UPPSALA
UNIVERSITET

UPTEC 22031
Examensarbete 30 hp
Juni 2022

The Binding Mechanism of Carbapenems in the Class A beta-lactamase IMI-1

A Molecular Dynamics Study of Ligand Stability

Isabell Lindahl

Civilingenjörsprogrammet i Kemiteknik

UPPSALA UNIVERSITET

The Binding Mechanism of Carbapenems in the Class A beta-lactamase IMI-1

Isabell Lindahl

Abstract

Antibiotic resistance is a global and accelerating matter. Over time, the bacteria have evolved several defense mechanisms against the antibiotics. One of the defense mechanisms is that the bacteria can produce enzymes with the ability to hydrolyze the characteristic β-lactam ring of the antibiotics. These enzymes are called β-lactamases. There are three different generations of antibiotics clinically available, and β-lactamases have co-evolved with the antibiotics over the generations. The third generation of antibiotics are called the carbapenems and β-lactamases which hydrolyze carbapenems are called carbapenemases. Carbapenemases are promiscuous, which means that they hydrolyze a variety of antibiotics. The β-lactamase IMI-1 is an imipenem-hydrolyzing enzyme and imipenem is a carbapenem, hence IMI-1 is a carbapenemase. In this project, IMI-1 was investigated in complex with the carbapenems imipenem, meropenem and biapenem using computational methods. More specifically, a homology model of IMI-1 was generated and the carbapenems were docked into the model. The system was then used for MD simulations where the important molecular interactions were identified, and the binding free energies were calculated using the LIE method. The results indicate that IMI-1 has flexible loops that enables an open and a closed conformation of IMI- 1. All three carbapenems were docked and simulated in both conformations of IMI-1. The results indicate that open and closed conformations confirms the promiscuity of carbapenemases since the flexibility enables various initial binding mechanisms. in other words, the hydrolysis may occur so quickly that the binding does not have much bearing of the activity of the enzyme. Furthermore, the calculated binding free energies indicate that IMI-1 is optimized for the catalytic process rather than the binding affinity. In conclusion, IMI-1 and similar systems requires further research using computational methods to counteract antibiotic resistance based on knowledge.

Teknisk-naturvetenskapliga fakulteten

Uppsala universitet, Utgivningsort Uppsala/Visby

Handledare: Yasmin Binti Shamsudin Ämnesgranskare: Erik Marklund

Examinator: Christian Sköld

Abbreviations

MD	Molecular dynamics
GLIDE	Grid-based ligand docking energetics
SBC	Spherical boundary constraints
SCAAS	Surface constrained all-atom solvent
LIE	Linear interaction energy method
LRA	Linear response approximation
l-s	Ligand-surroundings

Populärvetenskaplig Sammanfattning

Det kanske inte är lika högljutt som ett krig mellan stormakter, men det är ett krig som går till historieböckerna. Kriget mot bakterier och antibiotikaresistens accelererar och människan behöver nya vapen för att kunna vinna, inte bara slagen utan kriget i stort.

Föreställ dig kriget mellan människan och bakterierna. I alla tider har det bakteriella hotet hemsökt människans överlevnad, tills upptäckten av penicillinet. Penicillinet är människans första betydande vapen och man trodde att kriget mot bakterierna var vunnet. Tyvärr har bakterierna en evolutionär fördel. Med andra ord, de utvecklas i snabbare takt än människan och kan därmed utveckla nya egenskaper på kortare tid. Det tog därför inte lång tid innan bakterierna hade utvecklat ett försvar mot penicillinet. En av försvarsmekanismerna är att bakterierna själva kan tillverka enzym som bryter ner antibiotika. Enzymet fungerar på så sätt att människan har skickat iväg sitt vapen, penicillinet, med sikte mot bakterien men penicillinet kommer inte fram till målet utan blir stoppad och förstörd av enzymet. Försvaret fungerar därmed inte på avstånd liksom en pistol utan det krävs nära kontakt för att försvaret ska kunna förstöra vapnet. Människan behövde nu ännu starkare vapen och till slut upptäcktes människans andra vapen, cefalosporinerna. Cefalosporiner är samma sorts vapen som penicillin men med lite annorlunda sammansättning och innehåller därmed nya funktioner. De nya funktionerna gjorde cefalosporinerna starkare mot bakterierna och mer motståndskraftiga gentemot försvaret. Slaget var återigen vunnet, men inte kriget. Bakterierna utvecklas konstant och snart hade de utvecklat nya försvarsfunktioner till sina enzym.

Människan har just nu ett tredje vapen i bruk, carbapenemerna, vårt starkaste vapen mot bakterierna. Tyvärr utvecklas bakterierna i skrivande stund och som tidigare nämnt, i högre hastighet än människans vapen. Fler vapen med nya funktioner behövs för att människan ska kunna vinna kriget.

För att kunna designa nya vapen med nya funktioner krävs nya tillvägagångssätt– tillvägagångssätt som möjliggör vapen som är specialdesignade mot bakterierna och deras försvar. Detta tillvägagångssätt finns tillgängligt redan nu. Med hjälp av datorsimuleringar kan bakteriernas försvar inte bara byggas ihop och undersökas i 3D utan även funktionerna kan undersökas. Detta möjliggör att människans nästa vapendesign kan vara baserad på kunskapen om försvarets funktioner. Kriget sker på molekylär nivå, vapnen är jättesmå och klassiska metoder har inte kunnat berätta mer om bakteriernas försvar än hur beståndsdelarna sitter ihop med varandra. Med hjälp av de nya metoderna som utförs på datorer kan beståndsdelarna sättas ihop och funktionerna kan undersökas visuellt och med matematik. Med andra ord, det går att bygga verklighetsbaserade modeller av enzymen som sedan kan undersökas och avgörande funktioner kan identifieras. Enzymmodellerna behöver valideras, det vill säga, bekräftas som verklighetstrogna. Valideringen görs genom att beräkna hur mycket energi som krävs för avväpningen utifrån datorsimuleringarna och sedan jämförs den beräknade energin med experimentellt uppmätt energi.

I detta projekt har enzymet IMI-1 undersökts, vilken kan ta sönder många olika typer av antibiotika. Studien har fokuserat på att undersöka vilka funktioner som är avgörande i förstörandet av den tredje sortens vapen, carbapenemerna. Under den funktionella undersökningen av IMI-1 upptäcktes det att IMI- kan omforma sig till en öppen och en stängd form. Alla carbapenemer i den här studien datorsimulerades tillsammans med både den öppna och den stängda formen. Det visade sig att den öppna och stängda formen möjliggör olika inbindningsmekanismer och kan göra så att vissa av antibiotikan fångas in och hålls fast under avväpningsprocessen. Det upptäcktes också att delar av enzymet är aktiva vid attack av olika

carbapenemer. Sammanfattningsvis har försvaret en adaptionsfunktion gentemot människans vapen som gör det möjligt för enzymet att avväpna olika vapen. Fler avväpnings-apparater behöver undersökas med de nya datorsimuleringar för att kunna designa vapen som inte fastnar i enzymets gap.

1. Introduction

Bacterial infections were to be extinct as a response to the discovery of antibiotics [1]. However, bacteria have evolved defence mechanisms against antibiotics and bacterial infections are currently emerging and re-emerging due to overuse of microbial agents, among other reasons. In 2019 4.75 million deaths could be associated with antibiotic resistance [2]. Of those 4.75 million deaths, 1.27 million were directly caused by antibiotic resistance. Antibiotic resistance is a global threat as well as a combat between antibiotics and antibiotic resistance. The first β-lactam antibiotic, penicillin, was discovered in 1928 [3]. This 1st generation antibiotic was discovered in antibiotic-producing molds. Only a few years after penicillin was clinically applied, resistant bacteria started to emerge. In 1945, the 2nd generation antibiotic, cephalosporins, were discovered. History repeated itself by the emergence of further evolved bacterial defense mechanisms which could disarm cephalosporins as well. This led to the 3rd generation antibiotics, the so-called last resort antibiotics, carbapenems, which were clinically available in the 1980s. The carbapenems were working efficiently for around two decades. However, in 2001 the first promiscuous carbapenem-hydrolyzing enzyme, carbapenemase, was discovered.

The bacteria have several defence mechanisms against antibiotics [4]. One of the most widespread defense mechanisms is the production of deactivating enzymes called β-lactamases. Carbapenemase is a β-lactamase. β-lactamases hydrolyze the characteristic β-lactam ring in the antibiotics, hence the name, preventing the antibiotics from reaching its microbial target. There are four different classes of β-lactamases, A, B, C, and D, as well as numerous subclasses. The classification and characterization are based on the similarities and differences of the amino acid sequence [5]. Classes A, C and D are commonly called serine β-lactamases due to a serine at position 70, which is a major contributor to the mechanism of action. Class B differs in the sense that the mechanism of action includes a divalent zinc ion, hence they are called metallo-β-lactamases. The carbapenemase IMI-1 is a class A β-lactamase, an imipenem-hydrolyzing enzyme which was first isolated in 1984 in a hospital in California from the Gram-negative *Enterobacter cloacae* [6]. Carbapenemases are known to be promiscuous, in other words, they bind and hydrolyze a variety of antibiotics. Promiscuity of the β-lactamases increase as they evolve.

The active site residues of serine β-lactamases are Ser70-X-X-Lys73 and the mechanism of action occurs in two steps [7]. The first step, the acylation step, begins when serine 70 performs a nucleophilic attack on the β-lactam substrate, forming an acyl-enzyme adduct. In the second step, the diacylation step, the adduct is attacked by a water molecule which leads to hydrolysis as well as the release of the destroyed ligand from the active site (*Figure 1)*. It is the acyl-enzyme adduct provided in the first step that is being generated in this project using computational methods. More specifically, the acyl-enzyme adduct between IMI-1 and carbapenems.

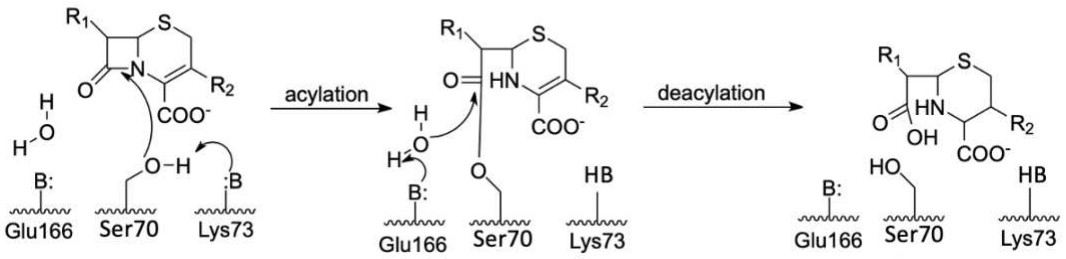


Figure 1. The mechanism of action of serine β-lactamases.

Imipenem (*Figure 2)* is a carbapenem, a last resort antibiotic. Other carbapenems in clinical use are meropenem and biapenem (*Figure 2),* among others. Carbapenems have a broad spectrum of activity and great potency against a large variety of pathogens. These pathogens include Gram-negative and Gram-positive aerobes as well as anaerobes [8]. Carbapenems differ among β-lactam microbial agents in the sense that they are generally less susceptible to hydrolysis by resistant bacteria. Furthermore, they are called last-resort antibiotics because they are administered to patients with severe infections or those who are suspected of harboring resistant bacteria. Imipenem, meropenem and biapenem are zwitter ions since they contain both a positive and a negative charge in biological pH. However, numerous studies available implies that resistance against carbapenems is increasing [8-11], which makes carbapenemases such as IMI-1 a valuable target for research.

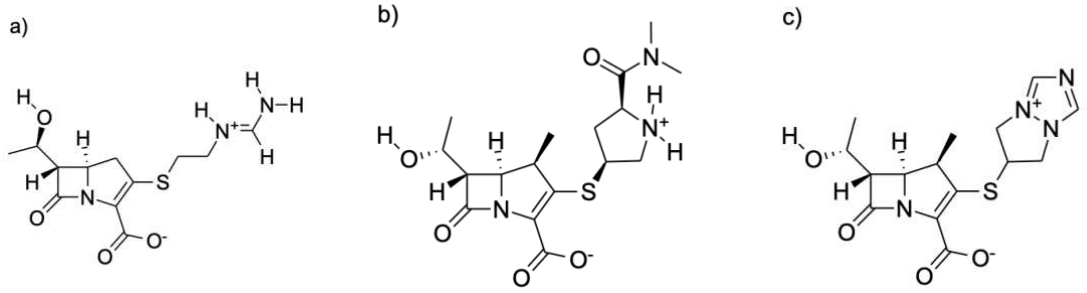


Figure 2. Chemical structures of carbapenems included in this project a) Imipenem b) Meropenem c) Biapenem.

Previous studies of IMI-1 determined the amino acid sequence, microbiological activities, and susceptibility of β-lactams, and kinetic parameters, such as k_{cat} k_m, and group affiliation [6]. Furthermore, docking and molecular interaction studies have been done on IMI-1 with inhibitors, but not carbapenems [12-13]. Hence, little is known in terms of molecular interactions and initial binding mechanism regarding IMI-1 with imipenem, meropenem and biapenem. Only the general mechanism of action of serine β-lactamases is known. Computational methods can be used to increase the knowledge of IMI-1 functionality, how it interacts with carbapenems, and in extension, how to disarm IMI-1. By creating a model of IMI-1 and dock carbapenems one can identify the important molecular interactions through molecular dynamics (MD) simulations. When the model is generated the binding free energy (ΔG_{bind}^{calc})of the system can be calculate using the linear interaction energy method (LIE). The calculated energy can be compared to experimental data (ΔG_{bind}^{exp}), validating the generated model. The validated model is thereafter used for studying and determining the initial binding mechanism of IMI-1 in complex with carbapenems.

The aims of this project are to identify important molecular interactions between carbapenems and the class A β-lactamase IMI-1, and to determine structural elements contributing to promiscuity in IMI-1.

2. Experimental – Computational Method

The general workflow of method used in this project is divided into five steps (*Figure 3.*). The first step is homology modelling where the 3D structure of IMI-1 is obtained. In the second step, the ligands, carbapenems, are prepared, followed by molecular docking of the protein model with the ligands. The docking generates poses of the ligand in the active site of IMI-1. The poses are thereafter used for MD simulations with the purpose of determining model

stability and interactions crucial for the stability. The final step is a combination of MD simulations and free energy calculations to determine the binding free energy of the generated systems. Docking and MD simulations are repeated until stability is achieved.

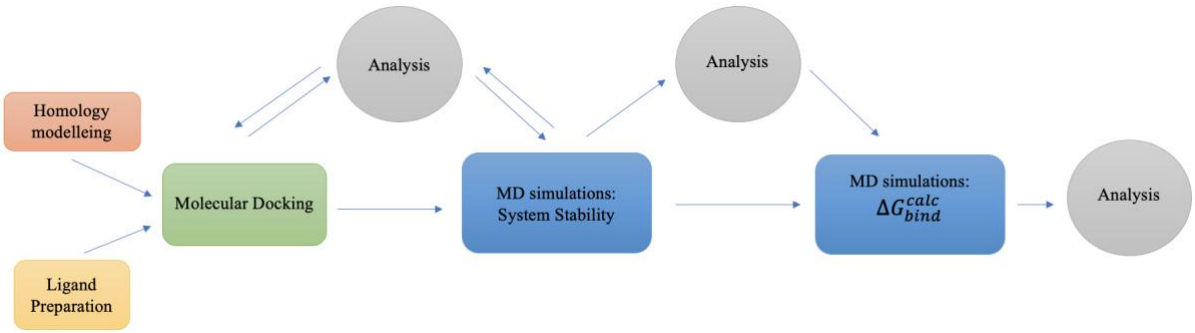


Figure 3. Schematic schedule of the different steps in the method.

2.1 Homology modeling

There is no crystal structure available of IMI-1 on RSCB. Therefore, a homology model of IMI-1 was created. The homology model of IMI-1 was created using the homology modeling program Prime (ver. 2021-4) provided in the Schrodinger Suite (ver. 2021-4). The model was generated based on the amino acid sequence retrieved from UNIPROT by entering the accession number Z21956 [6]. The amino acid sequence was imported to Prime and thereafter Prime automatically searched suggestions of 3D-templates with high sequence identity. The protein 1BUL was the suggestion with highest sequence identity and resolution, therefore 1BUL was chosen as a 3D template for the homology model of IMI-1 [14]. The generated model was validated by analyzing the Ramachandran plot and RMSD. The model was then prepared using the function Protein Preparation Wizard in Maestro. The model was pre-processed by adding missing hydrogens, creating disulfide bonds, and generating charged residues at pH 7.0. The model was thereafter refined with restrained minimization by the OPLS4 force field. The model of IMI-1 was then used in molecular docking with carbapenems as ligands.

2.2 Ligand Preparation

The ligands used in this project were generated by entering the SMILES of the ligands retrieved from PubChem in the 2D-sketcher function in Maestro. Thereafter, the function LigPrep was used to ionize the ligands at pH 7.0 and minimized by applying the OPLS4 force field. The prepared ligands was then docked to the homology model of IMI-1 using automated molecular docking.

2.3 Automated Molecular Docking

Automated molecular docking is solely used in this project and performed using the docking program GLIDE (grid-based ligand docking with energetics) (ver. 2021-04), provided in the Schrodinger Suite. When docking, the interactions and ligand positions are estimated by a molecular docking algorithm. Each interaction/position is provided with a score, hence GLIDE performs two different tasks, namely, conformational search of the ligand and ranking the pose by estimating the corresponding interaction energies by using a scoring function. GLIDE uses GlideScore SP and GlideScore XP and favors lower energy. The difference between the scoring functions is the penalties applied to the score of a pose from any violations, where XP applies harsher penalties than SP.

The ligands imipenem, meropenem and biapenem were docked into the active site of the generated IMI-1 model. Primarily, a docking grid was generated using the function Receptor Grid Generator. The grid constrains the location of the ligands to the active site of IMI-1 and the grid-box was centered at the β-carbon of Ser70. The box size depended on the size of the ligand and the conformation of the protein, to obtain productive poses. Therefore, grid-sizes between 10-14 Å were generated and used when docking each of the ligands. The generated grid was then used in the function Ligand Docking where inversion of ring conformations was excluded and the output of ligands where set to maximum ten poses for either scoring function. The generated poses were evaluated by identifying interactions and by measuring the distance of the targeted interaction between the ligand and Ser70. Shorter distance between Ser 70 and the β-lactam-ring of the ligand was favored. Poses with good score and favored interactions were selected for molecular dynamics (MD) simulations.

2.4 MD – Model System Stability

Selected poses from the docking process were used for MD simulations using the program Desmond provided in the Schrodinger Suite. The purpose of the MD simulations using Desmond is to determine which pose generates a stable system. The function System Setup was used to solvate the system using TIP3P as solvent in a Truncated Octahedron box with buffer distance of 10 x 10 x 10 Å, and minimized by applying the force field OPLS4. The solvated system was used in the task Molecular Dynamics where simulation settings were chosen, and the job-file was written. Simulation time was set to 10 ns, 50 ns and 200 ns depending on what system was simulated. The time step size of the MD simulations using Desmond was 2.0 fs and cutoff radius used was 9.0 Å for the non-bonded interactions and none for the ligand. The recording interval was set to 10.0 ps, the approximate number of frames were set to 1000, and energy output was set to 1.2 ps for all MD simulations performed using Desmond. Ensemble class used was NPT with MTK (constant pressure and temperature), barostat was τ=2.0, thermostat was τ=1.0 and the default pressure was used 1.01325 bar. The temperature for the simulation was set to 298 K, which is the same temperature that was used in experiments for retrieving the kinetic parameters of IMI-1 [6].

The simulations were run in five replicates per pose and the stability of the poses were determined by plotting the Protein-Ligand Root Mean Square Deviation (PL-RMSD). The PL-RMSD function is part of the Simulation Interaction Diagram function of the Schrödinger Suite. The PL-RMSD is used to determine the stability of the system since it is a measure of the average change in displacement throughout the simulation. Therefore, the PL-RMSD can be used to determine large conformational changes in the protein, ligand movement from the active site, change of the initial pose, and duration of poses. The important molecular interactions were identified, and the interaction distances were measured throughout the simulations.

2.5 MD -Binding Free Energy Determination with LIE

Use of the molecular dynamics program Q (ver. 5.06) enables calculation of the binding free energy of biomolecular systems. MD simulations on Q can be performed using spherical boundary conditions (SBC) which is a method that is based on the surface constrained all-atom solvent model (SCAAS) [15-16]. In the SBC model, the system is divided into three subshells; the outermost layer is constrained and resembles bulk water, the middle layer surrounds the inner most layer which contains the location of interest. The outermost layer is constrained in terms of position and polarization, in other words the outermost layer is frozen. The middle and the innermost subshells are unconstrained, representing a real biological system containing

freely moving solvents and residues. Thereby, the SBC model makes it possible to focus on a particular region of the biological system, which makes SBC advantageous when analysing molecular interactions between the ligand and protein in the active site.

The MD simulations performed using Q were all 10 ns and a 1 fs MD time step size was used. The outer sphere was manually set to 30 Å and the middle subshell was set to 24 Å. The simulations were run at 298 K with the OPLS-AA force field. The non-bonded interactions were calculated up to a 10 Å cutoff, except for the ligand atoms for which no cutoff was used. Furthermore, long-range electrostatics were treated with the local reaction field multipole expansion method. The parameters needed for the ligands that were not present in the original version of the force field were retrieved from automatic parametrization performed with Macromodel (ver. 13.3). The topology of the ligands were created by retrieving parameter-files of the ligands from the website LigParGen. The generated parameter-files were thereafter used in combination with the maestro-file and the OPLS4 force to create the topology-files of the ligands. As previously described, using the SBC model means that everything beyond the outer sphere is constrained and frozen. Therefore, polar residues that ended up outside of the sphere boundary were neutralized. In addition, residues in the middle layer were neutralized as well to obtain a net charge of zero for the system, when needed. Different residues were neutralized for the different ligands, and they differ based on the position of the center point of the system (the charged oxygen of the carboxyl-group on the ligands) and the center point differs based on the pose of the ligand. All solvent bonds and angles were constrained with the SHAKE algorithm.

The LIE method was thereafter applied to calculate the binding free energy (ΔG_{bind}^{calc}) of the simulated protein-ligand complexes [16-18]. The premise of LIE is to compare the ligand-surrounding energies in two different environments. The environments are represented by the two states; the ligand free in water and bound to the protein [19]. Both states must be considered and the ligand-surrounding interactions calculated to estimate the total difference in energy for the formation of the ligand-protein complex. Hence, the LIE method requires two MD simulations, one with the ligand bound to the target in a sphere of solvent and the other when the ligand is fixed in the solvated protein. The basis of the method can be described by a thermodynamic cycle shown in *Figure 4* and the expression for binding free energy (ΔG_{bind}) obtained from the thermodynamic cycle can be seen in equation 1.

$$\Delta G_{bind} = \Delta G_{bound}^{polar} - \Delta G_{free}^{polar} + \Delta\Delta G_{bind}^{non\text{-}polar} \quad (1)$$

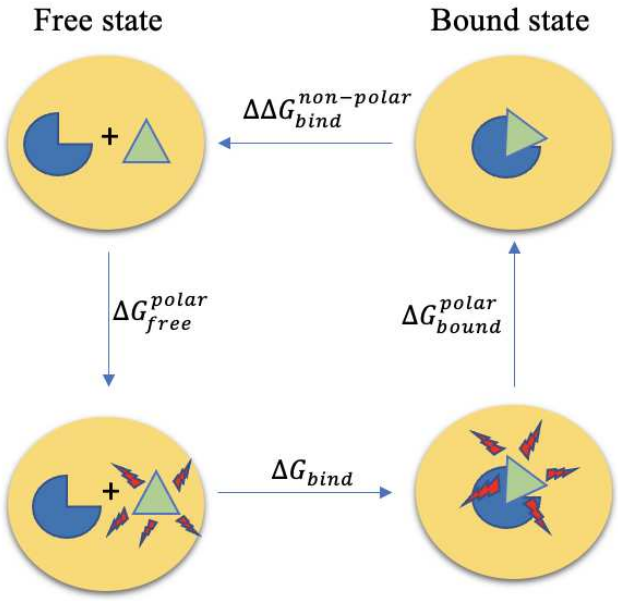


Figure 4. The thermodynamic cycle used for calculating LIE when estimating the binding free energies of complexes with IMI-1 and carbapenems. The blue structure represents the protein, the green triangle represents the ligand and the red lightnings represents the condition where electrostatic interaction is turned on.

The free energies ΔG_{bound}^{polar} and ΔG_{free}^{polar} comes from turning on the ligand-surrounding (l-s) electrostatic interaction energy when the ligand is bound to the target and when it is in the free state. The term $\Delta\Delta G_{bind}^{non-polar}$ represents the difference in energy between the ligand when it is bound to the target and when it is in the free state, as the ligand-surrounding interactions are turned off.

Both simulations, ligand in free state and ligand bound to protein, were run in five replicates where for each run, the l-s energies were gathered. The l-s energies, more specifically, are the non-polar, van der Waal (vdW) and the polar, electrostatic (el) contributions to the total binding energy. Thereafter, the average of the non-polar ($\langle U_{l\text{-}s}^{vdw}\rangle$) and the polar ($\langle U_{l\text{-}s}^{el}\rangle$) interactions were then used in the LIE equation (**equation 2**) to calculate the binding free energy ΔG_{bind}^{calc}.

The LIE equation (**equation 2)** is then used to calculate binding free energy (ΔG_{bind}^{calc}) from the MD-simulations. The LIE equation includes using the average of the non-polar ($\langle U_{l\text{-}s}^{vdw}\rangle$) and the polar ($\langle U_{l\text{-}s}^{el}\rangle$) interactions which are the gathered ligand-surrounding (l-s) energies of the two states, bound to the protein and free in water. The LIE equation also contains the scaling factors α and β, as well as the offset parameter γ.

$$\Delta G_{bind}^{calc} \approx \alpha\Delta\langle U_{l\text{-}s}^{vdW}\rangle + \beta\Delta\langle U_{l\text{-}s}^{el}\rangle + \gamma \qquad (2)$$

$$\approx \alpha\left(\langle U_{l\text{-}s}^{vdW}\rangle_{protein} - \langle U_{l\text{-}s}^{vdW}\rangle_{water}\right) + \beta\left(\langle U_{l\text{-}s}^{el}\rangle_{protein} - \langle U_{l\text{-}s}^{el}\rangle_{water}\right) + \gamma$$

The non-polar term, α, is used for scaling the non-polar interaction energies. It can be determined through the observation that solvation free energy of non-polar molecules is approximately proportional to their size [19]. The empirically determined value of α is 0.18.

Furthermore, the polar term, β, is used for scaling the electrostatic polar interactions. The electrostatic polar interactions are based on the linear response approximation (LRA). However, the ligands in this project are Zwitter ions and there are no well-defined β-values known for zwitter-ions and they are known to be outliers, so the values defined by Almlöf was used in this project [20]. For charged ligands, the value of β is the same [17] [21-22]. However, for neutral ligands (all ligands in this project) β is determined based on the number of hydroxyl-groups. The ligands used in this project are all neutral and contain one hydroxyl-group. Hence, the value of β for all ligands in this project is 0.37.

The offset parameter γ is a system-specific parameter and was used to fix the absolute scale between the experimentally determined and the calculated binding free energies [21-22]. Published values of γ range between 18 and -11. The experimental (ΔG_{bind}^{exp}) was then used to determine γ for the system

To generate a model representing a real biological system and to perform reliable computational experiments the calculated binding free energy (ΔG_{bind}^{calc}) can be compared with experimentally determined binding free energies. In this project experimental data of K_m was used to calculate the experimental binding free energy (ΔG_{bind}^{exp}) using equation (4) [6]. An approximation was made that Km is approximately the same as Ki [19].

$$\Delta G_{bind}^{exp} = -RT \ln K_i \approx RT \ln K_m \qquad (4)$$

3. Results and Discussion

3.1 Homology Modelling

The homology model of IMI-1 (*Figure 5*) was created using Prime and validated with Ramachandran plot and the RMSD (0.17 Å) between the model and the crystal structure 1BUL.

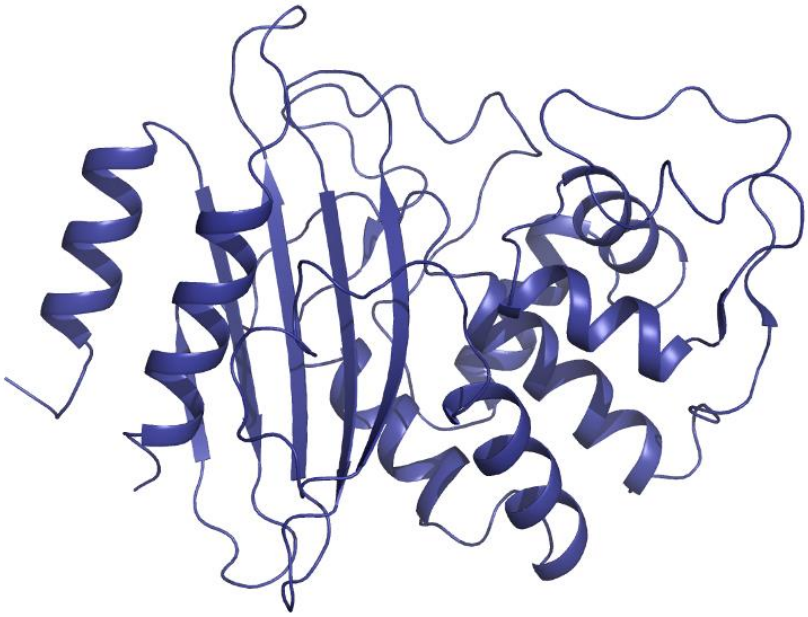


Figure 5. The created homology model based on the amino acid sequence of IMI-1 and the crystal structure 1BUL.

3.2 IMI-1 Flexibility

The homology model of IMI-1 was simulated using Desmond without ligand for 200 ns. The purpose of the simulation was to obtain a converged model and then use the converged conformation for molecular docking. Unexpectedly, the PL-RMSD *(Figure 6a)* of the simulation indicates large conformational movements in IMI-1. Trajectory analysis indicated that movements lead to different conformations of IMI-1, namely, an open conformation and a closed conformation with respect to the active site. Furthermore, the homology model was in between open and closed but more open than closed when generated.

The movement of the entire protein can be seen in *Figure 6a*, which is an overlay of four different frames from the simulation presented in *Figure 6b*. In *Figure 6*, the movement of the protein is consistent, except from the area marked with a blue circle. The two loops located opposite of each other, above the active site, show a distinct movement. The distinguished movement is that the loops move away from each other (open) and then back again (closed) as the simulation continues.

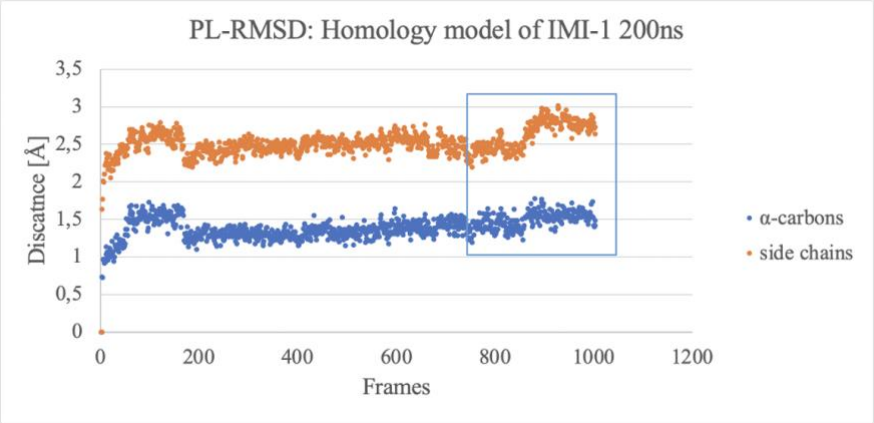


Figure 6a. The PL-RMSD of the homology model of IMI-1. The highlighted area marks the opening process.

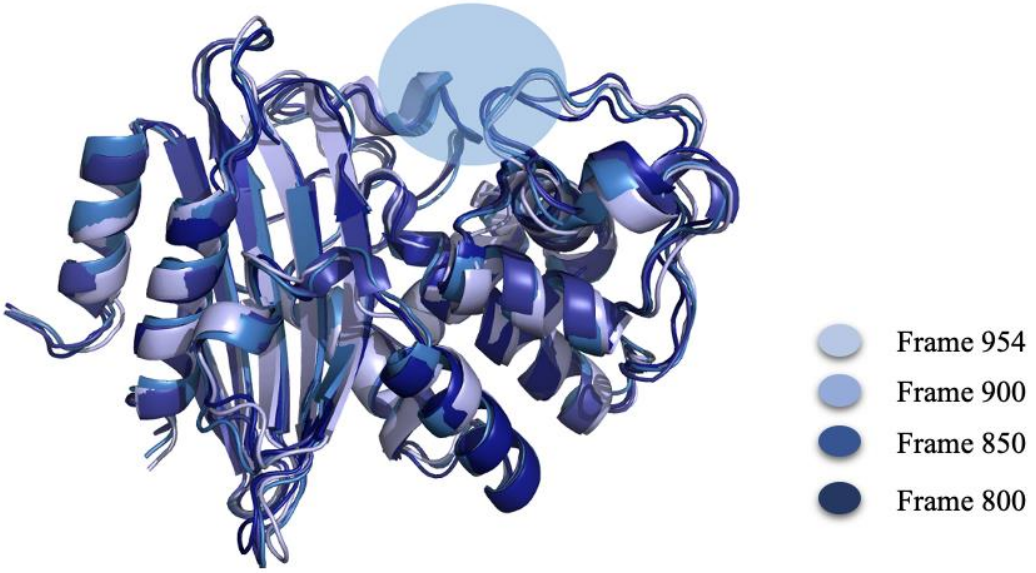


Figure 6b. Four different frames overlayed taken from the highlighted area of 5a, representing the opening process of IMI-1. The blue circle placed on the protein represents where the opening occurs.

The movement of the surrounding residues were analyzed and residues Phe105 and Leu169 had contributing movements to the opening process. During the opening they move away from each other and eventually direct in opposite directions. The distance between the α-carbons of Phe105 and Leu169 varies betweeen 8.3 Å to 14.6 Å and the distance between the outermost atoms throughout the simulation varies between 7.6 Å and 14.9 Å, as shown in *Figure 7*.

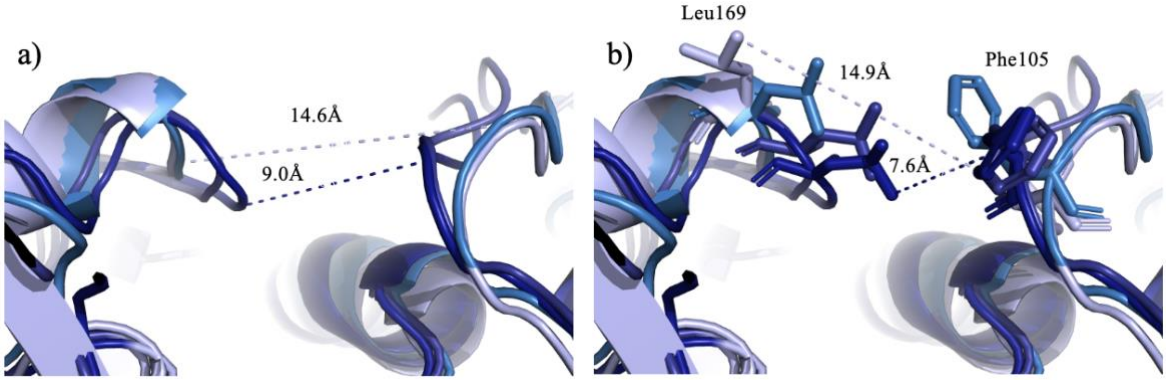


Figure 7. Difference is distance between open and closed conformation. Leu169 and Phe105 a) distance between α-carbons on the flexible loops b) distance between "top" atoms on Leu169 and Phe105 during the opening.

The two conformations of largest distance between residues 105 and 169 were used for molecular docking, *Figure 8*. All three ligands imipenem, meropenem and biapenem were docked into both IMI-1 conformations to identify if the molecular interactions varies depending on which conformation the ligand was docked into. However, it is uncertain which conformation of IMI-1 is found in nature.

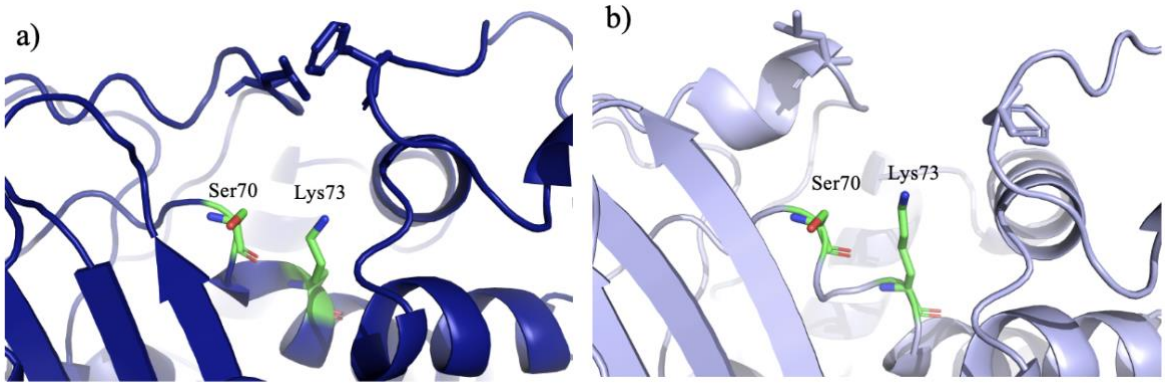


Figure 8. The conformation of the structures used for docking, Ser70-X-X-73 represents the active site; a) the closed conformation and b) the open conformation.

3.3 Docking

The intention of the docking procedure was to obtain poses which facilitated interaction between Ser70 and the oxygen substituent on the β-lactam ring, i.e. the productive pose. The difficulty of obtaining productive poses varied and depended on both protein conformation and ligand. Imipenem only generated a productive pose when docked into the closed conformation, whereas biapenem and meropenem only generated productive poses when docked into the open conformation (*Figure 9a)*. The large R-group of imipenem is projected out into the water when docked in the closed conformation, whereas in the open conformation the large R-group of

imipenem (*Figure 9b*) is curved and points into the protein. Both meropenem and biapenem are rotated 180 degrees in the different conformations of the protein.

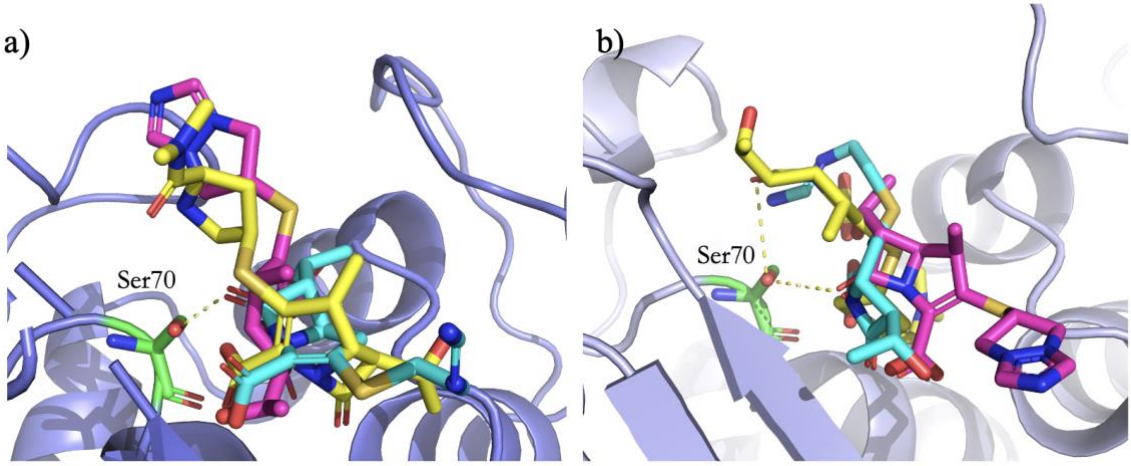


Figure 9a. The poses selected for MD simulation a) Imipenem (blue) meropenem (yellow) and biapenem (pink) docked into the closed conformation. b) Imipenem (blue) meropenem (yellow) and biapenem (pink) docked into the open conformation.

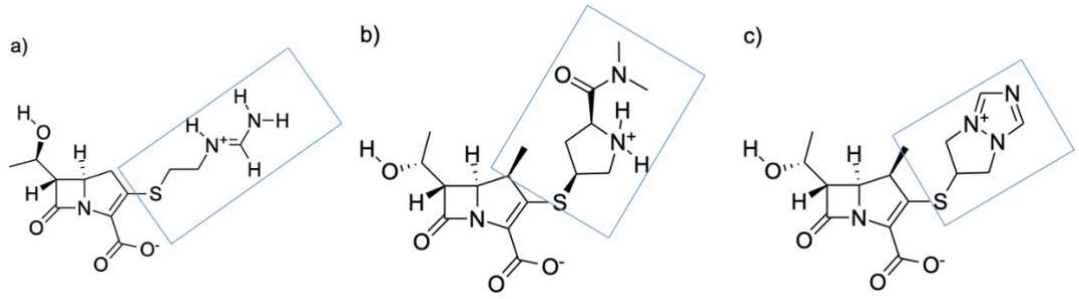


Figure 9b. The rectangle marks the large R-group of a) imipenem b) meropenem and c) biapenem.

3.4 Molecular Dynamics

Generating a stable system of IMI-1 with the carbapenems has been difficult. The aggravating factors are complex and flexible ligands as well as the flexibility in the protein. The opening and closing of the protein affect the ligand-protein interactions and cannot be controlled. Generally, 5-10 different poses (ca four productive) were simulated. Only one stable, productive pose for each ligand was obtained during MD simulations. In total, six stable MD simulations were obtained, three productive and three non-productive. Neither of the ligands generated a stable, productive pose in both conformations of IMI-1. Imipenem has a productive pose in the closed conformation of IMI-1 and a stable non-productive pose in open IMI-1. Whereas meropenem and biapenem has a stable, productive pose in the open conformation of IMI-1 and a stable non-productive pose in the closed conformation (*Figure 11*). A reason for this could be the structural differences of the large R-group. The open conformation of IMI-1 is better adapted to ligands with larger and more complicated R-groups whereas the closed conformation of IMI-1 favours ligands with smaller R-groups.

3.4.1 IMI-1 Flexibility Impact

One of the aggravating factors were IMI-1 flexibility. The first five MD simulation replicates were 50 ns, and usually only one replicate was stable. Thus, an exported frame from the stable

replicate was prepared and replicated, again for 50 ns. This procedure could be repeated 3-4 times for each ligand in both conformations. All five replicates were never stable from either of the rounds due to the flexibility of the protein. The movements were not predictable, meaning the opening and closing process could occur in any replicates in any of the rounds. When the protein opened, the molecular interactions between ligand and protein in the active site changes. For example, biapenem in the open protein conformation is shown in *Figure 10*. The ligand (blue) changes position, loose some important interaction as well as the productive pose as the protein is opening (green/grey).

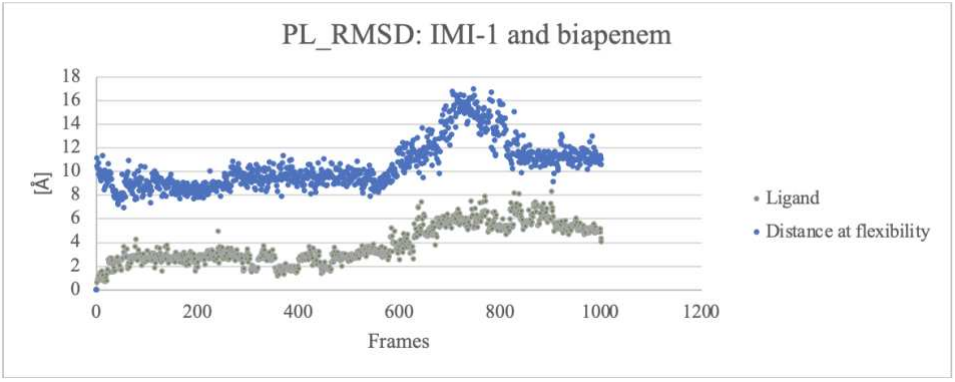


Figure 10. The PL-RMSD of biapenem during 50 ns MD simulation using Desmond and the distance between the flexible loops of IMI-1 during the simulation.

3.4.2 Important Molecular Interactions for Ligand Stability

The important interactions of all six poses are presented in *Table 1,* the poses are presented in *Figure 11* and the important residues in the two IMI-1 conformations are presented in *Figure 12*. The common important interactions between all three productive poses are Ser70 and Arg221. Other common important interactions between productive poses of meropenem and biapenem are Lys235 whereas, Thr236 is a common important interaction between the productive poses of biapenem and imipenem, see *Table 1*. The main contributor to the stable interactions in all three ligands is the carboxyl-group found on all three ligands. It generates stable salt bridges with Arg221, Lys73 and Lys235 depending on the ligand and protein conformation.

One big difference between the ligand-protein interactions when docked into the open and closed conformation is interactions with the large R-group (*Figure 9b*). When the ligands are docked into the open conformation of IMI-1 interactions with the large R-groups can be determined for all ligands, as presented in *Table 1*. Conversely, neither of the ligands have main interactions between the large R-group and the protein when docked into the closed conformation.

The productive pose and the non-productive pose are rotated almost 180 degrees for all ligands (*Figure 10*). The productive poses a), d) and f) all have the large R-group directed out from the binding pocket and into the water. Furthermore, the large R-group of imipenem and meropenem interacts with the protein only in the open form of IMI-1.

Table 1. Presenting the productive pose, important interactions and other specific interactions of the systems studied in this project.

	Protein conformation (*Figure 11*)	Productive pose	Important interactions	Interaction with large R-group of ligands. (*Figure 9b*)	Saltbridge with carboxyl
Imipenem	Closed (a)	Yes	Arg221, Thr236, Ser70, Ser238	No	Yes
	Open (b)	No	Thr236, Ser238, Lys73, Lys235, Ser70, Glu167	Yes	Yes
Meropenem	Closed (c)	No	Ser238, Glu167, Thr236, Ser70, Ser130	No	Yes
	Open (d)	Yes	Lys73, Ser70, Lys235	Yes	Yes
Biapenem	Closed (e)	No	Ser238, Ser70, Lys235, Ser130	No	Yes
	Open (f)	Yes	Lys235, Glu167, Thr236, Ser70, Arg221	Yes	Yes

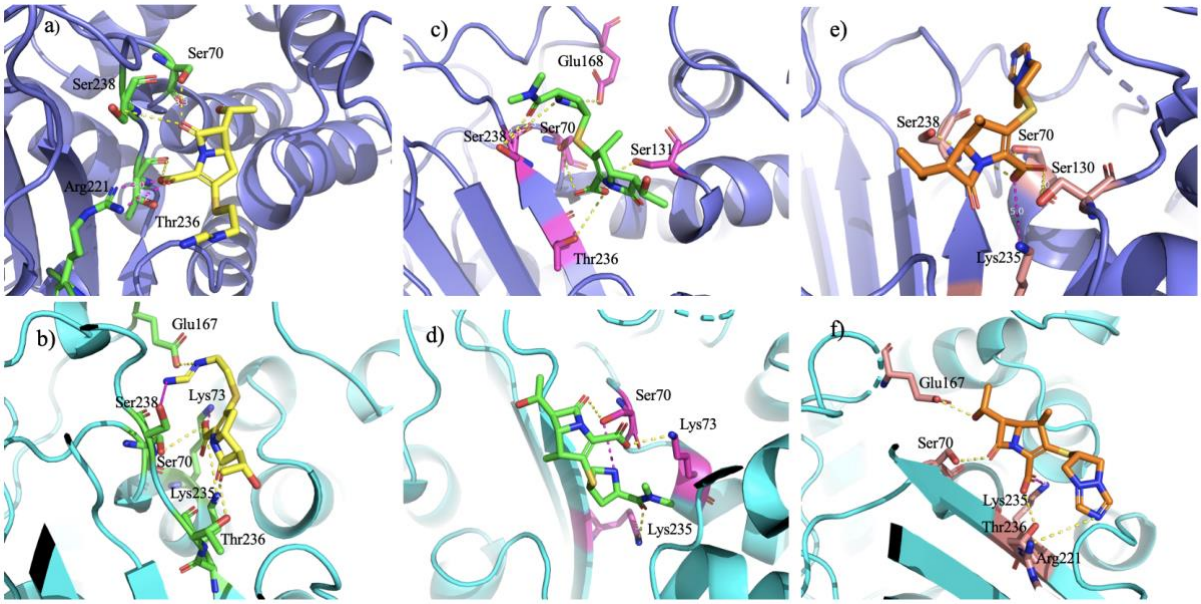


Figure 11. The most important interactions for all stable poses provided from MD simulations in Desmond a) imipenem in IMI-1 closed (productive) b) imipenem in IMI-1 open (non-productive) c) meropenem in IMI-1 closed (non-productive) d) meropenem in IMI-1 open (productive) e) biapenem in IMI-1 closed (non-productive) c) biapenem in IMI-1 open (productive).

As previously mentioned in the introduction, the general active site of a serine β-lactamase is Ser70-X-X-Lys73. In the closed conformation, Lys73 is not part of the important interactions for either of the ligands. However, it is part of the important interactions for imipenem and meropenem in IMI-1 open. Only meropenem in IMI-1 open is the productive pose.

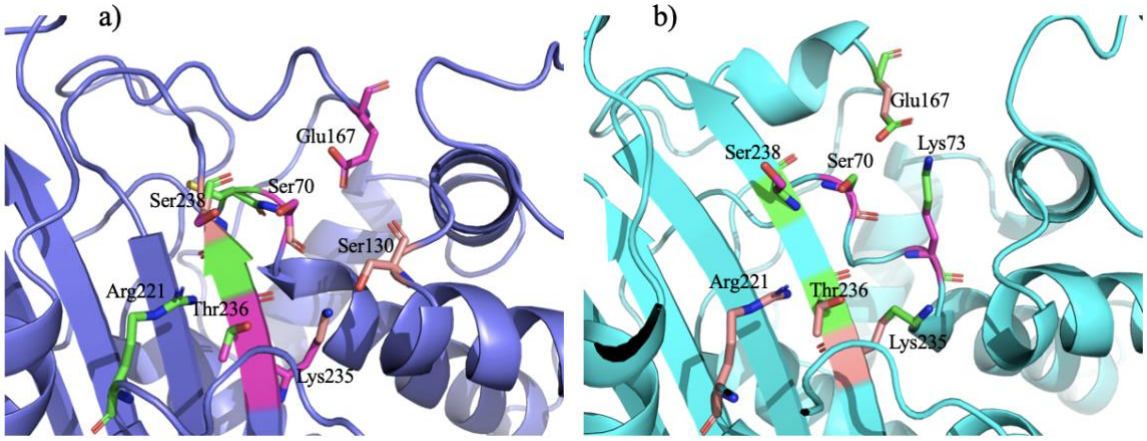


Figure 12. The residues part of the important interactions of imipenem (green), meropenem (pink) and biapenem (pale orange) in a) IMI-1 closed conformation and b) IMI-1 open conformation.

3.4.3 Binding Free Energies

The binding free energies varies from ligand to ligand, but also between protein conformation from different replicates of the same ligand. The offset parameter γ is not system specific in this case due to the flexibility of IMI-1 and flexibility of the ligand. Additionally, all ligands in this project are zwitter ions. Therefore, γ has not been determined in this project. If more time was available, an individual γ for each of the six simulations would have been determined.

The electrostatic interactions vary throughout the simulations and between replicates, especially for the water runs of imipenem and meropenem. The electrostatic interactions vary due to the ligand's flexibility, allowing intramolecular interactions (*Figure13*). Imipenem occurs in three different poses throughout the simulations while meropenem occurs in two different poses. Neither imipenem nor meropenem stays in the different poses for more than around 0.4 ns therefore, the average of the electrostatic interaction energies could be used when calculating the binding free energies. Biapenem has a rigid structure and thereby has one stable pose throughout the simulations in free state solvated in water.

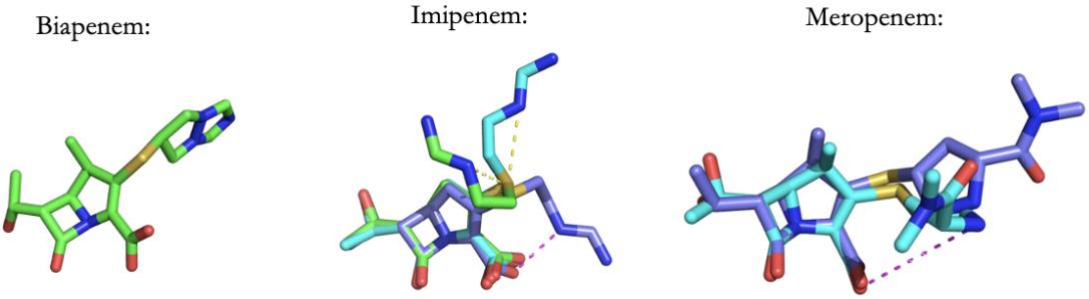
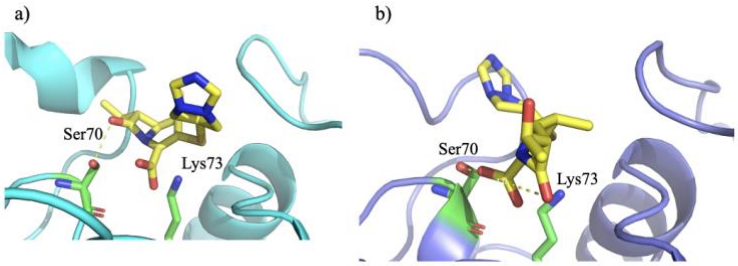


Figure 13 Different poses of the ligands when simulated in free state solvated in water. Three distinct conformations of imipenem in water runs. The yellow line between represents the interaction between N and S, the pink line represents the ionic bond between N and O. Two

distinct states of meropenem where the pink line represents ionic interaction between the N and O.

Biapenem

The calculated binding free energy of biapenem over the replicates indicates one stable pose of the ligand when docked into the open conformation of IMI-1 and another pose when docked into the closed conformation of IMI-1. The productive pose is observed when docked into the open conformation (*Figure 14*). The different poses and the calculated binding free energy as well as the experimental binding free energy is presented in *Figure 14.* As can be seen in the figure, biapenem is rotated 180 degrees in the different conformations and the most energetically favored (*Figure 14*) is non-productive. If γ would have been calculated, it would have been around -5 kcal/mol.

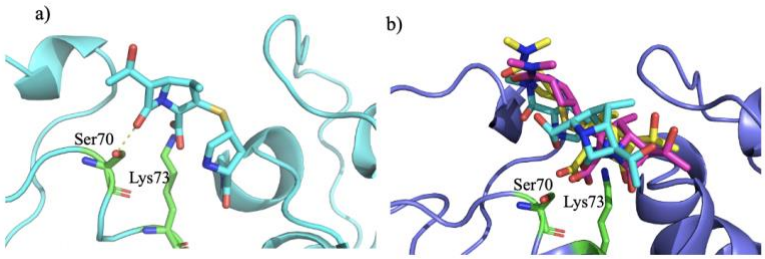


	IMI-1 conformation	ΔG_{bind}^{calc} [kcal/mol]	ΔG_{bind}^{exp} [kcal/mol]
Meropenem	Open (a)	2.21	-6.13
	Closed (b)	-1.26	

Figure 14. Presents the different poses of biapenem in open and closed, retrieved from simulations using Q, at the specific binding free energy presented in the table a) corresponds to the pose in the open conformation of IMI-1 (productive) and b) corresponds to the pose in closed conformation of IMI-1 (non-productive).

Meropenem

Meropenem in open IMI-1 conformation demonstrate one distinct pose, which is the productive pose. Furthermore, it is the only productive pose which has important interactions with both Ser70 and Lys73 (*Figure 15*). However, the calculated binding free energy is high which indicates that the pose is unlikely to appear in nature even though it is stable when simulated. Furthermore, meropenem show three distinct energy levels when simulated bound to the closed protein (*Figure 15*). The highest energy level (blue in *Figure 15*) has a different pose than the other two more energetically favored and is probably not stable in nature even though it is stable during a MD simulation of 10 ns. The yellow pose of meropenem (*Figure 15*) is the most energetically favored pose of them all, however, it is not the productive pose. Moreover, if γ for biapenem is used for meropenem, then the calculated binding free energies would still be significantly different from the experimental binding free energy. Hence, γ is not system specific in this case.

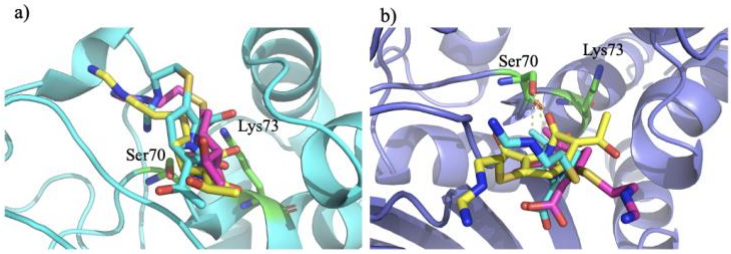


	IMI-1 conformation	ΔG_{bind}^{calc} [kcal/mol	ΔG_{bind}^{exp} [kcal/mol
Meropenem	Open (a)	23.4	-6.25
	Closed (a)	-19.5	
		-0.41	
		9.91	

Figure 15. Presents the poses of the different energy levels of meropenem in open and closed conformation of IMI-1, retrieved from MD simulations using Q a) Open conformation of IMI-1 energy lever 23.4 kcal/mol (blue) which is the productive pose and b) closed conformation of IMI-1 energy lever -19 kcal/mol (blue), energy lever -0.41 kcal/mol (pink) and energy level 9.91 kcal/mol (blue).

Imipenem

Imipenem shows three different energy levels separated by ca 10 kcal/mol when simulated bound to the protein (*Figure 16*). The productive pose is observed when imipenem was docked into the closed conformation of IMI-1 (*Figure 16*). A distinctive difference between imipenem in open and imipenem in closed, independent of energy level, is that they are rotated almost 180 degrees. The results in *Figure 15* indicates that all three energy levels, the productive pose b), is located further into the binding pocket. Comparison of the calculated binding free energies indicate that imipenem undertakes most poses with energetically favored interactions in general.



	IMI-1 conformation	ΔG_{bind}^{calc} [kcal/mol	ΔG_{bind}^{exp} [kcal/mol
Imipenem	Open (a)	-18,5	-5,14
		-10,9	
		1,50	
	Closed (b)	-11,9	
		-5,86	
		0,591	

Figure 16. Presents the poses of the energy levels of imipenem in open and closed conformation of IMI-1, retrieved from MD simulations using Q. a) Open form of IMI-1, energy level -18.5 kcal/mol (yellow), energy level -5.85 kcal/mol (pink), energy level 1.50 kcal/mol (blue) b) closed form of IMI-1, energy level -11.9 kcal/mol (yellow), energy level -5.86 (pink), energy level 0.591 kcal/mol (blue).

3.5 Conclusion

The purpose of IMI-1 is to form a complex with the antibiotics and perform hydrolysis. Conversely, the high obtained calculated binding free energies of the productive poses and the poor stability of the ligands in the active site indicates that the catalytical process is much faster than binding mechanism in this system. In other words, it is probably not necessary for the ligand to bind very long before it is catalyzed. Although the calculated binding-free energies are so high, the ligands remain in the binding pocket. Furthermore, imipenem is held in place when the binding pocket closes, suggesting a different binding mechanism than that of meropenem and biapenem.

Based on the docking-studies and analysis of ligand-protein interactions seen in the simulations, several possible binding mechanisms, depending on the ligand have been determined. This indicates that promiscuity may be the result of multiple binding modes in the enzyme.

If more time was available, further structural differences of the energy levels would have been examined for imipenem and meropenem. Additionally, a suggestion for further research is to study how to determine an accurate β for these ligands and zwitter ions generally. Furthermore, it would be interesting to determine the IMI-1 flexibility in other carbapenemases, and other classes of β-lactamases. Also, to determine whether the flexibility similarly affects the mechanism of inhibition by inhibitors available on the market. Lastly, determine how the acylated ligands interact with IMI-1, are they stable, or will they leave the binding pocket immediately?

In conclusion, despite the difficulty of modeling the IMI-1 system, calculation methods are a valuable tool for gaining a better understanding of how these enzymes look and work.

4. Acknowledgments

I would like to thank my family for support and patience. Foremost, I would like to thank my supervisor Yasmin Binti Shamsudin for the opportunity to learn, patience when I was learning, for valuable guidance and positive thinking during this project.

I would also like to thank my research group for the support, time, and the laughter that we have shared during this project.

5. References

[1] González-Lamothe, R., Mitchell, G., Gattuso, M., Diarra, M., Malouin F., Bouarab, K., *Int. J. Mol. Sci.* 2009, *10(8)*, 3400-3419.

[2] Murray, CJ., Ikuta, KS., Sharara, F., Swetschinski, L., Robles, Aguilar, G., Gray A., *The Lancet,* 2022, *399*, 629-655.

[3] Zaffiri, L., Gardner, J., Toledo-Pereyra, LH., *J Invest Surg*., 2012, *25(2)*, 67–77.

[4] Blair, JMA., Webber, MA., Baylay, AJ., Ogbolu, DO., Piddock, LJV., *Nat Rev Microbiol,* 2015,*13(1)*, 42–51.

[5] Bush, K., Jacoby, GA., *Antimicrob Agents Chemother*, 2010, *54(3)*, 969–76.

[6] Rasmussen, BA., Bush, K., Keeney, D., Yang, Y., Hare, R., O'Gara, C., *Antimicrob Agents Chemother*, 1996, *40(9)*, 2080–2086.

[7] He, Y., Lei, J., Pan, X., Huang, X., Zhao, Y., *Sci Rep.* 2020, *10(1)*.

[8] Chouchani C., Marrakchi R., and El Salabi A., *Crit. Rev. Microbiol.*, 2011, *37*, 167–177.

[9] Gaibani, P., Ambretti, S., Berlingeri, A., Gelsomino, A., Bielli, A., Landini, M, P., Sambri, V., *Euro Surveill.* 2011, *16*, 19800

[10] Gopalakrishnan, R., Sureshkumar, D., *J. Assoc. Physicians India,* 2011, 58, 25–31

[11] Livermore, D, M., Warner, M., Mushtaq, S., Doumith, M., Zhang, J., *Int. J. Antimicrob. Agents*, 2011, *37*, 415–419.

[12] Venkatesan, AM., Agarwal, A., Abe, T., Ushirogochi H., Yamamura, I., Ado, M., *J Med Chem* , 2006, *49(15)*, 4623–4637.

[13] Sehgal, SA., Adnan, Tahir, R., *J Theor Comput Sci*, 2014, *1(3)*.

[14] https://www.rcsb.org/structure/1BUL (Date accessed 2022-02-11)

[15] King, G., Warshel, A., *J. Chem. Phys*. 1989, *91(6),* 3647–3661.

[16] Åqvist, J., Medina, C., Samuelsson, J.-E. *Protein Eng Des Sel* 1994, *7,* 385-391.

[17] Hansson, T., Marelius, J., Åqvist, J. *J Comput Aided Mol Des* 1998, *12,* 27-35.

[18] Shamsudin, Khan, Y., Computational methods for calculating binding free energies of ligands in COX-1 [Internet] [Licentiate dissertation]. 2014. Available from: http://urn.kb.se/resolve?

[19] Almlöf, M., Carlsson, J., Åqvist, J. *J. Chem. Theory Comput.* 2007, *3*, 2162-2175.

[20] Åqvist, J., Hansson, T., *J. Phys. Chem.* 1996, *100*, 9512-9521.

[21] Carlsson, J., Boukharta, L., Åqvist, J., *J. Med. Chem.* 2008, *51*, 2648-2656.

[22] Mishra, S. K., Sund, J., Åqvist, J., Koča, J., *J. Comput. Chem.* 2012, *26*, 531-568.