragilis. *Biochemistry*, 40(6):1640–1650, Feb 2001.

[75] N. O. Concha, B. A. Rasmussen, K. Bush, and O. Herzberg. Crystal structure of the wide-spectrum binuclear zinc β-lactamase from Bacteroides fragilis. *Structure*, *4*:823–836, 1996.

[76] R M Rasia, M Ceolín, and A J Vila. Grafting a new metal ligand in the cocat-
     alytic site of b. cereus metallo-beta-lactamase: structural flexibility without
     loss of activity. *Protein Sci*, 12(7):1538–1546, Jul 2003.

[77] S McManus-Munoz and M W Crowder. Kinetic mechanism of metallo-
     beta-lactamase l1 from stenotrophomonas maltophilia. *Biochemistry*,
     38(5):1547–1553, Feb 1999.

[78] Z. Wang, W. Fast, and S.J. Benkovic. Direct observation of an enzyme-
     bound intermediate in the catalytic cycle of the metallo-beta-lactamase
     from bacteroides fragilis. *Journal of the American Chemical Society*,
     120(41):10788–10789, 1998.

[79] N V Kaminskaia, B Spingler, and S J Lippard. Intermediate in beta-lactam
     hydrolysis catalyzed by a dinuclear zinc(ii) complex: relevance to the mech-
     anism of metallo-beta-lactamase. *J Am Chem Soc*, 123(27):6555–6563, Jul
     2001.

[80] J D Garrity, B Bennett, and M W Crowder. Direct evidence that the reaction
     intermediate of metallo-beta-lactamase l1 is metal bound. *Biochemistry*,
     44(3):1078–1087, Jan 2005.

[81] J Spencer, A R Clarke, and T R Walsh. Novel mechanism of hydrolysis of
     therapeutic beta-lactams by stenotrophomonas maltophilia l1 metallo-beta-
     lactamase. *J Biol Chem*, 276(36):33638–33644, Sep 2001.

[82] D Suárez, E N Brothers, and K M Merz. Insights into the structure and
     dynamics of the dinuclear zinc beta-lactamase site from bacteroides fragilis.
     Biochemistry, 41(21):6615–6630, May 2002.

[83] M Dal Peraro, A J Vila, and P Carloni. Protonation state of asp120 in the
     binuclear active site of the metallo-beta-lactamase from bacteroides fragilis.
     Inorg Chem, 42(14):4245–4247, Jul 2003.

[84] J D Garrity, A L Carenbauer, L R Herron, and M W Crowder. Metal binding
     asp-120 in metallo-beta-lactamase l1 from stenotrophomonas maltophilia
     plays a crucial role in catalysis. *J Biol Chem*, 279(2):920–927, Jan 2004.

[85] Z Wang, W Fast, and S J Benkovic. On the mechanism of the metallo-beta-
     lactamase from bacteroides fragilis. *Biochemistry*, 38(31):10013–10023,
     Aug 1999.

[86] Z Wang and S J Benkovic. Purification, characterization, and kinetic studies
     of a soluble bacteroides fragilis metallo-beta-lactamase that provides multi-
     ple antibiotic resistance. *J Biol Chem*, 273(35):22402–22408, Aug 1998.

[87] M P Yanchak, R A Taylor, and M W Crowder. Mutational analysis
     of metallo-beta-lactamase ccra from bacteroides fragilis. *Biochemistry*,
     39(37):11330–11339, Sep 2000.

[88] A L Carenbauer, J D Garrity, G Periyannan, R B Yates, and M W Crowder. Probing substrate binding to metallo-beta-lactamase l1 from stenotrophomonas maltophilia by using site-directed mutagenesis. *BMC Biochem*, 3:4–4, 2002.

[89] J Spencer, J Read, R B Sessions, S Howell, G M Blackburn, and S J Gamblin. Antibiotic recognition by binuclear metallo-beta-lactamases revealed by x-ray crystallography. *J Am Chem Soc*, 127(41):14439–14444, Oct 2005.

[90] D de Seny, C Prosperi-Meys, C Bebrone, G M Rossolini, M I Page, P Noel, J M Frère, and M Galleni. Mutational analysis of the two zinc-binding sites of the bacillus cereus 569/h/9 metallo-beta-lactamase. *Biochem J*, 363(Pt 3):687–696, May 2002.

[91] Y Yamaguchi, T Kuroki, H Yasuzawa, T Higashi, W Jin, A Kawanami, Y Yamagata, Y Arakawa, M Goto, and H Kurosaki. Probing the role of asp-120(81) of metallo-beta-lactamase (imp-1) by site-directed mutagenesis, kinetic studies, and x-ray crystallography. *J Biol Chem*, 280(21):20824–20832, May 2005.

[92] M Dal Peraro, A J Vila, and P Carloni. Structural determinants and hydrogen-bond network of the mononuclear zinc(ii)-beta-lactamase active site. *J Biol Inorg Chem*, 7(7-8):704–712, Sep 2002.

[93] A. M. Davies, R. M. Rasia, A. J. Vila, B. J. Sutton, and S. M. Fabiane. Effect of pH on the active site of an Arg121Cys mutant of the metallo-β-lactamase from Bacillus cereus: Implications for the enzyme mechanism. *Biochem.*, 44:4841–4849, 2005.

[94] R. M. Rasia and A. J. Vila. Exploring the role and the binding affinity of a second zinc equivalent in B. cereus metallo-β-lactamase. *Biochem.*, 41:1853–1860, 2002.

[95] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, Th. Frauenheim, S. Suhai, and G. Seifert. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B*, 58(11):7260–7268, Sep 1998.

[96] D Xu, Y Zhou, D Xie, and H Guo. Antibiotic binding to monozinc cpha beta-lactamase from aeromonas hydropila: quantum mechanical/molecular mechanical and density functional theory studies. *J Med Chem*, 48(21):6679–6689, Oct 2005.

[97] D Xu, D Xie, and H Guo. Catalytic mechanism of class b2 metallo-beta-lactamase. *J Biol Chem*, 281(13):8740–8747, Mar 2006.

[98] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B):B864–B871, Nov 1964.

[99] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133–A1138, Nov 1965.

[100] W. J. Carr. Energy, specific heat, and magnetic properties of the low-density electron gas. *Phys. Rev.*, 122(5):1437–1446, Jun 1961.

[101] W. J. Carr and A. A. Maradudin. Ground-state energy of a high-density electron gas. *Phys. Rev.*, 133(2A):A371–A374, Jan 1964.

[102] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45(7):566–569, Aug 1980.

[103] J. P. Perdew and Alex Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23(10):5048–5079, May 1981.

[104] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.*, 58:1200–1211, 1980.

[105] John P. Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B*, 45(23):13244–13249, Jun 1992.

[106] F. Sim, A. St. Amant, I. Papai, and D. R. Salahub. Gaussian density functional calculations on hydrogen-bonded systems. 114(11):4391–4400, 1992.

[107] C. Tuma, A. D. Boese, and N. C. Handy. Predicting the binding energies of h-bonded complexes: A comparative dft study. 1:3939–3947, 1999.

[108] K. Capelle. A bird's-eye view of density-functional theory. *arxiv:cond-mat/0211443*, 2006.

[109] R. O. Jones and O. Gunnarsson. The density functional formalism, its applications and prospects. *Rev. Mod. Phys.*, 61(3):689–746, Jul 1989.

[110] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38(6):3098–3100, Sep 1988.

[111] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37(2):785–789, Jan 1988.

[112] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Harcourt, Orlando, 1976.

[113] Glenn J. Martyna and Mark E. Tuckerman. A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *The Journal of Chemical Physics*, 110(6):2810–2821, 1999.

[114] W. H. Press, B. P. Flannery S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge, 1 edition, 1986.

[115] N. Troullier and José Luriaas Martins. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B*, 43(3):1993–2006, Jan 1991.

[116] Leonard Kleinman and D. M. Bylander. Efficacious form for model pseudopotentials. *Phys. Rev. Lett.*, 48(20):1425–1428, May 1982.

[117] R. P. Feynman. Forces in molecules. *Phys. Rev.*, 56(4):340–343, Aug 1939.

[118] H. Hellmann. Einführung in die quantenchemie. *Phys. Rev.*, 1937.

[119] D. Marx and J. Hütter. *Modern Methods and Algorithms of Quantum Chemistry*, volume 1, chapter "Ab initio molecular dynamics: Theory and implementation", pages 301–449. John von Neumann Institute for Computing, Jülich, NIC Series, 2000.

[120] A Warshel and M Levitt. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol*, 103(2):227–249, May 1976.

[121] D. Bakowies and W. Thiel. Hybrid models for combined quantum mechanical and molecular mechanical approaches. *Journal of Physical Chemistry*, 100(25):10580–10594, 1996.

[122] M. Eichinger, P. Tavan, J. Hutter, and M. Parrinello. A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields. *The Journal of Chemical Physics*, 110(21):10452–10467, 1999.

[123] Alessandro Laio, Joost VandeVondele, and Ursula Rothlisberger. A hamiltonian electrostatic coupling scheme for hybrid car–parrinello molecular dynamics simulations. *The Journal of Chemical Physics*, 116(16):6941–6947, 2002.

[124] D. A. Case, D. A. Pearlman, J. W. Caldwell, T.E. Cheatham IIIW. S., J. Wang, Ross, C.L. Simmerling, T. Darden, K.M. Merz, R. V. Stanton, A.L. Cheng, J. J. Vincent, M. Crowley, V. Tsui, H. Gohlke, R. J Radmer, Y. Duan, J. Pittera, I. Massova, G. L. Seibel, U. C. Singh, P.K. Weiner, and P. A. Kollman. AMBER7. *University of California, San Francisco*, 2002.

[125] J. Gao, P. Amara, C. Alhambra, and M.J. Field. A generalized hybrid orbital (gho) method for the treatment of boundary atoms in combined qm/mm calculations. *Journal of Physical Chemistry A*, 102(24):4714–4721, 1998.

[126] I. Antes and W. Thiel. Adjusted connection atoms for combined quantum mechanical and molecular mechanical methods. *Journal of Physical Chemistry A*, 103(46):9290–9295, 1999.

[127] Yingkai Zhang, Tai-Sung Lee, and Weitao Yang. A pseudobond approach to combining quantum mechanical and molecular mechanical methods. *The Journal of Chemical Physics*, 110(1):46–54, 1999.

[128] N. Reuter, A. Dejaegere, B. Maigret, and M. Karplus. Frontier bonds in qm/mm methods: A comparison of different approaches. *Journal of Physical Chemistry A*, 104(8):1720–1735, 2000.

[129] V. Kairys and J.H. Jensen. Qm/mm boundaries across covalent bonds: A frozen localized molecular orbital-based approach for the effective fragment potential method. *Journal of Physical Chemistry A*, 104(28):6656–6665, 2000.

[130] 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(29):7300–7307, 2002.

[131] D. Frenkel and B. Smit. *Understanding Molecular Simulations: From Algorithms to Applications*. Academic Press, San Diego, CA, 2 edition, 2002.

[132] Wendy D. Cornell, Piotr Cieplak, Christopher I. Bayly, Ian R. Gould, Kenneth M. Merz, David M. Ferguson, David C. Spellmeyer, Thomas Fox, James W. Caldwell, and Peter A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, 1995.

[133] C. J. Cramer. *Essentials of Computational Chemistry*. Wiley, New York, 2 edition, 2002.

[134] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, 1989.

[135] P. P. Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Annalen der Physik*, 369(3):253–287, 1921.

[136] T. Darden, D. York, and L. Pedersen. Particle mesh ewald: An nlog(n) method for ewald sums in large systems. 98:10089–10092, 1993.

[137] R. W. Hockney and J. W. Eastwood. *Computer Simulations Using Particles*. McGraw-Hill, New York, 1981.

[138] S. Nosé. A molecular dynamics method for simulations in the canonical ensemble. 52:255–268, 1984.

[139] W. G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. 31:1695–1697, 1985.

[140] G. J. Martyna, M. L. Klein, and M. Tuckerman. Nosé–hoover chains: The canonical ensemble via continuous dynamics. 97(4):2635–2643, 1992.

[141] M. T. Nelson, W. Humphrey, A. Gursoy, A. Dalke, L. V. Kale, R. D. Skeel, and K. Schulten. NAMD: A parallel, object oriented molecular dynamics program. *Int. J. High Perform C.*, *10*:251–268, 1996.

[142] C. Chipot and A. Pohorille, editors. *Free Energy Calculations: Theory and Applications in Chemistry and Biology*, volume 86 of *Springer Series in Chemical Physics*. Springer, 2007.

[143] N. Foloppe and R. Hubbard. Towards predictive ligand design with free-energy based computational methods? *Curr Med Chem*, 13(29):3583–3608, 2006.

[144] E. A. Carter, G. Ciccotti, J. T. Hynes, and R. Kapral. Constrained reaction coordinate dynamics for the simulation of rare events. *Chem. Phys. Lett.*, 156(5):472–477, April 1989.

[145] G. M. Torrie and J. P. Valleau. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. 23:187–199, 1977.

[146] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman. Multidimensional free-energy calculations using the weighted histogram analysis method. *J. Comp. Chem.*, 16(11):1339–1350, November 1995.

[147] Benoît Roux. Ion conduction and selectivity in k(+) channels. *Annu Rev Biophys Biomol Struct*, 34:153–171, 2005.

[148] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, *78*:2690–2693, 1997.

[149] C. Jarzynski. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Phys. Rev. E*, *56*:5018–5035, 1997.

[150] E. Darve and A. Pohorille. Calculating free energies using average force. 115(20):9169–9183, 2001.

[151] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler. Transition path sampling and the calculation of rate constants. 108(5):1964–1977, 1998.

[152] G. Ciccotti, M. Ferrario, J. T. Hynes, and R. Karpal. Constrained molecular dynamics and the mean potential for an ion pair in a polar solvent. *Chem. Phys.*, *129*:241–251, 1989.

[153] Michiel Sprik and Giovanni Ciccotti. Free energy from constrained molecular dynamics. *The Journal of Chemical Physics*, 109(18):7737–7744, 1998.

[154] G. Garau, C. Bebrone, C. Anne, M. Galleni, J. M. Frere, and O. Dideberg. A metallo-β-lactamase enzyme in action: Crystal structures of the monozinc carbapenemase CphA and its complex with biapenem. *J. Mol. Biol.*, *345*:785–795, 2005.

[155] P. A. Crawford, K. W. Yang, N. Sharma, B. Bennett, and M. W. Crowder. Spectroscopic studies on cobalt(II)-substituted metallo-β-lactamase ImiS from Aeromonas veronii bv. sobria. *Biochem.*, *44*:5168–5176, 2005.

[156] S. D. B. Scrofani, J. Chung, J. J. A. Huntley, S. J. Benkovic, P. E. Wright, and H. J. Dyson. NMR characterization of the metallo-β-lactamase from Bacteroides fragilis and its interaction with a tight-binding inhibitor: Role of an active-site loop. *Biochem.*, *38*:14507–14514, 1999.

[157] J. J. A. Huntley, S. D. B. Scrofani, M. J. Osborne, P. E. Wright, and H. J. Dyson. Dynamics of the metallo-β-lactamase from Bacteroides fragilis in the presence and absence of a tight-binding inhibitor. *Biochem.*, *39*:13356–13364, 2000.

[158] F. R. Salsbury, M. F. Crowley, and C. L. Brooks. Modeling of the metallo-β-lactamase from B-fragilis: Structural and dynamic effects of inhibitor binding. *Proteins Struct. Funct. and Bioinf.*, *44*:448–459, 2001.

[159] D. Suarez and K. M. Merz. Molecular dynamics simulations of the mononuclear zinc-β-lactamase from Bacillus cereus. *J. Am. Chem. Soc.*, *123*:3759–3770, 2001.

[160] P. Oelschlaeger, R. D. Schmid, and J. Pleiss. Modeling domino effects in enzymes: Molecular basis of the substrate specificity of the bacterial metallo-β-lactamases IMP-1 and imp-6. *Biochem.*, *42*:8945–8956, 2003.

[161] D. Suarez, N. Diaz, and K. M. Merz. Molecular dynamics simulations of the dinuclear zinc-β-lactamase from bacteroides fragilis complexed with imipenem. *J. Comput. Chem.*, *23*:1587–1600, 2002.

[162] N. Diaz, D. Suarez, K. M. Merz, and T. L. Sordo. Molecular dynamics simulations of the TEM-1,β-lactamase complexed with cephalothin. *J. Med. Chem.*, *48*:780–791, 2005.

[163] M. Dal Peraro, A. J. Vila, and P. Carloni. Protonation state of Asp120 in the binuclear active site of the metallo-β-lactamase from Bacteroides fragilis. *Inorg. Chem.*, *42*:4245–4247, 2003.

[164] M. Dal Peraro, A. J. Vila, and P. Carloni. Structural determinants and hydrogen-bond network of the mononuclear zinc(II)-β-lactamase active site. *J. Biol. Inorg. Chem.*, *7*:704–712, 2002.

[165] M. Dal Peraro, A. J. Vila, and P. Carloni. Substrate binding to mononuclear metallo-β-lactamase from Bacillus cereus. *Proteins Struct. Funct. and Bioinf.*, *54*:412–423, 2004.

[166] H. Park, E. N. Brothers, and K. M. Merz. Hybrid QM/MM and DFT investigations of the catalytic mechanism and inhibition of the dinuclear zinc metallo-β-lactamase CcrA from Bacteroides fragilis. *J. Am. Chem. Soc.*, *127*:4232–4241, 2005.

[167] P. Oelschlaeger, R. D. Schmid, and J. Pleiss. Insight into the mechanism of the IMP-1 metallo-β-lactamase by molecular dynamics simulations. *Protein Eng. Des. Sel.*, *16*:341–350, 2003.

[168] G. Estiu, D. Suárez, and K.M. Merz. Quantum mechanical and molecular dynamics simulations of ureases and zn β-lactamases. *J. Comput. Chem.*, *27*(*12*):1240–1262, 2006.

[169] F Simona, A Magistrato, D M Vera, G Garau, A J Vila, and P Carloni. Protonation state and substrate binding to b2 metallo-beta-lactamase cpha from aeromonas hydrofila. *Proteins*, 69(3):595–605, Jul 2007.

[170] R. Car and M. Parrinello. Unified approach for molecular-dynamics and density-functional theory. *Phys. Rev. Lett.*, *55*:2471–2474, 1985.

[171] A D Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A*, 38(6):3098–3100, Sep 1988.

[172] C Lee, W Yang, and R G Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Phys Rev B Condens Matter*, 37(2):785–789, Jan 1988.

[173] P Carloni, U Rothlisberger, and M Parrinello. The role and perspective of ab initio molecular dynamics in the study of biological systems. *Acc Chem Res*, 35(6):455–464, Jun 2002.

[174] M Dal Peraro, P Ruggerone, S Raugei, F L Gervasio, and P Carloni. Investigating biological systems using first principles car-parrinello molecular dynamics simulations. *Curr Opin Struct Biol*, 17(2):149–156, Apr 2007.

[175] C M Bebrone, C Anne, F Kerff, G Garau, K De Vriendt, R Lantin, B Devreese, J Van Beeumen, O Dideberg, J M Frère, and M Galleni. Mutational analysis of the zinc and substrate binding sites in the cpha metallo-beta-lactamase from aeromonas hydrophila. *Biochem J*, May 2008.

[176] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *Journal of Physical Chemistry*, 98(45):11623–11627, 1994.

[177] S. Piana, D. Bucher, P. Carloni, and U. Rothlisberger. Reaction mechanism of hiv-1 protease by hybrid car-parrinello/classical md simulations. *Journal of Physical Chemistry B*, 108(30):11139–11149, 2004.

[178] A. Magistrato, J. VandeVondele, and U. Rothlisberger. Three- and four-center trans effects in triply bonded ditungsten complexes: An ab initio molecular dynamics study of compounds with stoichiometry w2cl4(nhet)2(pme3)2. *Inorganic Chemistry*, 39(24):5553–5560, 2000.

[179] B Tamames, S F Sousa, J Tamames, P A Fernandes, and M J Ramos. Analysis of zinc-ligand bond lengths in metalloproteins: trends and patterns. *Proteins*, 69(3):466–475, Nov 2007.

[180] L A Abriata, L J González, L I Llarrull, P E Tomatis, W K Myers, A L Costello, D L Tierney, and A J Vila. Engineered mononuclear variants in bacillus cereus metallo-beta-lactamase bcii are inactive. *Biochemistry*, 47(33):8590–8599, Aug 2008.

[181] M Dal Peraro, A J Vila, and P Carloni. Substrate binding to mononuclear metallo-beta-lactamase from bacillus cereus. *Proteins*, 54(3):412–423, Feb 2004.

[182] J M González, F J Medrano Martín, A L Costello, D L Tierney, and A J Vila. The zn2 position in metallo-beta-lactamases is critical for activity: a study on chimeric metal sites on a conserved protein scaffold. *J Mol Biol*, 373(5):1141–1156, Nov 2007.

[183] L I Llarrull, S M Fabiane, J M Kowalski, B Bennett, B J Sutton, and A J Vila. Asp-120 locates zn2 for optimal metallo-beta-lactamase activity. *J Biol Chem*, 282(25):18276–18285, Jun 2007.

[184] L I Llarrull, M F Tioni, and A J Vila. Metal content and localization during turnover in *B.cereus* metallo-β-lactamase. *J Am Chem Soc*, *in press*.

[185] F.L. Gervasio, V. Schettino, S. Mangani, M. Krack, P. Carloni, and M. Parrinello. Influence of outer-shell metal ligands on the structural and electronic properties of horse liver alcohol dehydrogenase zinc active site. *Journal of Physical Chemistry B*, 107(28):6886–6892, 2003.