Master Thesis in Pharmacology

Theoretical and experimental studies of Thermolysin inhibition

Yimingjiang Wuxiuer
May 2008

Department of Pharmacology
Institute of Medical Biology
Faculty of Medicine
University of Tromsø
Norway

ACKNOWLEDEMENTS

I had the great good fortune of learning from faculty members who are both accomplished scholars and dedicated teachers. Each of them contributed to my interest in Phamacology and molecular modeling. But two of them stand out in my memory as patient, tolerant mentors and intellectual role models: Professor Ingebrigt Sylte and Stipendiat Mahmud Tareq Hassan Khan. Professor Ingebrigt encouraged me as a young master student and quite literally introduced me to the world of molecular modeling. Mr. Tareq showed me what it means to be an experienced, flexible, energetic, diligent and self motivated scientist. Without whom, I would not have had one paper published (as a second author). I can no longer estimate the number of points in the process of writing and research where the comments of them led to significant improvements in clarity and more meaningful treatments of the project. Their influence on my life was profound, and I thank them deeply. I am no less indebted to other colleagues and teachers of the molecular modeling group at the Department of Pharmacology at the University of Tromsø. I would especially like to thank Osman Gani, Olayiwola Adedotun Adekoya and Kurt Kristiansen for giving me advice and helping me in the computer graphics room. I'm very lucky to work and study in such a resourceful, efficient and productive team.

I would also like to thank my parents and my brothers, for all of their love and support.

Thank you!

TABLE OF CONTENTS

SUMMARY

Zinc-metalloproteinases play a key role in the biosythesis and metabolism of different bioactive peptides. As a member of zinc-metalloproteinases, thermolysin has served as a model system to study the inhibition mechanism of other metalloproteinases. Inhibitors of thermolysin have considerable potential as therapeutic agents. In the present master thesis, docking calculations were performed and reported for 25 potent non-peptidal thermolysin inhibitors retrieved from literatures. Docking software ICM^{TM} was used. 8 out of 25 compounds were from X-ray crystal structure complexes with thermolysin. ICM^{TM} reproduced 63% of the binding modes from 8 X-ray crystal complexes. Experimental assay were done with MS, R and M compounds. 2 out of 13 MS compounds were found to inhibit thermolysin (one was later found to be competitive inhibitor and the other one was found to be a non-competitive slow inhibitor). 12 out of 37 R compounds inhibited thermolysin with varying affinity and 2 out of 7 M compounds had thermolysin inhibition property. One of them was found to be very potent inhibitor, with IC_{50} value of $4{,}411 \times 10^{-11}$ mM. Docking calculations were performed with all MS and active R compounds to predict the binding poses. However, the experimental verification showed that only two of the MS compounds are thermolysin binders. For R compounds, the correlation between binding energies and IC_{50} is not linear, which indicates that the docking poses were not yet 'perfect'. The presence of metal ions as zinc is a problem for docking studies. Parameterization of zinc ion need to be taken directly into account by the docking program, and special treatment may be necessary.

ABBREVIATIONS

ACE	Angiotensin-converting enzyme;
CATH	Classes/ Architecture/ Topology / Homologous Superfamily;
CoV	Coronavirus;
DMF	Dimethyl formamide
EC	Enzyme Commission;
ECE	Endothelin converting enzyme;
ECE-I	Endothelin converting enzyme I;
ECM	Extracellular matrix;
EMBL	European Molecular Biology Laboratory;
FAGLA	FA-glycyl-L-leucine amide
HTS	High-throughput screening;
IC_{50}	The 50% inhibitory concentration of a substance
ICM	Internal coordinate mechanics;
KD	Dissociation constant;
K_i	Inhibition constant;
LMWP	Low molecular weight protamine;
MD	Molecular Dynamics;
MEROPS	An information resource for peptidases
MMPs	Matrix metallopeptidases;
MMFF	Merck Molecular Force Field
M_r	Relative molar mass;
NEP	Neprilysin;
NMR	Nuclear Magnetic Resonance;
NP	*Bacillus subtilis*
PDB	Protein Databank;
RMSD	Root mean square deviation;
SARs	Structure-activity relationships;
SARS	Acute respiratory syndrome;
SFs	Scoring functions;
TLN	Thermolysin;
TLPs	Thermolysin -like peptidases;
VLS	Virtual ligand screening;

VAB	Vibriolysin of Antartic *bacterium*;
ZBG	Zinc binding group

1 INTRODUCTION

1.1 Proteinases in general

In 1946 Linus Pauling first formulated the basic principle underlying enzyme catalysis, namely, that an enzyme increases the rate of chemical reaction by binding and stabilizing the transition state of its specific substrate tighter than the ground state [1]. The presence of enzymes may increase the biological catalyzed reaction by factors of 10^{15}, while enhancements in the range 10^{3}-10^{9} are more typical [2]. Proteinases are widely distributed enzymes in nature, where they perform a variety of different functions. They selectively catalyze the hydrolysis of polypeptide bonds. In higher organism, proteinases are involved in regulating physiological processes such as control of blood pressure, immunological defense, wound healing, digestion, fertilization, differentiation, growth, cell signaling and migration, apoptosis, and blood clotting [1]. All the well-characterized proteinases so far, belong to one or other of families classified by MEROPS (http://MEROPS.sanger.ac.uk/). They are aspartic, cysteine, glutamic, metallo-proteinases, serine, threonine proteinases and others of unknown catalytic type [3]. This classification is based on a functional criterion, namely, the nature of the most prominent functional group in the active site. Proteinases are also classified by clan. According to MEROPS, a clan contains all peptidases from a single evolutionary origin. It represents one or more families that show evidence of their evolutionary relationship by having similar tertiary structures, or by the order of catalytic-site residues in the polypeptide chain and often by common sequence motifs surrounding the catalytic residues [3]. Clan MA contains a variety of metallopeptidases (Table1.1). The metallopeptidases in family M2,M4 and M13 have been subjects of particularly intense research.

Table 1.1: Types of pepidase activity within MA clan.

MA	Aminopeptidase A(M1, M61)
	Carboxypeptidase (M2, M32)
	Peptidyl-dipeptidase (M2)
	Oligopeptidase (M3, M13)
	Endopeptidase (M4, M10, M12 and others)

1.2 Metalloproteinases

Zn^{2+} containing metalloproteinases constitute an expanding list of structurally related proteases which are widely distributed in nature. They are involved in biochemical events of extreme importance, such as digestion (carboxypeptidase A, astacin), tissue remodeling and extracellular matrix (ECM) degradation (matrix metalloproteinases, MMPs), blood-pressure regulation (neprilysin), formylation and deformylation in bacterial protein synthesis (peptide deformylases), etc. [4]. In last decades, zinc peptidases with known 3D structure have attracted increased attention for structure-based design of drugs [5, 6] used for treatments of a number of diseases such as infectious disease, hypertension, tumor invasion, arthritis and bone destruction. Metalloproteinases are the most diverse of the eight main types of proteinases, with 76 families identified so far [3].

1.3 TLN family

A family is a set of homologous peptidases. The proteins that have evolved evolutionarily from a common ancestor are said to be homologous [7]. Although the sequence similarity in homologous proteins is less preserved, 3D structures of homologous proteins have been remarkably conserved during their evolution, because the common structure is crucial for the specific function of the protein [1]. Thermolysin (TLN, EC 3.4.24.27) belongs to a family of metalloproteinases [3] that chemically function at neutral pH [8]. It is a bacterial endopeptidease of M_r 34,600 containing 316 residues in its single polypeptide chain [9].

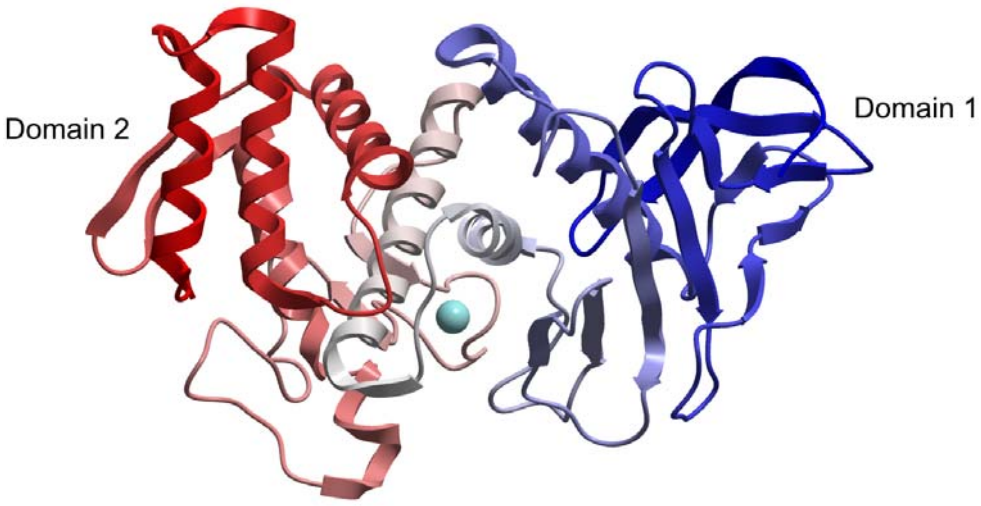


Figure 1.1: TLN structure in ribbon mode (PDB: 1gxw, 316 residues), domain 1 (blue) and domain 2 (red).

TLN is an endopeptidase rather than an exopepetidase. This is manifested in the nature of the binding site, instead of a deep pocket as in carboxypeptidases [10], TLN has an open extended cleft that can bind the polypeptide chain of the substrate on both sides of the peptide bond which would be cleaved (this bond named as scissile bond) [9, 11]. According to CATH (Classes/ Architecture/ Topology / Homologous Superfamily) protein structure classification (http://www.cathdb.info/latest/index.html), this wide cleft separates TLN into two large domains (Figure 1.1). Table 1.2 shows more detailed classification based on secondary structure elements (α helix and β strand).

Table 1.2: The CATH Structure classification of TLN [12] (PDB: 1gxw).

Domain	CATH NO	Class	Architecture	Start Res	Stop Res	Length
1	3.10.170.10	αβ	Roll	6	154	149
2	1.10.390.10	Mainly α	Orthogonal Bundle	155	315	161

Ramachandran plots of TLN (Figure 1.2) show that 96.5% (303/314) of all residues are in energetically favored light blue regions [13]. Further more, it is evident that higher proportion of amino acids is distributed in the lower left quadrant which refers to right handed α-helix. Whereas amino acids fall in the upper left quadrant form the β strands, another major secondary structure elements.

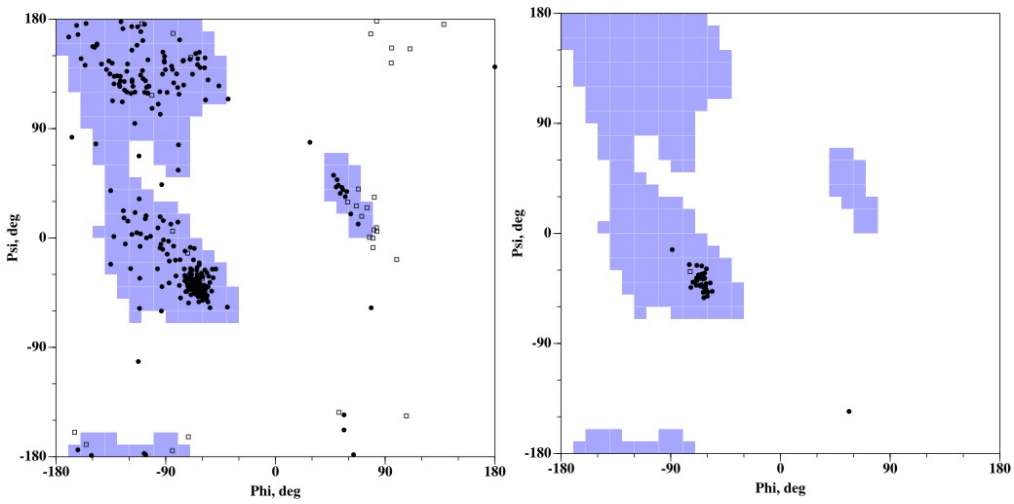


Figure 1.2: Ramachandran plots showing the preferred combinations of the torsion angles phi (Φ) and psi (Ψ) for the positions of the amino acids of TLN (Left) and two major helix around active site (Right). (PDB: 1gxw)

There are some important features in the TLN active site (shown in Figure 1.3).

- The active site zinc ion is coordinated by three amino acid residues and a water molecule at two alternative positions [14]. These amino acids include two histidines (His142, His146) and a glutamic acid (Glu166).
- In addition, one glutamic acid residue (Glu143) and one histidine residue (His231) interact with a water molecule at the active site and are required for catalysis [15, 16]

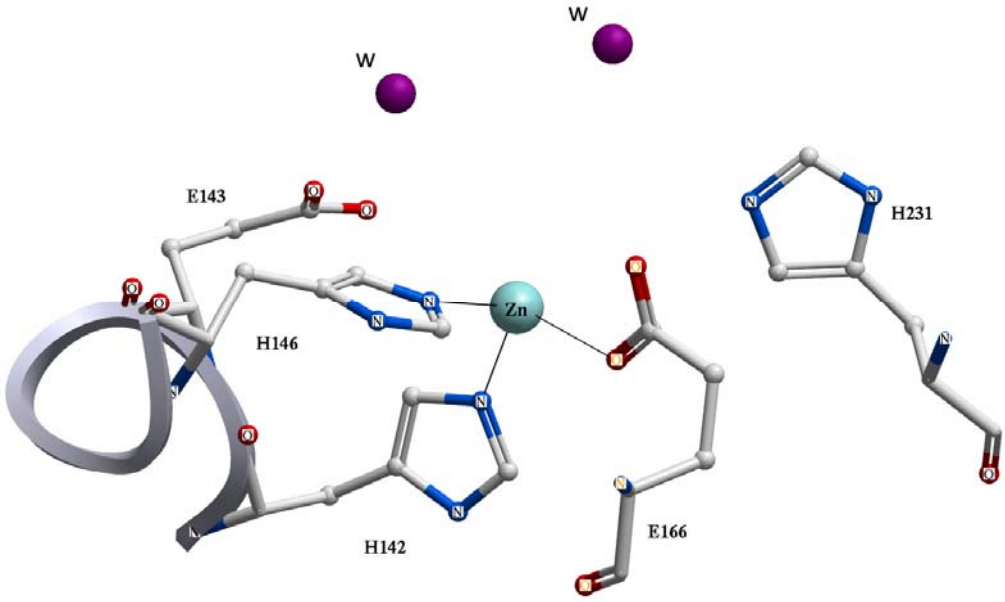


Figure 1.3: Catalytic site of TLN indicating the water molecule at two alternative positions.

The crystal structure complex of the enzyme with one thiocyanate ion in the active site, has been solved [17] (PDB code; 1gxw). The structure of TLN contains a single catalytic zinc ion that is essential for hydrolytic activity, and four calcium ions that are required for thermostability [18]. This extracellular endopeptidase catalyses hydrolysis of the peptide bond specifically on the imino side of large hydrophobic residues [19, 20], in particular leucine, isoleucine and phenylalanine.

According to the results from European Molecular Biology Laboratory (EMBL) www.ebi.ac.uk , TLN active site reside between helix 3 and helix 4 (Figure 1.3). Helix 3 starts with Ile137, ends with Tyr151. Both of the helices fall in lower left quadrant of the Ramachandran plot (Figure 1.2, Right), which indicates highly conserved sequence residing in energetically favorable region [1]. Helix 4 starts with Asn159, ends with Tyr179. Two kinds of turns, β-turn (Tyr151-Gly154) and γ-turn (Ile156-Gln158) connect these two helices. Turns play an important role in globular proteins from both structural and functional points of view [1]. A polypeptide chain cannot fold into a compact structure without the component of

turns. In TLN, these turns occur on the exposed surface and hence form the side of peptide loading grove.

The TLN family is the proteinase family M4 of the MA clan of metalloproteinases [3]. Other TLN-like metalloproteinases (TLPs) such as, neprilysin (NEP) [21], angiotensin converting enzyme (ACE), and endothelin converting enzyme I (ECE-I) [22] regulate the cardiovascular system in human physiology. These enzymes do not belong to the M4 family but have huge mechanistic and structural similarities with TLN as well as other metalloproteinases. Characterization with peptide substrates and high performance liquid chromatography analysis of β-casein digests shows that the M4 family is a homogeneous family in terms of catalysis, even though there is a significant degree of amino acid sequence variation [23]. They posses a consensus sequence HEXXH that constitutes the zinc-containing catalytic domain [5, 24]. Based on similarities on catalytic site mentioned above, TLN have served as a test vehicle to identify proposed inhibitor interactions within the active site of zinc-metalloproteinases [5, 25].

1.3.1 Therapeutic possibilities of TLN like enzymes

Among the enzymes associated with human pathogenic fungi and bacteria, TLPs seem to play a predominant role during pathogenesis and cause increases in vascular permeability, hemorrhagic edema, and sepsis [26, 27]. TLPs of the M4 family such as aureolysin, pseudolysin, and bacillolysin are virulence factors of diverse bacterial pathogens [28]. They promote development within the infected host, and they are used to suppress or avoid its innate immune system [28], while TLN functions intracellulary for bacterial nutrition purposes [29]. A recent study of the cell entry mechanism of Acute respiratory syndrome-coronavirus (SARS-CoV) reveals that proteases such as trypsin and TLN facilitated a 100- to 1,000-fold higher efficient SARS-CoV infection than without these proteases [30]. TLN like proteinases are also key factors in the pathogenesis of various diseases, including several types of bacterial infections [31-33], cholera [34], gastritis and peptic ulcer [35] and gastric carcionoma [36]. Since particular metalloproteinases associated with human pathogens have been recognized as prominent virulence factors, their therapeutic inhibition has become a novel strategy in the development of second-generation antibiotics [37, 38]. On the other hand, in an attempt to reduce severe adverse reactions of protamine in cardiovascular surgeries, Lee [39] and his colleagues developed *in vitro* low molecular weight protamine (LMWP) as a potentially effective and less toxic heparin antagonist by using enzymatic digestion of

protamine with TLN. Animal test showed that it could completely neutralize all these different anticoagulant functions of heparin in dogs. These preliminary findings indicated that LMWP could potentially provide an effective and safe means to control both heparin- and protamine-induced complications[39].

1.3.2 Industrial applications

Natural enzymes are attractive as catalysts in industrial processes. They are often more effective and selective, and produce more pure products with minimal wastage. Proteinases of the TLN family have many unique characteristics with a huge potential for industrial processes. Besides a purely physiological activity, TLN has also been found to be a useful catalyst in protein engineering [40, 41].

Enzyme characteristics of TLN family include,

- Thermophilic enzymes with catalytic activity at temperatures up to 80°C;
- Enzymes that exhibits maximum activity under extreme alkaline conditions [42];
- Enzymes with high stability in organic solvents such alcohols [43]; and enzymes being unique due to their strict substrate specificity [44].

A bioinformatic study of Vibriolysin from *Antartic bacterium* strain 643 (VAB) has shown that VBA has the structural features of a cold adapted protein [45]. Several enzymes of the TLN family are already in industrial use. TLN is used as a bio-catalyst in the synthesis of the artificial sweetener aspartame, as a peptide and ester synthetase, and as a non-specific proteinase to obtain fragments for peptide sequencing [46-48]. Vimelysin from *Vibrio* str.T1800, has a potential application in peptide condensation reactions [49], while vibriolysin from *Vibrio proteolyticus* is used in the production of aspartame and for removal of necrotic tissue from wounds such as burns or cutaneous ulcers [50].

1.3.3 A model system for other metalloproteinases

The zinc-metalloproteinases secreted by the gram-positive thermophilic bacterium *Bacillus thermoproteolyticus* [3] is the prototype of the TLN family and has served as a model system to study the inhibition mechanism of other metalloproteinases. Crystallographic data for TLN and various TLN-inhibitor complexes have been used in efforts to model the active site of other TLN-like enzyme [25, 51]. Because Zinc-metalloproteinases play a key role in

the biosythesis and metabolism of different bioactive peptides, inhibitors of these enzymes have considerable potential as therapeutic agents [52].

Angiotensin I-converting enzyme (ACE, EC 3.4.15.1) belongs to the M2 family of Zinc-binding metalloproteinases, within the MA clan [3]. ACE plays an important role in blood pressure homeostasis by cleaving the C-terminal dipeptide from angitensin I to produce the potent vasopressor peptide angiotensin II [53]. In addition, ACE inactivates the vaso dilatory peptide bradykinin by the sequential removal of two C-terminal dipeptides [54]. ACE contains two Zinc-coordinating catalytic domains (N and C domains) each bearing the HEXXH motif where the two histidines form two of the three amino acid ligands, while a glutamate 24 residues downstream forming the third ligand [55]. Apart from the HEXXH motif, root mean square deviation (RMSD) of the Zinc-binding sites between TLN and ACE is 0.52Å [56]. The role of a Zn ion in ACE catalysis was thought to be analogous to that in TLN [57]. As a consequence of these structural and functional similarities and the role of ACE in the metabolism of those two vaso active peptides, the active site of TLN has been used as a model to develop highly potent and specific ACE inhibitors. These inhibitors have been used as orally active drugs in the treatment of hypertention and congestive heart failure [58].

Neprilysin (neutral endopeptidase; NEP, EC 3.4.24.11) is an integral plasma membrane ectopeptidase of the M13 family of Zinc metalloproteinases [3]. Like TLN and ACE, NEP possesses the HEXXH motif as well as a consensus sequence EXIXD in which the glutamate (Glu) serves as the third Zn ligand [55]. NEP is involved in the metabolism of a number of regulatory peptides of the mammalian cardiovascular, inflammatory, nervous and immune systems [59]. So, NEP is potential therapeutic targets in cardiovascular and inflammatory disorders. Before the crystal structure of NEP-inhibitor complex was solved, the structure and function similarities between NEP and TLN served as the basis of NEP inhibitor design by using TLN as a test model. Selective inhibitors, such as phosphoramidon have contributed to understanding NEP enzyme function [52].

The merits of modeling proteins based on homology to a known structure have been applied a decade ago [60], and it is clear that model building by homology is a valuable technique to construct three dimensional coordinates of protein structures when the sequence identity between the model and the structure is high [7]. With 47% sequence identity with the neutral protease of *Bacillus subtilis* (NP-sub), TLN was used as main structural template to build and optimize a three-dimensional model of NP-sub in associated with site-directed mutagenesis techniques [61].

1.4 TLN and ligand interactions

1.4.1 The catalytic site of TLN and HEXXH motif

TLN is one of the most studied metalloenzymes and, for a number of free enzyme and enzyme-inhibitor complexes, detailed X-ray structures have been obtained. The first was published in 1972 [62]. These structures show that the zinc-binding residues are His142, His146 and Glu166, while Glu143 act as the catalytic residue (Figure 1.4). There is a distorted tetrahedral coordination of the Zn ion of TLN, where one position (the fourth) is exposed to solvent water and turns out to always be present for catalytic zinc metalloproteins [63-65]. The Glu143 residue, which is positioned in the second coordination shell of the zinc cation, forms together with the above-mentioned two histidines the consensus HEXXH motif of the enzyme amino-acid sequence[4]. Although the overall sequence and backbone conformations of the related enzymes display very different binding modes, the zinc-binding HEXXH motif is highly conserved between all the members of clan MA, thus implying evolutionary converged or inherited reaction mechanisms [4, 6].

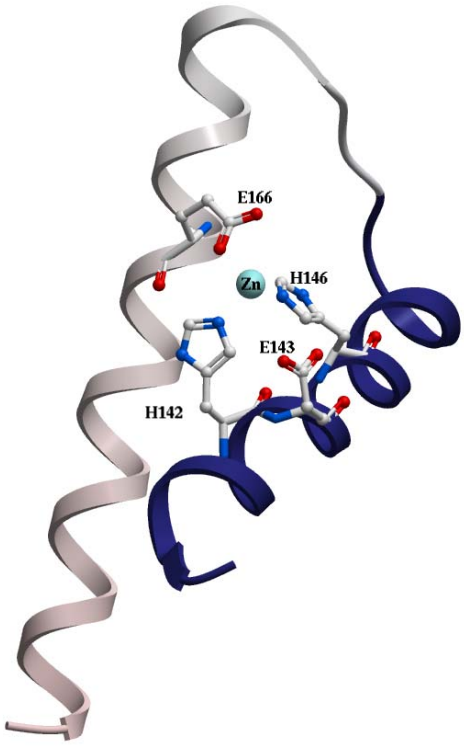


Figure 1.4: The TLN active site structure(PDB:1gxw), generated from X-ray data[17].

.

Jongeneel [66] identified mono-catalytic zinc metallopeptidases from a number of families. The Zn^{2+} is bound by a glutamate (Glu166), 20-33 residues *C*-terminal to the HEXXH motif. The metallopeptidases in which the zinc is bound by HEXXH+Glu are known as 'Glu-zincins' (Figure 1.5). The Zn^{2+} is tetrahedrally coordinated, and the fourth ligand is a water molecule which forms the nucleophile in the catalytic process [67]. Four M4 family members have been crystallized and the 3D structure solved. Their sequence has been compared (Figure 1.5).

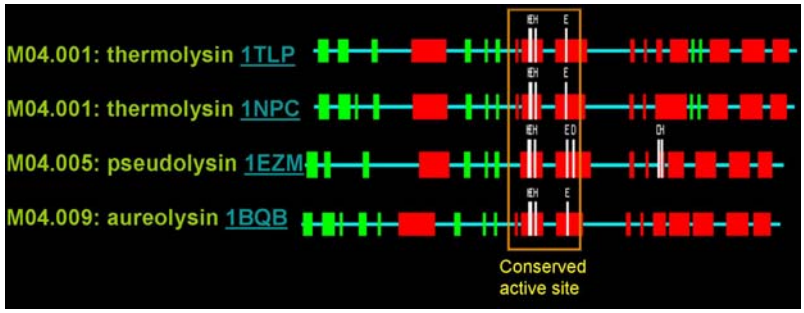


Figure 1.5: Structural alignment of some members of this 'Glu-zincins'. The 3D structures were taken from Brookhaven Protein Databank (PDB, codes 1tlp, 1npc, 1ezm, and 1bqb) where the box is showing the highly conserved active site motif HEXXH (+Glu). Green boxes stand for β sheets and red boxes refer to α helixes.

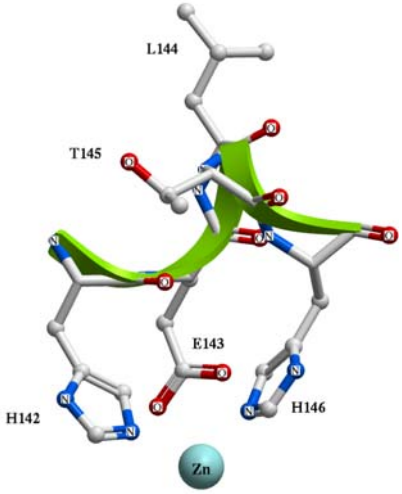


Figure 1.6: The folding pattern of the HEXXH motif in the three-dimensional space of the protein (PDB: 1gxw).

The families of clan MA are united by the presence of an HEXXH motif (Figure 1.6) in which the two His residues are zinc ligands and the Glu has a catalytic function [3]. The fourth ligand is a molecule of water that becomes activated and mediates the nucleophilic attack on the scissile peptide bond (Figure 1.3).

Nevertheless, it is important to recognise that the pentapeptide HEXXH also occurs in many proteins that are not peptidases. Other non-peptidase proteins contain motif HEXXH too, but a more specific motif can be defined for clan MA. A longer consensus sequence that is more reliable in detecting metallopeptidases was described by Jongeneel *et al*. [66] and further refined by Rawlings and Barrett [68].

In the folded metallopeptidases, the HEXXH motif is part of a helix-turn-helix region around the active site, and this turn is required between the two helices to bring the ligands together (Figure 1.4).

1.4.2 The catalytic mechanism

Despite extensive crystallographic studies, site-directed mutagenesis and kinetic investigations accumulated for TLN and homologous enzymes over the years, the catalytic reaction mechanism is still under debate, particularly regarding to which amino acids that take part in catalysis. Usually, Glu143 was originally considered a proton donor, but later considered to be an electrophile [69]. Other residues found to be essential for catalysis are Tyr157, Asp226, and His231 (Figure 1.7).

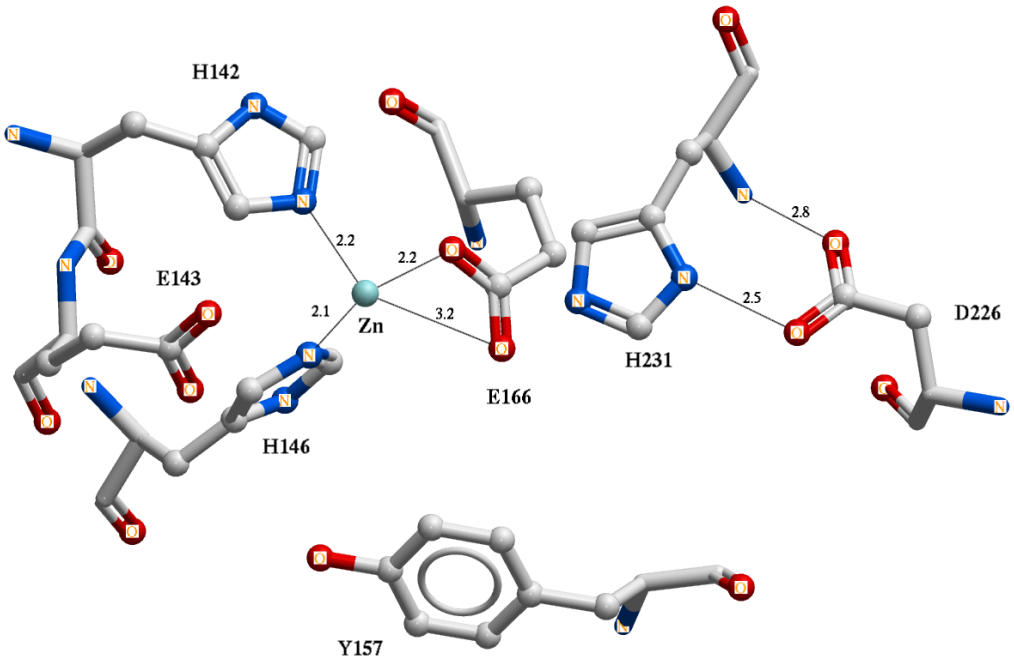


Figure 1.7: The TLN active site structure(PDB:1gxw), generated from X-ray data[17].

The residue His231 is proposed to be the proton donor and general base, while Asp226 is thought to orient the imidazolium ring of His231 by the salt bridge between either OD1 and OD2 of Asp226 and ND1, N of His231 [9, 45, 67] .

Tyr157 is suggested to act as a general acid and thereby contribute to positioning the nucleophilic water molecule. Asp226 is completely conserved amongst M4 or all metallopeptidases [70].

1.5 Experimental studies of enzyme inhibition

1.5.1 Enzyme inhibition

Compounds that influence the rates of enzyme catalyzed reactions either reduce the rate (inhibition) or increase the substrate turnover (activation). Accordingly, the compounds are termed inhibitors or activators. Enzyme inhibition plays a vital role in controlling catalytic reaction. A number of techniques are used for searching the basic explanation of the mechanism of a substrate involved in enzymatic reactions. On the basis of these kinetic observations, inhibitors are usually divided into two main classes, reversible and irreversible, depending upon the manner in which the inhibitor is attached to the enzyme [2]. The reversible inhibitors form noncovalent interactions with the enzyme surface, which can be easily reversed by dilution or dialysis, while the irreversible inhibitors interact with functional groups on the enzyme surface by forming strong covalent bonds that often persist even during complete protein breakdown [71].

The relationship between a single substrate and enzyme concentration was interpreted in 1913 by Leonor Michaelis and Maud Menten with their classic Michaelis-Menten kinetics (Equation 1.1) [72] .

$$v = \frac{V_{\max}[S]}{K_m + [S]} \quad \text{..... (1.1)}$$

V_{max}-maximum rate, K_m-Michaelis-Menten constant, v-reaction rate, [S]-substrate concentration.

The prerequisite of this kinetic model is that the concentration of enzyme is much lower than the concentration of substrate [S] (i.e. where enzyme concentration is the limiting factor), and when the enzyme is not allosteric. From equation (1) we know that the reaction rate v is the number of reactions per second catalyzed per mole of the enzyme. The algebraic rearrangement of the Michaelis-Menten equation is the mathematical basis for deducing Lineweaver-Burk plot [73]. Lineweaver-Burk analysis is one method of linearizing substrate-velocity data so that the kinetic constants K_m and V_{max} can be determined. The reaction rate increases with increasing substrate concentration [S], asymptotically approaching the

maximum rate V_{max} (Figure 1.8). One creates a secondary, reciprocal plot (Figure 1.9): 1/v vs. 1/[S].

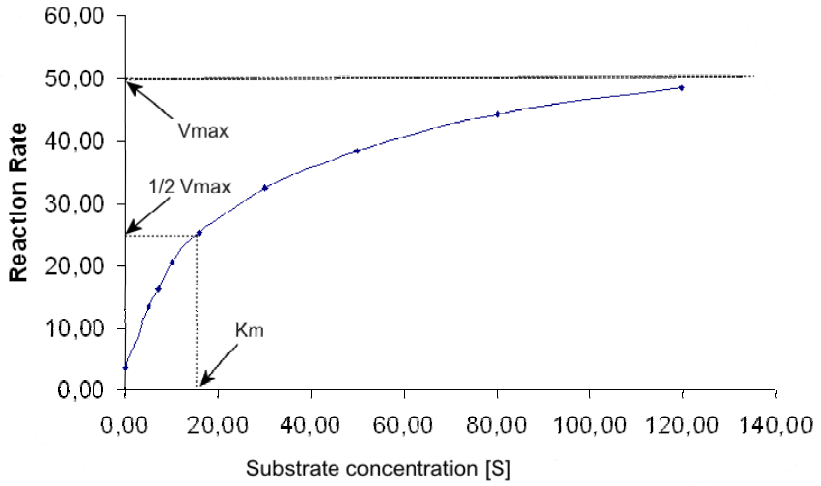


Figure 1.8: Reaction rate V as the function of substrate concentration [S].

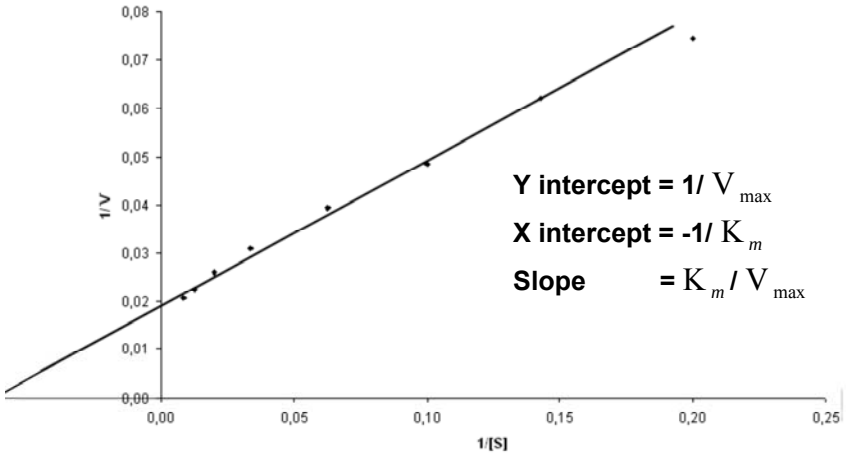


Figure 1.9: Lineweaver-Burk analyses of 1/ V vs. 1/[S].

Although this technique has been widely applied to study various enzyme systems, its utility is somewhat limited because it is most appropriately applied to enzymes that utilize only a single substrate. *In vivo*, a common form of enzyme inhibition involves the competition between substrate and inhibitor for an active site. This is also the case involved in the interfering or removal of TLN disease related activity on its substrate by introducing inhibitors. All the ligands used in this thesis contain common TLN zinc binding groups (ZBG): carboxyl, hydroxyl, sulphydryl, dialkylsilanediols or phosphonate. We assume that they all are competitive inhibitors; so that they compete with the substrate for the active site of the TLN with their ZBG. By forming an inactive enzyme-inhibitor complex, they decrease

the rate of catalysis. Therefore, the Michaelis-Menten equation for the rate (v) of an enzyme catalyzed reaction in the presence of an inhibitor is given by

$$v = \frac{V_{max}}{1 + \frac{K_m}{[S]}\left(1 + \frac{[I]}{K_i}\right)} \quad \text{..... (1.2)}$$

where it shows that, in the presence of inhibitor, the extent to which the reaction is slowed is dependent on the inhibitor concentration [I] and the dissociation constant, K_i, for the enzyme inhibitor complex. It is seen that, smaller K_i value indicates strong binding of the inhibitor to the enzyme. But with competitive inhibitor, the inhibition may be overcome, at a fixed inhibitor concentration, by increasing the substrate concentration.

1.5.2 IC_{50} values and affinity prediction

In the initial screening of inhibitors, it is convenient to compare affinities within tested compounds as percentage inhibition. An IC_{50} value (Concentration of inhibitor required to inhibit enzyme activity by 50% *in vitro*) is a convenient measure of potency. The IC_{50} of a drug can be determined by constructing a dose-response curve and examine the effect of different concentrations of antagonist (inhibitor) on reversing agonist (substrate) activity. This is called a functional inhibitor assay [74]. In the present study a single concentration of the substrate [FA-glycyl-L-leucine amide (FAGLA)] was used in every assay tube for the competition binding assays. The level of specific hydrolysis of the FAGLA was then determined in the presence of a range of concentrations of putative inhibitors, in order to measure the concentration with which they could compete with FAGLA for the binding. Competition curves may also be computer-fitted to a logistic function to reduce the calculating time.

The IC_{50} value is converted to an absolute inhibition constant K_i by using the Cheng-Prusoff equation [75].

$$IC_{50} = K_i\left(1 + \frac{[S]}{K_m}\right) \quad \text{..... (1.3)}$$

Where K_i is the inhibition constant for a drug, and indicates the binding affinity of the inhibitor. [S] is substrate concentration and K_m is the affinity of the substrate for the enzyme. Whereas the IC_{50} value for a compound should be used with care when comparing inter laboratory results for competitive inhibition, because it is dependent on the substrate

concentration and the assay you are using. For non-competitive and uncompetitive inhibitors, IC_{50} equals K_i value and is independent of substrate concentration. K_i is independent of substrate and inhibitor concentrations for all classes of reversible inhibitors [76].

1.6 Theoretical studies of protein ligand-interactions

1.6.1 Docking

Structure based drug design approaches have increasingly demonstrated their value since the first biologically relevant X-ray structures became available 30 years ago. The impact of these methods and technologies on early lead discovery and lead optimization is significant. Protein-ligand docking aims to predict and rank the structure(s) arising from the association between a given ligand and a target protein of known 3D structure [77]. Pioneered during the early 1980's [78], it remains a field of vigorous research, having become a useful tool in drug discovery programs [79]. In particular, protein-ligand docking occupies a very special place in the general field of docking, because of its applications in medicine [80]. From the docking of both protein and ligand as rigid bodies [78], protein-ligand docking has developed to an area where full flexibility on the ligand is commonly evolved. Even though protein flexibility is known to increase affinity between a given drug and its target [81], and the fact that proteins are in constant motion between different low energy conformational states. Protein flexibility has not been well considered by new docking programs (Table 1.3). Furthermore, binding orientation and binding-site location can be greatly influenced by protein flexibility. The traditional lock-and-key and induced-fit theories have given their way to more modern theories that bestow a greater weight to the receptor flexibility issue [82]. The current idea interprets a protein as an ensemble of differently populated conformational states in equilibrium, rather than the stable conformation appearing in X-ray crystal structure [82]. An important notion that needs attention is that the conformation of receptor in the protein-ligand complex are not necessarily the most populated conformations in an unbound state [81-83]. So, when it comes to docking, these aspects imply that instead of targeting a single pose of a given ligand on a single receptor structure, one should ideally look for the most populated alternatives from an ensemble of solutions comprising several different binding conformations [77].

Table 1.3: Most common used ligand-protein docking programmes, official website, docking algorithms, scoring function and system compatibility.

Program	Docking Algorithms	Scoring function	Scoring Function Class	System compatibility	Website
AUTODOCK [84, 85]	MC & GA	Autodock SF [84]	Force Field or Empirical	Linux, Windows, and MacIntosh OSX systems.	www.autodock.scripps.edu/
GOLD [86, 87]	GA	ChemScore and GoldScore [88, 89]	Empirical / Force Field	Linux and Windows system	www.ccdc.cam.ac.uk/products/life_sciences/gold/
FLexX [90]	ICA	Modified empirical Böhm's scoring function [90]	Empirical	Linux, SGI Irix and Windows system	www.biosolveit.de/
DOCK [91-94]	ICA	Dockscore [95]	Force Field	Linux, Windows, and MacIntosh OSX systems.	www.dock.compbio.ucsf.edu/
ICM [96, 97]	MC	ICM scoring function [98]	Force Field	Windows Vista/XP/NT/2000 ,Linux/i386/AMD 64, SGI IRIX, Mac OS X	www.molsoft.com/docking.html
GLIDE [99, 100]	HSS & MC	GLIDE_Emodel scoring function [99, 101]	Empirical	Unix, Linux, Windows, SGI IRIX/Altix, IBM AIX, Intel hardware	http://www.schrodinger.com/ProductDescription.php?mID=6&sID=24&cID=0
Fred [102]	RBD	ChemScore, [88]	Empirical	LINUX, IRIX, Windows, OS X, AIX, HP-UX, Solaris, Tru64UNIX	http://www.eyesopen.com/products/applications/fred.html
CHARMM [103]	GA & MC	CHARMM score [103]	Force Field	Unix and Linux	www.charmm.org/
eHits [104]	RBD of fragments followed by reconstruction	eHits scoring function [105]	Empirical	Linux, SUN Solarix, SGI-IRIX and IBM-AIX	http://www.simbiosys.ca/ehits/
PatchDOCK [106]	SP	geometric shape complementarity score [107]	Knowledge based	Web server or Linux	http://bioinfo3d.cs.tau.ac.il/PatchDock/
Surflex [108]	ICA & MA	Hammerhead scoring function [109]	Empirical	Windows	http://jainlab.ucsf.edu

Note: MC, Monte Carlo method; GA, genetic algorithm; ICA, Incremental construction algorithms; HSS, hierarchical systematic search method; RBD, rigid body-docking; SP, shape complementarities; MA, matching algorithm.

There are currently more than 49,000 (by Mar 11, 2008) crystallographic or NMR structures of proteins or nuclei acids available from the Protein Data Bank (PDB) [110], and the rate of 3D macromolecular structure determination continues to increase every year, particularly with the development of new techniques such as high-throughput X-ray crystallography [111]. Many of these macromolecules play vital roles in critical metabolic pathways and may be regarded as potential therapeutic targets, offering unparalleled opportunities for structure-based drug design and discovery. In this context, the ideal of using a high resolution structure of a target protein to design the perfect ligand binding modes is challenging. Completely de novo design of a lead molecule produced clinical candidates as zanamivir, commercial name RELENZA® , by GlaxoSmithKline company against common cold [112], and AG-331 as an anti-cancer drug [113].

In terms of protein-ligand docking methods, the goals can be rationalized as first, search for precise ligand conformations and poses by minimizing the total energy of the protein ligand complex, within a given targeted protein when the structure of the protein is known or can be estimated. Second, to predict the binding affinity of any protein ligand complex.

So, docking, as a result, usually involves two independent steps:

- Determining the orientation of a ligand relative to the receptor.
- Evaluate (score) the ligand orientation.

1.6.2 Scoring

The binding affinity prediction problem denotes the question of how well the ligands bind to the protein (scoring) [77]. In a biological environment, the system would include not less than the solvent molecules, the ligand, and the macromolecular receptor. Solvent molecules are commonly excluded from docking approaches, because of the tremendous number of degrees of freedom with the solvent molecules, or in special cases implicitly modeled in the scoring functions as a way to address the solvent effect. However, even the remaining part of the system - ligand and receptor - has a computational untreatable number of degrees of freedom, and therefore, the dimensionality of the problem has to be reduced through the application of different approximations, allowing the search space to be more effectively sampled.

In the last 15 years, a number of scoring functions (SFs) have been reported and implemented in docking programmes. The SFs of most often used docking programs are

listed in Table 1.3. In general, SFs attempt to predict the binding free energy or to rank-order compounds by their bioactivity. These SFs are classified as force field (FF)-based, empirical and knowledge-based [114-117]. Since the accuracy of SFs plays key role in a successful docking approach [118], numerous progress over the years has been made. However, commonly used SFs are known to have limitations as the protein-ligand complex often includes subtleties not captured by SFs [119]. The SFs should enable the distinction between the true binding modes and all the other alternative modes explored, or between active and random compounds. However, a very rigorous scoring function would be computationally too expensive, rendering the analysis of the several binding modes unfeasible. While, many oversimplifications, in the name of speeding up screening time, are believed to be one of the main causes of poor ranking. This in turn, causes major difference of docking successfulness of different docking programmes (Figure 10). For this reason, the lack of a suitable scoring function, both in terms of speed and accuracy, is the major bottleneck in docking [79]. To overcome this hurdle, it is ideal to combine several more scoring methods in docking programmes. For example, ChemScore and GoldScore have been implemented in GOLD. This has been found to be superior to the use of a single function in some cases [120].

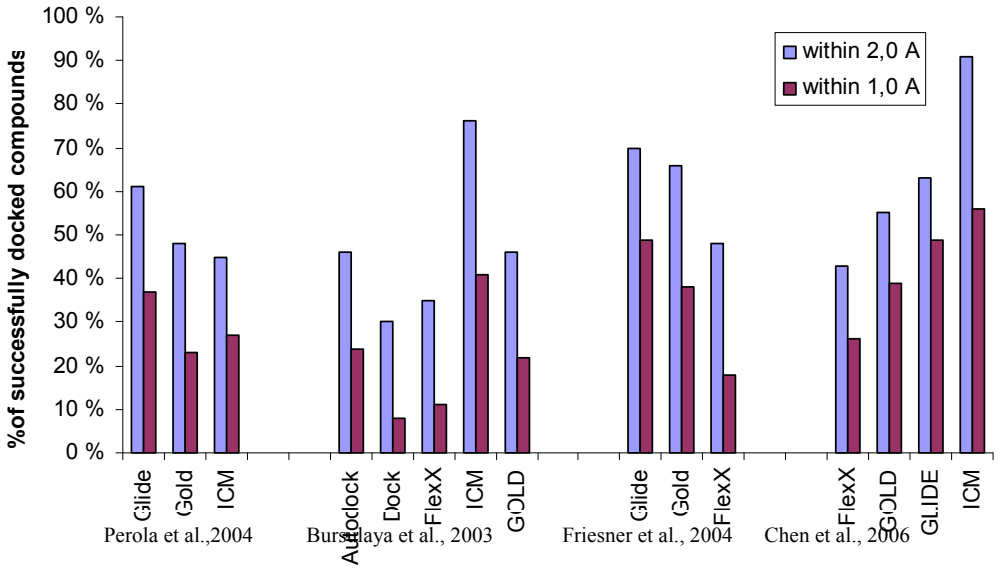


Figure 10: Comparative studies of different docking program, reporting the present of docked compounds within an RMSD of 1 and 2 Å from corresponding X-ray structure.

1.6.3 Different programs

The Internal Coordinates Mechanics ICM™ program (www.molsoft.com) [121] performs flexible ligand docking using a Monte Carlo minimization procedure in internal coordinates to find the global minimum of the energy function. It is also fit for protein-protein docking. Some frequently used protein-protein docking programs are listed in Table 1.4. The ICM program was compared with AutoDock, DOCK, FlexX, and GOLD in several studies [122, 123]. ICM program provided the highest accuracy and has also outperformed DOCK and FlexX in virtual library screening tests [122].

Molecular docking methodologies ultimately seek to predict (or often retrospectively reproduce) the best mode by which a given compound fit into a binding site of a macromolecular target. On the basis of the results for this data set, ICM therefore appears to be the most versatile VLS and Docking tool. Comparing low throughput approaches (those based on molecular dynamics), until now, more than 60 docking programmes and over 30 scoring functions have been developed [124].

Table 1.4: Commonly used protein-protein docking programs.

Software	Algorithm
ICM [121]	Force field
CHARMM [103]	Force field
MolFit [125]	Fast Fourier Transformation
FTDOCK [126]	Fast Fourier Transformation
DOT [127]	Fast Fourier Transformation
Hex [128]	Shortest path first
ZDOCK [129]	Fast Fourier Transformation
BUDDA/PPD [107]	Geometric Hashing
Guided Docking [130]	Force field
BiGGER [131]	Global scoring function
GAPDOCK [132]	Genetic Algorithm Approach
Surfdock [133]	Fourier correlation of spherical harmonics
GRAMM [134]	Fast Fourier Transformation
PatchDOCK [106]	Fast geometric shape complementary

2 Aims of the study

It is clear that for any modeling strategy to be validated, it must at the very least rationalize (either qualitatively or semi-quantitatively) the existing structure-active relationships (SARs). Furthermore, it must be capable of making predictions (either extending SARs or leading to the birth of new SARs) that can be tested through feasible experimental strategies. An essential component of the development of such a model is the availability of reliable biological assays which preferably yield consistently reproducible results and which are based on a finite number of mechanisms. Thus, inhibition of binding by a group of ligands to an enzyme would constitute an excellent example of a reliable assay which could form the basis of the pharmacophore design and protein inhibitors library building. We gathered a certain amount of test compounds from several of our collaborators. The main goals of this task fall in:

- Determine the molecular interactions of inhibitors from the literature with TLN by using docking and scoring.
- Discover novel TLN inhibitors by using experimental and theoretical approaches.
- Determine the molecular interaction of these inhibitors with TLN.

As computational potency and knowledge are constantly increasing, the integration of ICM™ program with experimental assays would make it compatible with current trend in drug discovery and add another dimension in the evaluation of novel targets.

3 Methods

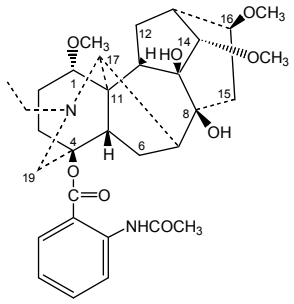
3.1 In vitro experiment

3.1.1 Chemicals and compounds

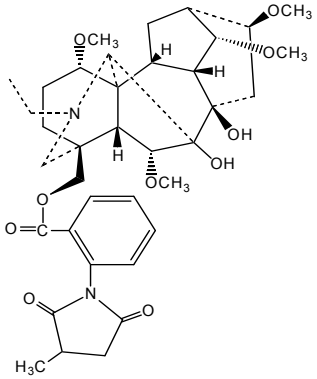
All the chemicals used for the binding assays were purchased from Sigma-Aldrich (US); three-times crystallized thermolysin was purchased from CalBioChem (E-Merck, Germany). The tested compounds were from different collaborators on the basis of joint collaborative research projects. MS compounds are from Prof. Mukhlis Sultankhudzaev, Uzbek Academy of Science, Tashkent, Uzbekistan; R compounds are from Dr. Rasool Khan, Department of Chemistry, University of Peshawar, Pakistan; M compounds are from Dr. Lenta Ndjakou Bruno, Department of Organic Chemistry, University of Bielefeld, Germany.

Table 3.1: List and codes of compounds tested in this study.

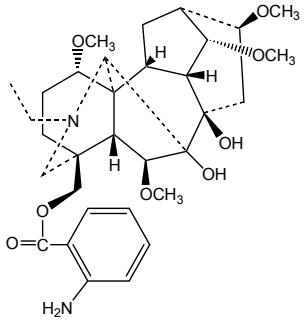
Groups	Compounds			
MS	MS1, MS2, MS3, MS4, MS6, MS8, MS9, MS10, MS11, MS12, MS13, MS14, MS15.			
R	RK2-16	RSH 57	RSH1	RSH39(d)
	RSH11	RS 7	RSH 42	RS1
	RS10	RSH 66	RK1-4	J1
	RSH22(b)	RSH30d2	RSH77(c)	RM19
	RS17	RK2-10	RSH28(b)	RS3
	RSH79(a)	RS20(R)	RSH 16	RSH12
	RSH78(f)	RSH19(b)	RS 14	RSH44a
	J 11	RSH35a	RSH 41	
	RSH 10	RSH30d	RSH 23	
	RS 12	RS 25	RSH78(g)	
M	MEA2	SZA1		
	MEA3	SZA4		
	MTA8	MTA9		
	MTA8AC			



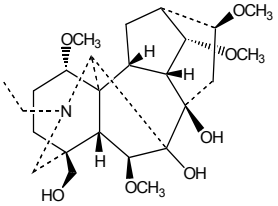
MS1



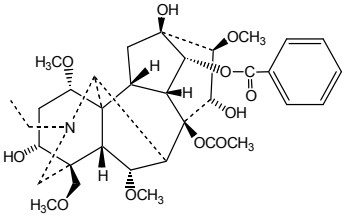
MS2



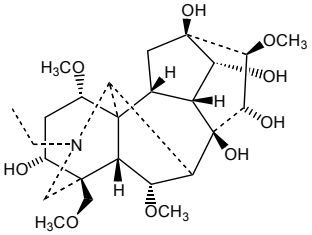
MS3



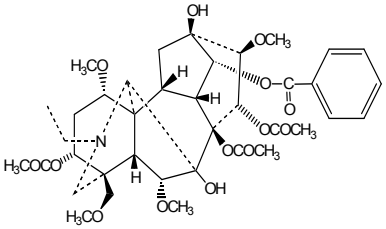
MS4



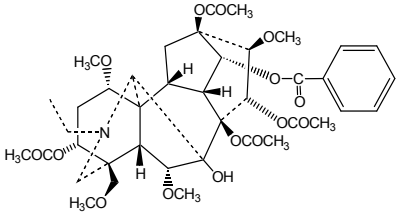
MS6



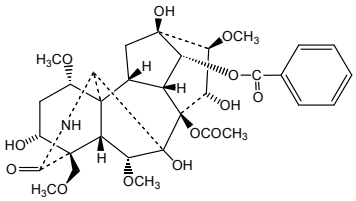
MS8



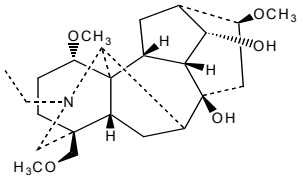
MS9



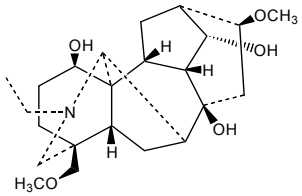
MS10



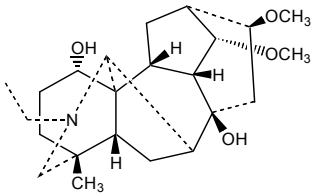
MS11



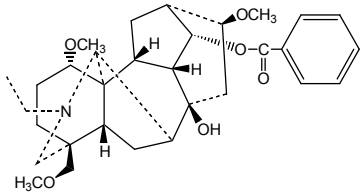
MS12



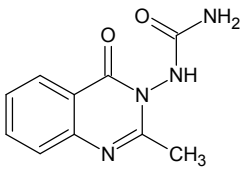
MS13



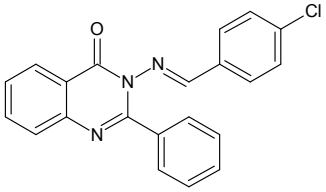
MS14



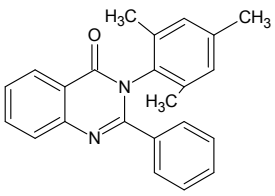
MS15



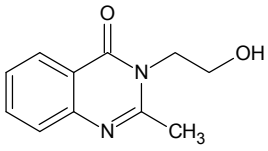
RK2-16



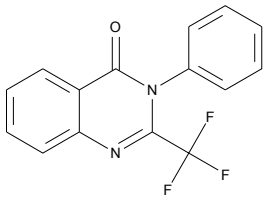
RSH11



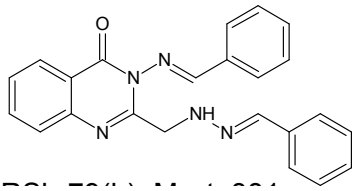
RS10



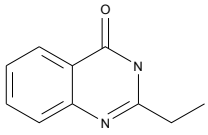
RSH22(b)



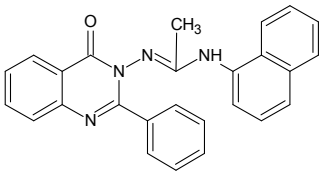
RS17



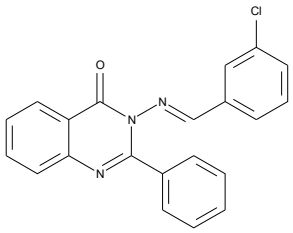
RSH79(b)



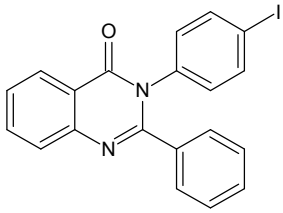
RSH78f



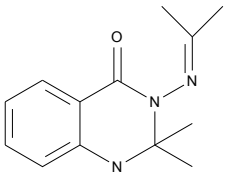
J11



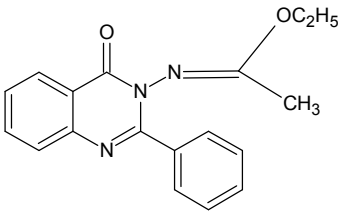
RSH10



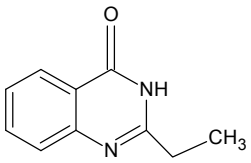
RS12



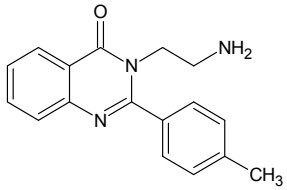
RSH57



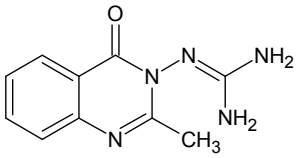
RS7



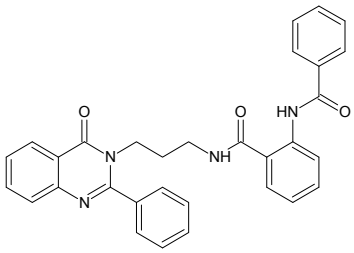
RSH66



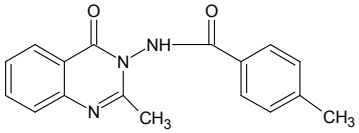
RSH30d-2



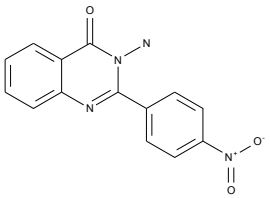
RK2-10



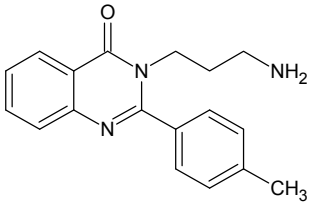
RS20(R)



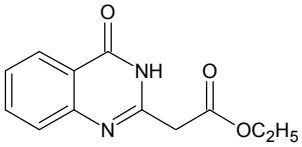
RSH19(b)



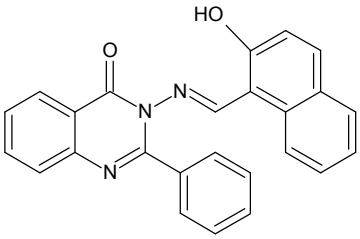
RSH35a



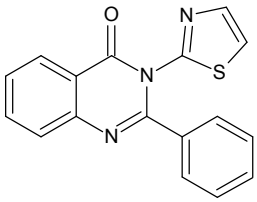
RSH30(d)



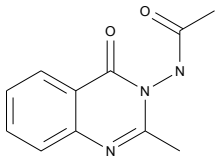
RS25



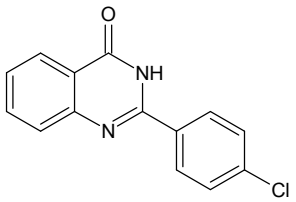
RSH1



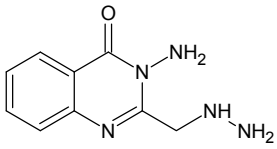
RSH42



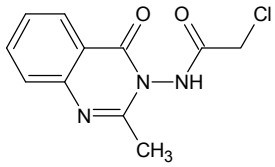
RK1-4



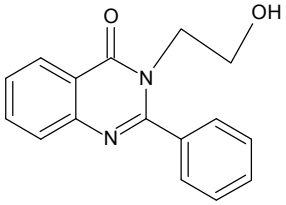
RSH77(c)



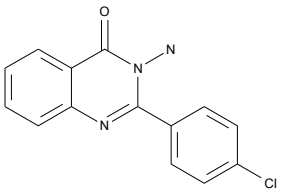
RSH28(b)



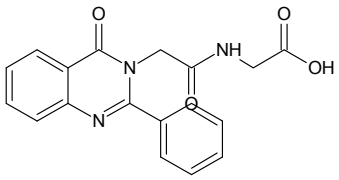
RSH16



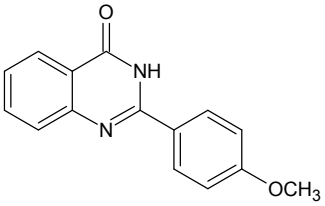
RS14



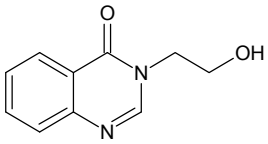
RSH41



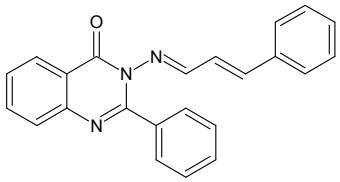
RSH23



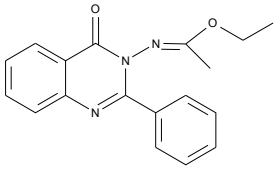
RSH78(g)



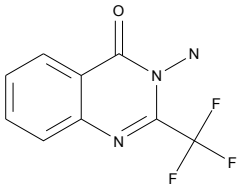
RSH39(d)



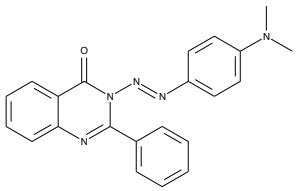
RS1



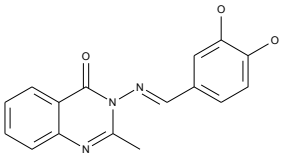
J1



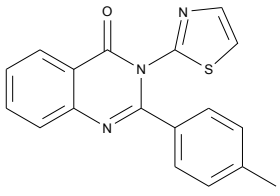
RM19



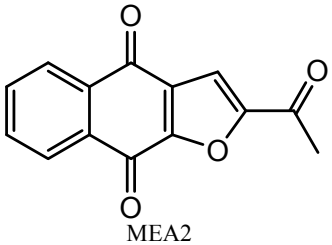
RS3



RSH12



RSH44a



MEA2

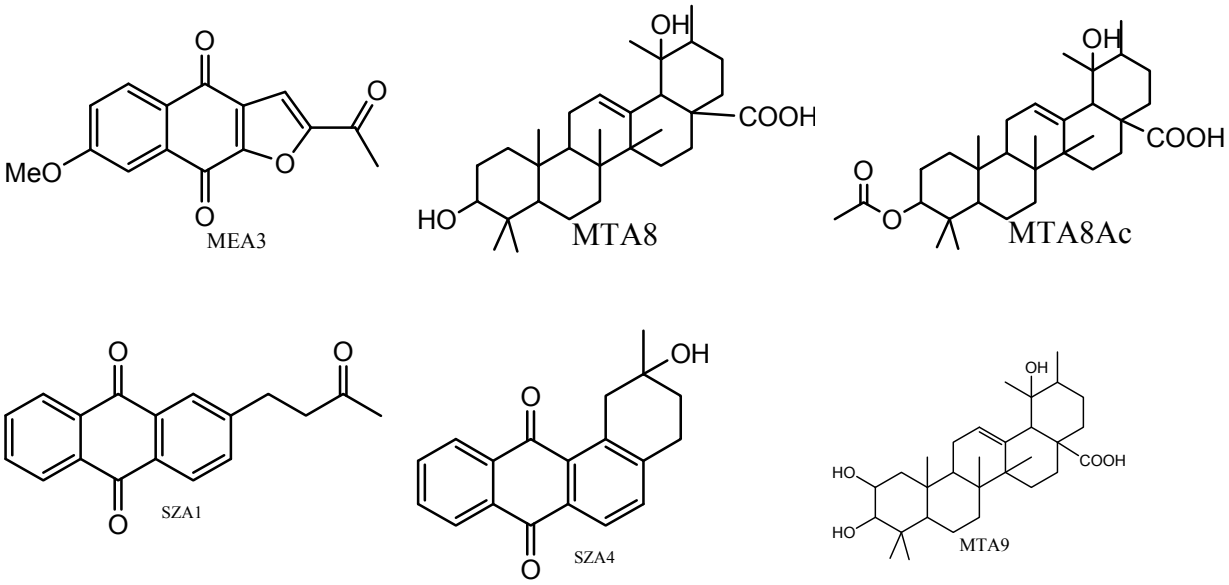


Figure 3.1: Structures of all experimental compounds.

3.1.2 Assay method

All the steady-state enzyme assays were performed at 25 °C using the modified spectrophotometric method of Feder and Schuck [135]. The method was modified such that a 96-well microplate instead of single curette could be used, thereby increasing the through-put of the experiments. The principle of the assay is a "mix and measure". Thermolysin is mixed with substrate FAGLA that is digested by thermolysin. The changes in absorbance due to the compound concentration by a spetctrophotometer which is proportional to the activity of thermolysin is detected. Thermolysin inhibition is measured by adding a fixed amount of thermolysin and a putative inhibitor to the well. Inhibitor blocks the activity of thermolysin. Therefore less substrate will be degraded, which reduces the signal detected by the spectrophotometer. The percentage of inhibition is calculated by the equation (3.1).

For all assays, the concentrations of enzyme and substrate were 50 nM and 1.0 mM, respectively (the concentration in M). The thermolysin activity was determined by following the decrease in absorption at 346 nm due to the enzymatic hydrolysis of the substrate FAGLA. Three times crystallized thermolysin was obtained from a commercial supplier (Novabiochem, Germany) and used without purification. Stock solutions of Tris (50 mM), NaBr (2.5 M), and $CaCl_2$ (10 mM), pH 7.0 were prepared and stored at 4 °C. A stock solution of FAGLA was prepared in DMF (dimethyl formamide), and diluted with buffer to a final concentration of 0.1 M Tris, 0.1 M NaBr, and 2.5 mM $CaCl_2$, pH 7.0 (final concentration of DMF, 2.5%). The enzyme and compounds were incubated for 15 min at 25 °C in a temperature-regulated 96-

well microplate, and then the FAGLA solution was added to give a substrate concentration of 1.0 mM. Initial velocities for <10% reaction were determined [136, 137].

These were in duplicate at three different concentrations (0.5, 0.05, and 0.005 mM) of all experimental compounds to get a view of the inhibition pattern. For compounds found to inhibit thermolysin activity was calculated by equation (3.1).

$$\%\text{Inhibition}=\left(\left(\frac{S}{B}\right)-1\right)\times 100 \quad \text{(3.1)}$$

Where *S* and *B* stand for absorbance of experimental samples and blank.

3.2 Docking calculation using ICM

Three groups of ligand (25 ligands from literatures, 15 from MS compounds and 13 from RS compounds) were docked to target protein. The regular docking methodology was used during the first time docking of 25 ligands from literature and initial docking of all 15 ligands of MS group (Table 3.1). A modified method was applied in the docking of 13 ligands from R group, second time docking of 25 ligands form literature and docking of 2 MS compounds.

3.2.1 Preparation of ligands before docking

All the docking calculations of thermolysin inhibitors were performed using the ICM™ [97] docking module. First, the 2D structures of the literature retrieved inhibitors were drawn using the CS-ChemOffice (www.chembridgesoft.com) and converted to 3D using Discovery Studio (www.accelrys.com) and saved in PDB format.

The X-ray coordinates of inhibitors that have been crystallized with thermolysin were extracted directly from the PDB file and energetically minimized using ICM. Then ligands were examined for bond order and protonation state.

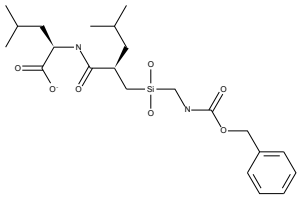
All carboxylic acids were deprotonated, tertiary amines were positively charged, phosphonates were partially deprotonated. Merck molecular force field (MMFF) [138] partial charges were assigned to the ligands.

In the modified docking method, the structures were written out as mol table files from CS-Chemoffice ready for docking. Without manually preparations indispensable in the case of PDB structures, ligand structures from mol file can be automatically treated with all necessary optimizing processes on the fly after running docking simulations.

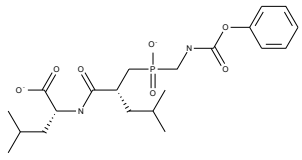
Based on structure similarity to ligands in X-ray complexes with thermolysin, the 25 ligands were rearranged into 7 sub-groups (Table 3.2). The structure of th 25 ligands are shown in Figure 3.2.

Table 3.2: Seven groups of 25 ligands from the literature.

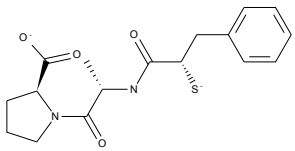
Group No.	Ligand number	X-ray structure within the group
Group 1	6, 7, 8 ,10,11	1qf1
Group 2	3, 4, 9	1qf2
Group 3	5	1qf0
Group 4	12,13	1os0
Group 5	1,18	1y3g
Group 6	2, 14, 15, 16, 17, 19, 20, 21	5tmn,6tmn
Group 7	22, 23, 24, 25	1thl



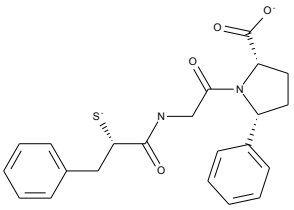
1



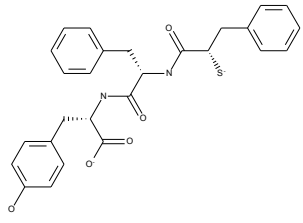
2



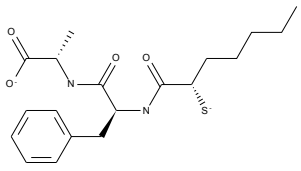
3



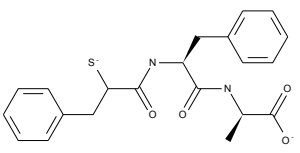
4(1qf2)



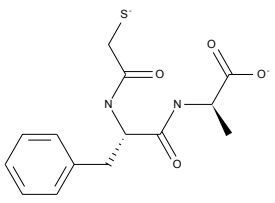
5(1qf0)



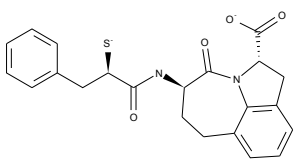
6(1qf1)



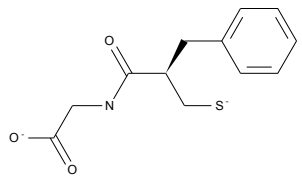
7



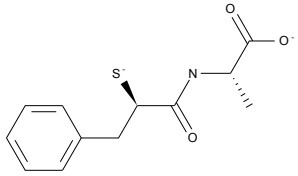
8



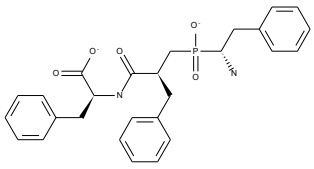
9



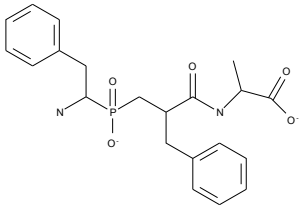
10



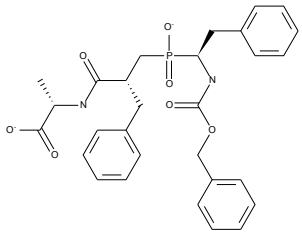
11



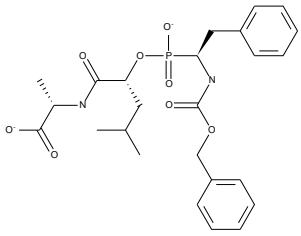
12(1os0)



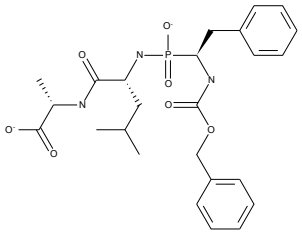
13



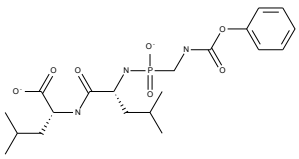
14



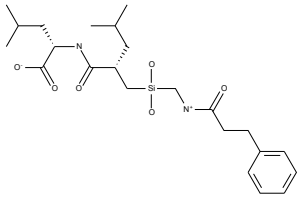
15



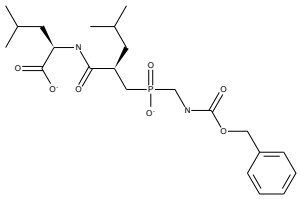
16



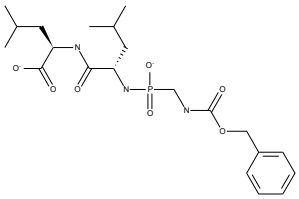
17



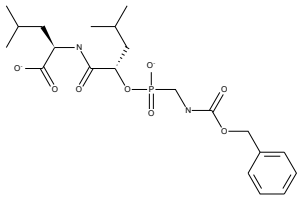
18



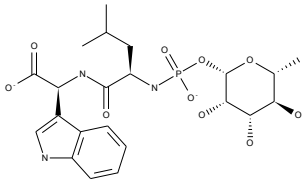
19



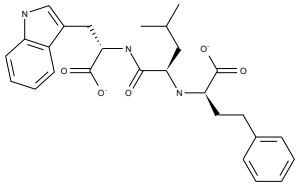
20(5tmn)



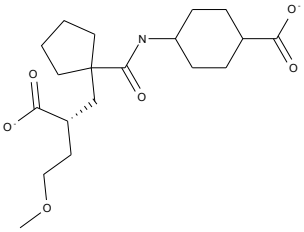
21(6tmn)



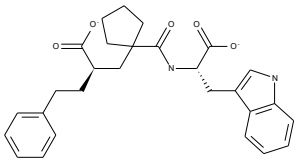
22



23



24



25(1thl)

Figure 3.2: Structure lists of 25 thermolysin ligands from the literature. The PDB code is given for those obtained from X-ray complexes with thermolysin.

3.2.2 Preparation of the target protein

In the meantime thermolysin (Target PDB code 1gxw) was prepared for the docking calculations with ICM. The retrieved 1gxw file from Brookhaven Protein databank (PDB, www.rcsb.org/pdb) was converted to ICM-object. 2266 missing hydrogens were added, and 79 polar hydrogens were optimized. Metal ions were preserved, while crystallographic waters were deleted. This was followed by a local minimization to relieve potential bad contacts. During the first docking of the 25 ligands known from the literature and initial docking of MS compounds, the 3D space for docking was specified to include the entire binding pocket of the target (Figure 3.3). Zn was used as a core. All amino acids within 5 Å from the core were included in the 3D space of the pocket. A modified docking method was used for the other dockings. In this approach the ICM pocket finder module was used to determine the binding pocket, such that enough residues were included for obtaining correct receptor-ligand interactions. The measured volume of binding pocket for this method was 285.9 cubic angstroms, and radius 4.087 angstroms. Amino acids surrounding binding pocket were Asn112, Trp115, Gly117, Phe130, Leu133, Val139, His142, Glu143, His146, Asp150, Tyr157, Glu166, Ile188, Leu202, Arg203 and His231. The grid map was adjusted manually for all docking, and five potential maps (electrostatic, hydrogen bond, hydrophobic, van der Waals attractive and repulsive) were calculated for the receptor.

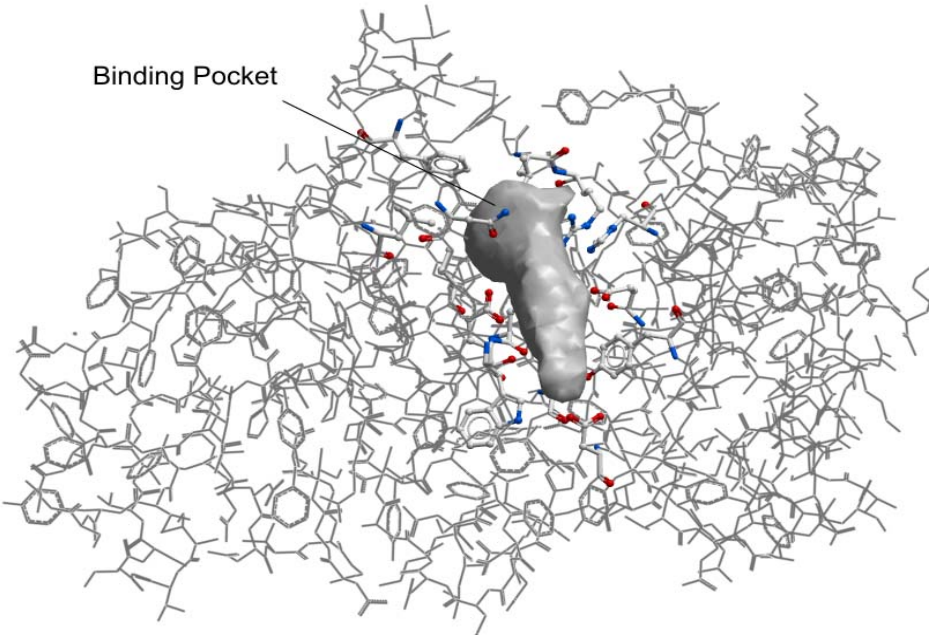


Figure 3.3: The binding pocket of thermolysin (PDB: 1gxw).

3.2.3 Docking process

Regular docking method is usually docking of one ligand at a time. To speed up the docking process, the modified method was using the 'batch docking' approach. Batch docking utilizes the ligands that are assembled as mol tables in one 'sdf' file. Instead of docking each ligand separately, the docking of multiple ligands can be done in one run and the results are displayed in an output file. It is ideal for large-scale docking jobs. Monte Carlo (MC) [139] docking runs were then performed. Based on our experiences from the regular docking of 25 ligands from literature and 15 MS compounds, we performed 5 runs of docking for the rest of the ligands using the modified batch docking method. All dockings by the modified method were using the same grid maps. ICM stochastic global optimization algorithm attempts to find the global minimum of the energy function which includes five grid potentials describing interaction of the ligand. A torsional or positional conformation change, followed by local minimization, was performed at each step. During this process a stack of alternative low energy conformations was saved. The docked solutions (stacks) were checked visually. Multiple stack conformations were chosen based on their docking energies, and pose similarities to crystal structures.

The ligands those have been through 5 runs of docking were analyzed differently. First, the best poses of each of 5 runs were saved based on docking energy, and similarity to x-ray structures, and then their docking energies were averaged. The one with a docking energy closest to the average value was chosen as reference which was sequentially used to calculate RMSD to remaining best poses (one from each run). Then a selection was done between the reference conformation and the one closest in RMSD, and the best conformation for each compound was finally used for calculation of their binding energies using the ICM script (briefly described in the following 'Calculation' section).

Each docking process gave multiple conformations depending on the number of flexible side chains. The final conformation was chosen based on several points, like docking energy, hydrogen bonding, interactions with Zn, and interaction with the HEXXH motif.

3.2.4 Calculation of the free energy of binding

In the regular docking method, the free energies of binding (Cal.ΔG) between the protein and ligand was calculated for all possible conformations in each stack file. It was used as one of the guidelines for determining the best docking poses. But, in the modified docking

method, Cal.ΔG was calculated only for the best poses and compared with the experimental free energies of binding (Exp.ΔG) which was calculated from the experimental *Ki* values using equation 3.2.

$$\Delta G = -RT\ln K_i \quad (3.2)$$

Where R is the gas constant (0.00198 Kcal/mol) and T is the temperature in Kelvin (298 K).

The calcBindingEnergy macro of ICM was used to calculate the free energy of binding. This approach evaluates the binding of each ligand to target protein by estimating the electrostatic, hydrophobic and entropic energy terms. The evaluation of binding energy has the following features: van der Waals/hydrogen bonding interaction is excluded since it has close magnitudes for protein-protein and for protein-solvent interactions; electrostatic free energy change is calculated by the REBEL method; side-chain entropy change is calculated by standard ICM entropic term based on exposed surface area of flexible side-chains; hydrophobic energy change is calculated using surface term. Generally, the free energies of binding (Cal.ΔG) between the protein and ligand was calculated using ICM script utilizing equations 3.3 and 3.4 [140].

$$\Delta G = \Delta G_H + \Delta G_{EL} + \Delta G_S + C \quad (3.3)$$

$$\Delta G = \Delta G_H + \Delta G_{COUL} + \Delta G_{DESOLV} + \Delta G_S + C \quad (3.4)$$

where ΔG_H is the hydrophobic or cavity term, which accounts for the variation of water/non-water interface area, ΔG_{EL} is the electrostatic term composed of coulombic (ΔG_{COUL}) interactions and desolvation (ΔG_{DESOLV}) of partial charges transferred from an aqueous medium to a protein core environment, and ΔG_S is the entropic term which results from the decrease in the conformational freedom of functional groups buried upon complexation; and finally *C* is a constant that accounts for the change of entropy of the system due to the decrease of free molecules concentration (cratic factor), and loss of rotational/translational degrees of freedom [140]. The energy unit is kcal/mol.

3.2.5 Plotting the interactions using LigPlot

For easy interpretations of the interactions between the receptors and the compounds LigPlot [141] was used to generate 2D plots of the ligand-thermolysin complexes.

4 Result

4.1 Molecular modeling of 25 thermolysin ligands from the literature

To assess the prediction of binding modes and affinity by ICM and to characterize the docking accuracy, several parameters were calculated using both regular (default) and modified docking methods. The first parameter was docking energy which stands for the total of all the terms, both energy (i.e. " van der Waals, 1-4 van der Waals, hydrogen bonding energy, electrostatic energy, torsion energy, hydrogen bonding energy, bond stretching energy, surface term ") and penalty part (i.e. " distance restraints, tethers, multidimensional variable restraints ") [97]. The second parameter was the calculated binding energy (Cal.ΔG) between the docked ligand and thermolysin. This value was calculated by using equation 3.3 and 3.4. The third parameter, experimental free energies of binding (Exp.ΔG) was calculated from the experimental *Ki* values using the equation 3.2. The last calculated parameter was static Root Mean Square Deviation (sRMSD) from X-ray crystallographic structures. In the present, sRMSD gives the RMSD between the binding modes of two selected chemical (hetero) molecules according to the optimal chemical match by superimposing target protein coordinates, sRMSD was calculated between ligand structure conformation of the best ranked docking pose and the corresponding crystallographic protein coordinates. In our analysis, to maintain simplicity and evaluate the usefulness and true predictive capability of all sets of ligands, we chose to focus on both docking energy, sRMSD and the position of the putative ZBG as the basis for selecting between top ranked poses. We were very much interested in whether we could repeatedly rely on these criteria as an indication of the binding mode prediction. We determined the negative log of the experimental K_i as pK_i [142] value of ligands to compare with calculated ΔG of the docking poses. The flexibility of the docked ligand was defined by the number of rotatable bonds. The experimental 3D coordinates were available only for those ligands co-crystallized with thermolysin. The Cambridge structural database (CSD) [143] was searched for ligand structures of the 25 ligands in this part of the study, but structures were found. To avoid that parameters and optimizing algorithm could

induce deviation due to different chemical structure drawing program of small molecules, only 2D mol files were used when using the modified docking method. All the optimizing and converting of 2D to 3D approaches were done by the ICM environment. The structure of 25 ligands are shown in Figure 3.2.

4.1.1 Initial docking of 25 ligands

The ligands were divided into 7 sub groups according to their structural similarity with ligand co-crystallized with thermolysin. All possible stack conformations for each ligand were screened by calculating ICM binding energy (Cal.ΔG). Based on the similarity to crystal structure complexes with thermolysin and to the similarity of the calculated binding energy, the best poses were chosen (2 to 3 for each ligand). These poses were then subjected to further refinements and similar evaluating approachs were performed for the refined poses, and final poses were selected. Table 4.1 demonstrates the changing of binding energy before and after refinement of the best poses. The correlation between the Cal. ΔG and Exp. ΔG was calculated by regression analysis indicated no correlation at all (R^2 = 0.1874). The reason for that may be the presence of diverse chemical moieties in their skeletons.

Table 4.1: Results of the first docking of 25 ligands (ligands were not charged). The energy values are given in kcal/mol.

X-ray Structure	Comp.	Docking energy before refinement	Cal. ΔG before refinement	Docking energy after refinement	Cal. ΔG after refinement	K_i (in M)	pK_i	Exp. ΔG
	1	-131.7	-10.90	-1095	-9.38	4.1×10^{-8}	7.4	-10.0
	2	-82.0	-7.79	-1068	-11.29	1.0×10^{-8}	8.0	-10.9
	3	-79.1	-7.44	-986.5	-11.57	2.0×10^{-7}	6.7	-9.1
1QF2	4	-91.6	-7.53	-1042	-7.45	1.2×10^{-6}	5.9	-8.0
1QF0	5	-102.4	-6.08	-1067	-8.84	4.2×10^{-8}	7.4	-10.0
1QF1	6	-76.7	-10.66	-1058	-7.83	4.8×10^{-8}	7.3	-9.9
	7	-86.1	-10.58	-1072	-6.19	1.9×10^{-8}	7.7	-10.5
	8	-76.1	-7.97	-1066	-4.70	1.0×10^{-6}	6.0	-8.2
	9	-87.0	-8.05	-1052	-9.46	1.6×10^{-6}	5.8	-7.9
	10	-79.6	-5.60	-1031	-5.08	1.6×10^{-6}	5.8	-7.9
	11	-64.2	-3.87	-1046	-5.62	3.5×10^{-7}	6.5	-8.8
1OS0	12	-112.2	-8.88	-1118	-13.74	9.3×10^{-7}	6.0	-8.2
	13	-105.8	-7.55	-1116	-8.43	7.8×10^{-7}	6.1	-8.3
	14	-136.8	-9.38	-1106	-13.05	2.6×10^{-8}	7.6	-10.3
	15	-122.9	-7.57	-1073	-8.83	4.5×10^{-8}	7.4	-10.0
	16	-127.3	-8.78	-1015	-9.59	6.8×10^{-11}	10.2	-13.8
	17	-106.9	-9.78	-1069	-7.70	9.1×10^{-9}	8.0	-10.8
1Y3G	18	-118.3	-10.31	-1082	-9.65	4.1×10^{-8}	7.4	-10.1
	19	-109.4	-8.27	-1029	-6.72	1.0×10^{-8}	8.0	-10.9
5TMN	20	-120.3	-6.31	-1090	-12.02	2.0×10^{-7}	8.0	-10.9
6TMN	21	-123.9	-2.52	-1072	-3.94	1.2×10^{-6}	2.1	-2.8
	22	-134.8	-10.99	-1119	-11.72	4.2×10^{-8}	7.1	-9.6
	23	-99.7	-7.87	-976.2	-6.20	4.8×10^{-8}	6.3	-8.5
	24	-72.4	-5.77	-1073	-7.34	1.9×10^{-8}	5.7	-7.7
1THL	25	-112.9	-7.91	-1069	-9.18	1.0×10^{-6}	6.4	-8.7

4.1.2 Re-docking of 8 ligands from X-ray structure complexes

The ligands known from X-ray structure complexes of thermolysin were re-docked using two approaches. Glu143 in the TLN typically interacts with acceptor oxygen of bound ligand and most probably is in a protonated form. This residue therefore needs to be protonated, during docking as may His231, which forms a salt bridge with Asp 226 and typically places its second ring nitrogen atom within 3 Å of a second zinc bound oxygen of a carboxylate or phosphonate ligand [101] [144]. So, in one of the two docking approaches, Glu143 was protonated on OE2 position, while His142 and His146 were protonated on δ-N position. In the other re-docking approach, Glu143 was deprotonated while His142 and His146 were still protonated. A set of 8 ligands known from the thermolysin-inhibitor X-ray crystal structures were docked into two scenarios (un-protonated Glu143 and protonated Glu143).The binding pocket was assigned by ICMpocket finder algorithm [4]. Each of the 8 ligand-thermolysin complexes was docked in 5 runs. The best poses were selected based on RMSD to original crystal structrures (Table 4.4 and 4.5). Docking energy was averaged over the 5 (Table 4.2 and 4.3). To make the comparing more objective, chart illustration were made for both energy value and RMSD (Figure 4.1, 4.2, 4.3 and 4.4). The general trend between the ligand flexibility and docking accuracy were analyzed.

Table 4.2: Docking energy (kcal/mol) after docking into thermolysin with charged Glu143. The ligands were also charged.

Complex	Rotatable bond	Previous docking energy	Re-dock 1	Re-dock 2	Re-dock 3	Re-dock 4	Re-dock 5	Average Energy
1QF2	**8**	-91.6	-89.9	-85.8	-90.2	-75.4	-83.7	-85.0
1THL	**12**	-112.9	-110.8	-103.6	-111.5	-100.1	-106.0	-106.4
1QF0	**13**	-102.4	-97.1	-91.6	-94.1	-94.1	-97.1	-94.8
1OS0	**13**	-112.2	-109.1	-104.5	-99.9	-99.3	-102.7	-103.1
1QF1	**13**	-76.7	-87.6	-88.0	-77.8	-88.3	-89.8	-86.3
5TMN	**16**	-120.3	-116.1	-94.1	-107.6	-100.7	-96.0	-102.9
6TMN	**16**	-123.9	-105.9	-92.7	-93.5	-100.0	-113.5	-101.1
1Y3G	**16**	-118.3	-105.8	-99.0	-110.6	-108.3	-119.3	-108.6

Table 4.3: Docking energy (kcal/mol) after docking into thermolysin with neutral Glu143 and charged ligands.

Complex	Rotatable bond	Previous docking energy	Re-dock 1	Re-dock 2	Re-dock 3	Re-dock 4	Re-dock 5	Average Energy
1QF2	**8**	-91.6	-81.2	-81.6	-98.3	-87.1	-83.9	-86.4
1THL	**12**	-112.9	-94.2	-113.2	-101.6	-78.8	-107.6	-99.1
1QF0	**13**	-102.4	-98.6	-98.4	-108.5	-104.8	-104.7	-103.0
1OS0	**13**	-112.2	-105.0	-112.1	-111.8	-110.0	-105.2	-108.8
1QF1	**13**	-76.7	-83.2	-93.2	-86.8	-88.7	-89.0	-88.2
5TMN	**16**	-120.3	-105.3	-118.0	-104.7	-109.4	-105.5	-108.6
6TMN	**16**	-123.9	-109.5	-107.9	-108.0	-109.3	-99.9	-106.9
1Y3G	**16**	-118.3	-119.0	-121.2	-102.8	-100.4	-111.4	-111.0

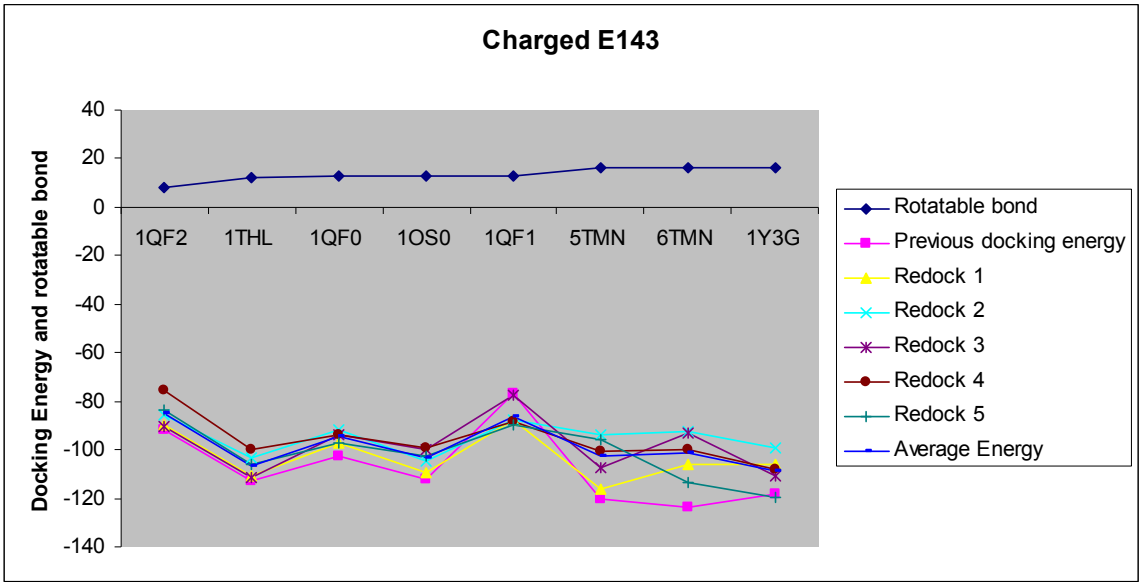


Figure 4.1: Chart illustration of docking energy values (kcal/mol) with charged Glu143 receptor.

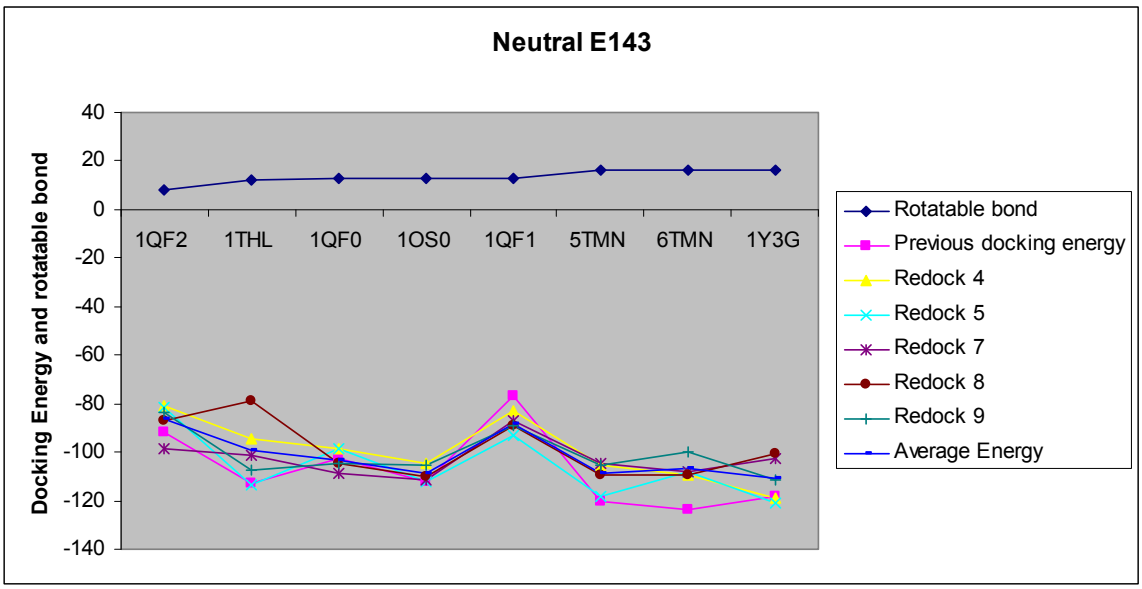


Figure 4.2: Chart illustration of docking energy values (kcal/mol) with neutral Glu143 receptor.

Table 4.4: Corresponding sRMSD values relative to the x-ray structure with charged Glu143 and charged ligand.

Complex	Rotatable bond	Previous sRMSD	Re-dock 1	Re-dock 2	Re-dock 3	Re-dock 4	Re-dock 5	Average sRMSD
1QF2	8	6.29	1.59	2.29	5.41	4.32	6.54	4.03
1THL	12	5.91	0.71	3.71	4.00	1.66	2.47	2.51
1QF0	13	2.32	3.15	2.52	1.97	2.47	2.81	2.58
1OS0	13	1.65	1.15	1.41	2.24	2.23	3.00	2.01
1QF1	13	4.96	0.91	1.04	1.69	1.03	1.60	1.25
5TMN	16	2.61	2.64	2.95	2.09	2.23	1.74	2.33
6TMN	16	2.64	2.50	1.78	3.34	2.21	4.61	2.89
1Y3G	16	4.81	2.79	2.23	1.96	1.90	1.52	2.08

Table 4.5: Corresponding sRMSD values relative to the x-ray structure with neutral Glu143 and charged ligand.

Complex	Rotatable bond	Previous sRMSD	Re-dock 1	Re-dock 2	Re-dock 3	Re-dock 4	Re-dock 5	Average sRMSD
1QF2	8	6.29	0.55	2.15	2.44	3.27	3.27	2.33
1THL	12	5.91	1.68	2.68	2.48	3.12	1.31	2.25
1QF0	13	2.32	1.39	2.94	1.70	2.22	2.49	2.15
1OS0	13	1.65	1.58	1.40	1.39	1.39	1.60	1.47
1QF1	13	4.96	1.47	2.71	0.97	0.96	1.05	1.43
5TMN	16	2.61	2.18	2.38	2.75	2.35	2.31	2.39
6TMN	16	2.64	4.77	2.20	4.87	4.52	2.74	3.82
1Y3G	16	4.81	2.54	3.00	3.43	2.87	2.86	2.94

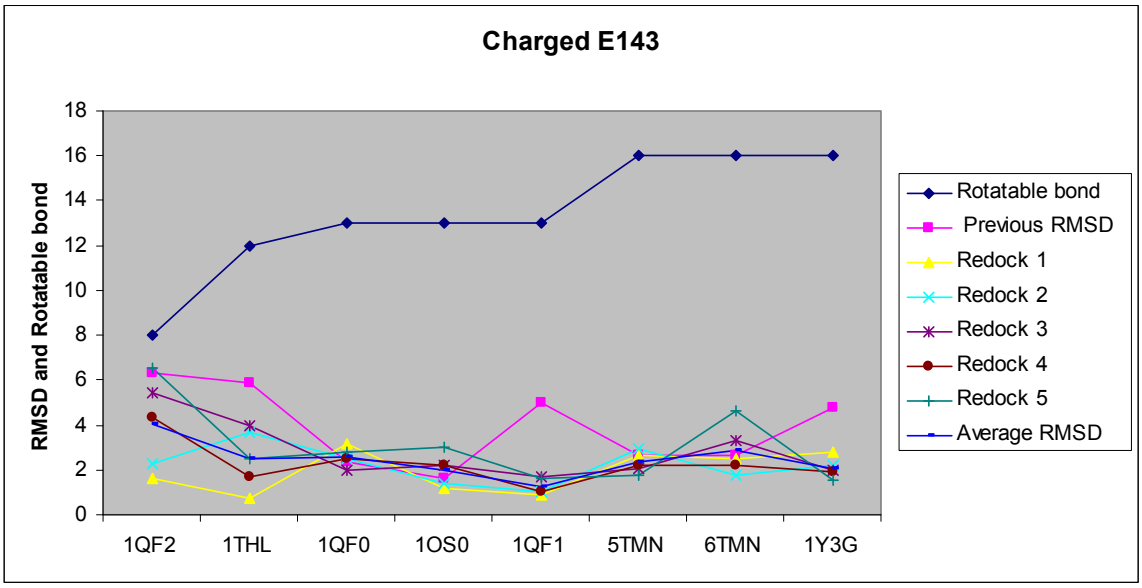


Figure 4.3: Chart illustration of sRMSD values with charged Glu143 receptor.

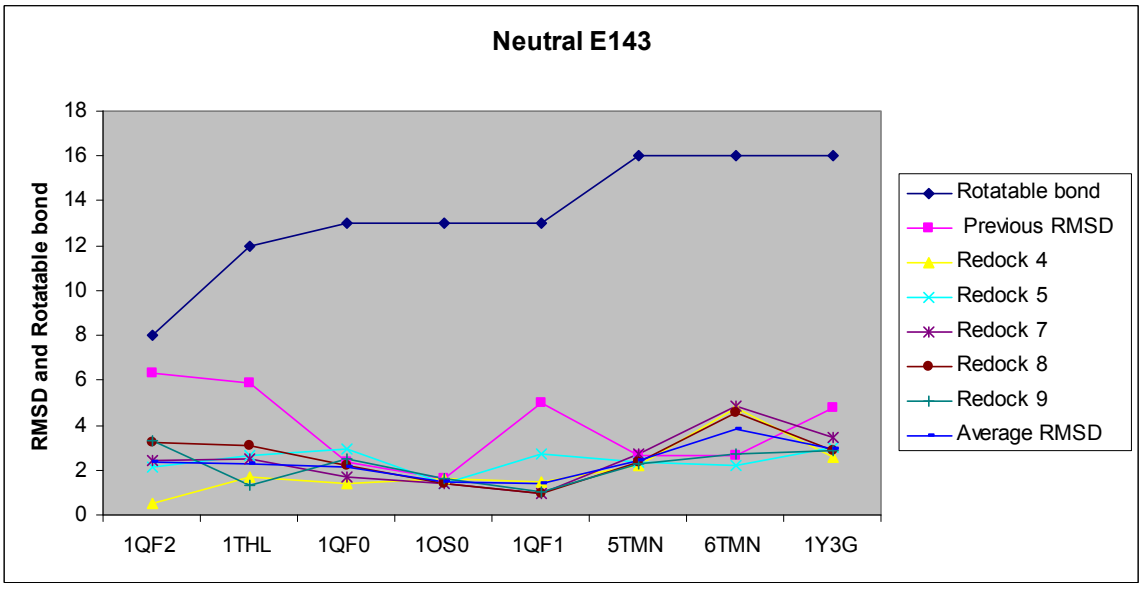


Figure 4.4: Chart illustration of sRMSD values with neutral Glu143 receptor.

Re-docking of the 8 ligands with two different scenarios (charged or neutral Glu143) indicated that Glu143 most probably should be in a neutral form. The docking energies did not favor one of the scenarios in front of the other, but the sRMSD values were in favor of a neutral Glu143.

4.1.3 Re-docking of the remaining 17 compounds of the 25 ligands

Inspired by the relatively better docking accuracy compared with initial docking for the 8 ligands (most similar with x-ray crystal structures) using neutral Glu143, the rest of 17 ligands were re-docked with similar grid maps and settings. The best poses were detected as described in the method part. The necessity of doing this was based on the fact that target protein and ligand are in constant motion between different conformational states with almost similar energies. So, instead of targeting a single pose of a given ligand in thermolysin, we tried to look for the most populated alternatives from an ensemble of stack conformations. Only small deviations of the best poses were seen between the 5 docking runs, but the energy was varying for some of the ligands (Table 4.6 and 4.7). The final results of the 25 ligands are shown in one Table 4.8.

Table 4.6: Highest scored docking energy values (kcal/mol) of the best pose in each of the 5 docking runs for the remaining 17 of the 25 ligands from the literature. Reference docking indicates the docking run with the docking energy of the best scored pose closet to the average of the best scored poses.

Ligand	Rotatable bond	Re-dock 1	Re-dock 2	Re-dock 3	Re-dock 4	Re-dock 5	Reference docking run	Average Energy
1	16	-127.1	-121.7	-128.3	-125.7	-120.8	4	-124.7
2	15	-107.8	-107.5	-107.7	-105.9	-107.2	5	-107.2
3	9	-68.3	-71.2	-68.7	-71.3	-71.2	5	-70.2
7	11	-88.6	-87.2	-87.5	-91.4	-88.0	1	-88.5
8	9	-70,9	-73,3	-71,6	-70,9	-67,9	4	-70,9
9	6	-84	-79,9	-84	-82,3	-82.3	5	-82.5
10	7	-60.7	-70.5	-70.8	-60.7	-71.4	2	-66.8
11	6	-65.4	-65.6	-65.3	-65.6	-65.6	2	-65.5
13	11	-99.5	-97.1	-96.5	-97.3	-96.6	4	-97.4
14	16	-125.8	-123.8	-128.9	-127.4	-127.8	1	-126.7
15	16	-119.5	-130.5	-124.6	-124.5	-133.8	3	-126.6
16	16	-128.7	-126.1	-132.9	-130.9	-130.7	5	-129.9
17	15	-107.2	-101.0	-104.3	-103.3	-107.2	3	-104.6
19	16	-120.6	-119.0	-115.8	-121.5	-117.3	2	-118.8
22	11	-124.9	-129.3	-136.1	-126.6	-127.6	5	-128.9
23	14	-103.5	-102.5	-104.6	-105.3	-106.9	4	-104.6
24	10	-83.0	-85.1	-82.5	-85.9	-84.3	2	-84.2

Table 4.7: sRMSD values of 5 runs for the remaining 17 of 25 compounds. This sRMSD was calculated between the reference docking run and the remaining best scored docking runs. These values indicate the popularity of the reference docking pose within the 5 runs (the lower of this sRMSD values, the more popular of the reference docking poses).

Ligand	Rotatable bond	Re-dock 1	Re-dock 2	Re-dock 3	Re-dock 4	Re-dock 5	Reference docking run.	Average sRMSD
1	16	3.39	3	2.79	/	3.91	4	3.27
2	15	2.27	0.09	0.67	1.99	/	5	1.25
3	9	0.16	0.01	0.21	0.02	/	5	0.1
7	11	/	2.8	1.47	5.78	0.46	1	2.63
8	9	0.02	1.78	1.31	/	2.2	4	1.33
9	6	1.81	3.42	1.82	0.01	/	5	1.77
10	7	2.02	/	2.02	0.13	0	2	1.04
11	6	0.05	/	0.06	0.01	0.01	2	0.03
13	11	2.43	2.36	2.06	/	1.19	4	2.01
14	16	/	1.87	1.73	1.15	1.12	1	1.47
15	16	1.83	4.07	/	0.81	2.27	3	2.24
16	16	0.13	3.14	2.98	1.21	/	5	1.87
17	15	2.17	1.21	/	1.96	1.68	3	1.76
19	16	1.5	/	1.89	1.53	2.85	2	1.94
22	11	2.19	2.97	2.78	3.08	/	5	2.76
23	14	2.45	1.91	2.76	/	3.11	4	2.56
24	10	1.29	/	2.53	2.24	2.03	2	2.02

4.1.4 Final result of the docking of 25 ligands

The final correlation between the Cal. ΔG and Exp. ΔG was calculated for regression analysis. The result showed that, the experimental and predicted binding energies (shown in Table 4.8) of the molecules were not significantly correlated ($R^2 = 0.56$, $p = 0.001$), but was improved compared with the initial docking (Table 4.1). Diversified chemical moieties may still contribute to the un-ideal R^2 values, but modified docking method may be the reason for improvement in binding prediction compared with the previous dockings.

Table 4.8: Ultimate poses and corresponding values for 25 compounds.

X-ray structure (if available)	Comp.	Rotatable bonds	K_i (in M)	pK_i	Docking energy (kcal/mol)
	1	16	4.1×10^{-8} [137]	7.4	-125.7
	2	15	1.0×10^{-8} [137]	8.0	-107.5
	3	9	2.0×10^{-7} [145]	6.7	-71.2
1QF2	**4**	8	1.2×10^{-6} [145]	5.9	-81.2
1QF0	**5**	13	4.2×10^{-8} [145]	7.4	-108.5
1QF1	**6**	13	4.8×10^{-8} [145]	7.3	-88.7
	7	11	1.9×10^{-8} [145]	7.7	-88.6
	8	9	1.0×10^{-6} [145]	6.0	-70.9
	9	6	1.6×10^{-6} [145]	5.8	-82.3
	10	7	1.6×10^{-6} [145]	5.8	-70.5
	11	6	3.5×10^{-7} [145]	6.5	-65.6
1OS0	**12**	13	9.3×10^{-7} [146]	6.0	-111.8
	13	11	7.8×10^{-7} [146]	6.1	-97.3
	14	16	2.6×10^{-8} [146]	7.6	-128.9
	15	16	4.5×10^{-8} [147]	7.4	-124.6
	16	16	6.8×10^{-11} [147]	10.2	-130.7
	17	15	9.1×10^{-9} [148]	8.0	-104.3
1Y3G	**18**	16	4.0×10^{-8} [14]	7.4	-119.0
	19	16	1.0×10^{-8} [14]	8.0	-119.0
5TMN	**20**	16	9.1×10^{-9} [14]	8.0	-118.0
6TMN	**21**	16	9.0×10^{-3} [14]	2.1	-107.9
	22	11	8.0×10^{-8} [52]	7.1	-127.6
	23	14	5.6×10^{-7} [52]	6.3	-105.3
	24	10	2.2×10^{-6} [52]	5.7	-85.1
1THL	**25**	12	3.8×10^{-7} [52]	6.4	-94.2

The major drawback of hydrogen bond calculator by the LigPlot program is that it was designed for computing hydrogen bonds between protein side chains. Therefore LigPlot does not plot hydrogen bonds between metal ion and ligand. Manual inspections of the complexes in Table 4.8 were therefore needed to detect zinc interactions. Hydrogen bonding and zinc interactions of re-docked 25 ligands in Table 4.9 are shown schematically in the following figures.

For compound **1**, all the detailed interactions are shown in Figure 4.5.1 and 4.5.2. The atoms C1, C7 and C9 of compound 1 showed hydrophobic interactions with His231. The atoms C4 and C6 interacted with Asn112 (CG), Phe130 and Leu202; C10 and C11 interacted with Phe114; C12 interacted with Tyr157 and Trp115 (CD1); C14 interacted with Asn116, Tyr157 and Trp115 (CD1); C16 interacted with Asn116; C17 interacted with Phe114 and Asn116; C19 and C22 interacted with His142 and Leu202.

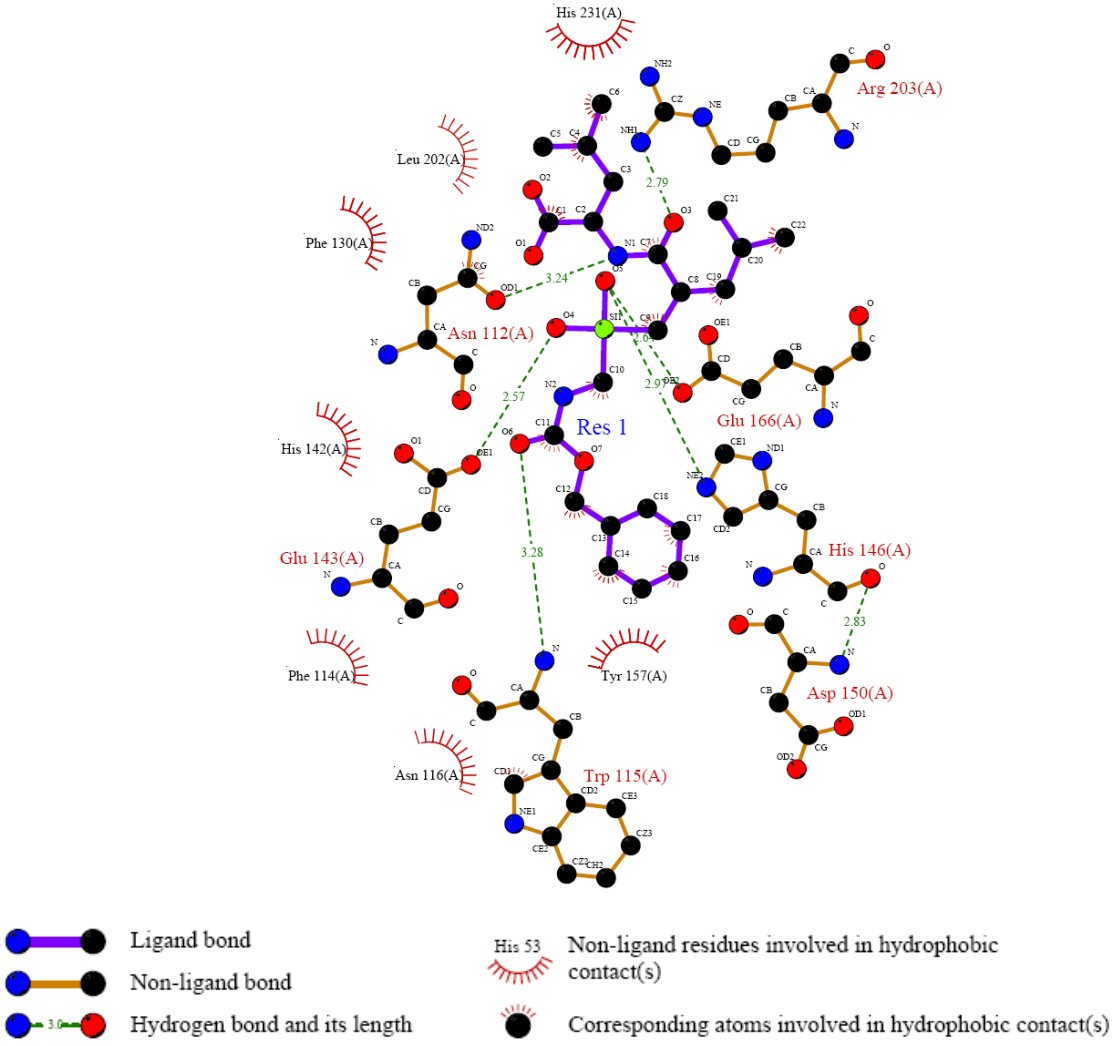


Figure 4.5.1: Schematic diagram of the interactions of compound **1** within the active site of thermolysin (Target PDB code 1gxw). The schematic 2D interaction plot between the thermolysin active site and compound **1** was produced using the program LigPlot [141]. Labels of different bond type, hydrophobic residues and corresponding atoms involved in hydrophobic contact (s) interpreted in this figure apply to the whole LigPlot schemes in this study.

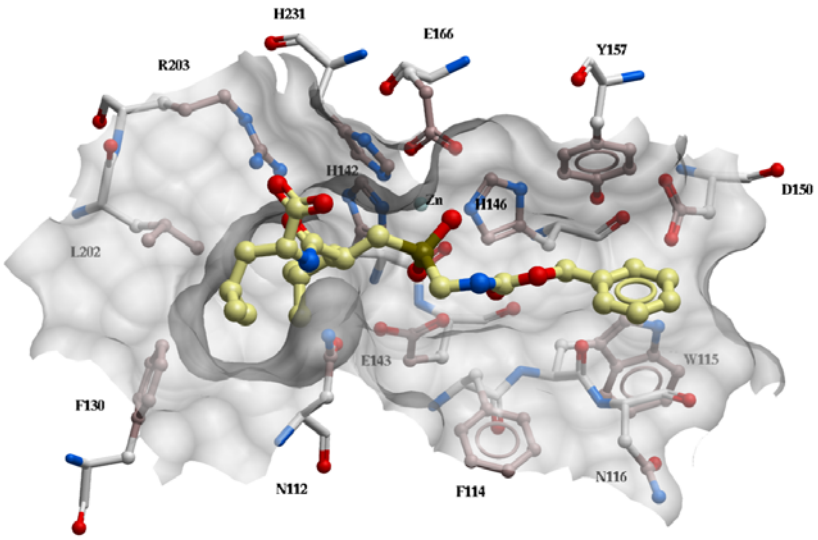


Figure 4.5.2: Stereo-chemical relationship between compound 1 and active site of thermolysin. Binding pocket is shown in transparent mode. Amino acids are corresponding to Ligplot of compound 1.

The detailed interactions of compound **2** are shown in Figure 4.6.1 and 4.6.2. For compound 2, atoms C3, C13, C14, and C15 showed hydrophobic interactions with His231, while C5 and C8 interacted with His142; C11 interacted with Phe130; C17 and C18 interacted with Tyr157 and His146 (CG, CD2), and C19 and C20 interacted with Tyr157 only.

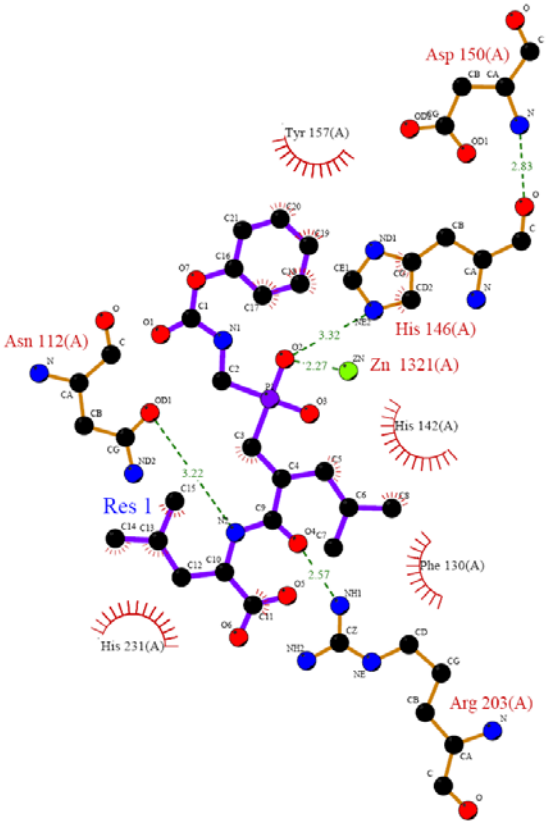


Figure 4.6.1: Schematic diagram of the interactions of compound **2** with the active site of thermolysin (Target PDB code 1gxw).

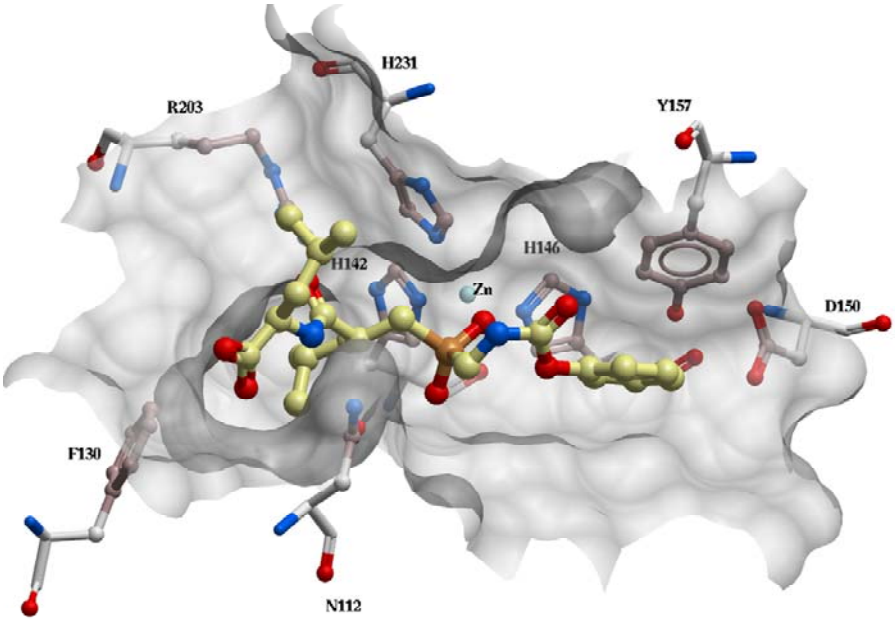


Figure 4.6.2: Corresponding (to Figure 4.6.1) stereoview of compound 2 at the active site of thermolysin.

The detailed interactions of compound **3** are shown in Figure 4.7. In case of compound **3**, S1 showed hydrophobic interactions with His231 (CD2); C4 also showed hydrophobic interactions with His231 but with other atoms (CG, CB, and CD2) of the same residue, C9 also interacted with His231 (atoms CD2, CD, and CB). C8 interacted with Asn112 (CG); C13 interacted with Leu133 and Phe130; C14 and C15 interacted with Leu133, Phe130, Val139 and Leu202; C16 interacted with Val139 and Leu202; and C17 interacted with Leu202.

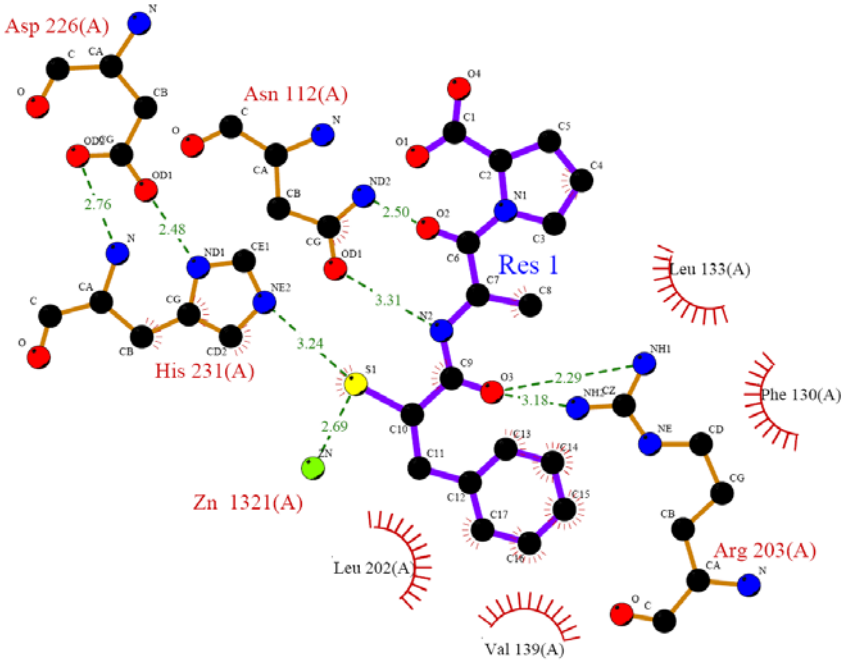


Figure 4.7: Schematic diagram of the interactions of compound **3** with the active site of thermolysin (Target PDB code 1gxw).

The detailed interactions between compound 4 and thermolysin are shown in Figure 4.8. For compound **4**, S1 showed hydrophobic interactions with His231; C4 interacted with Leu133, Phe130 and Val139; C5 interacted with Leu202; C6 and C8 interacted with Phe130, Leu133, Val139 and Leu202; C7 interacted with Leu133, Val139, Leu202 and His231; C9 interacted with His231 and Leu202; atom C10 interacted with Asn112 (CG); C13 and C14 interacted with His231 and atoms C17 and 19 interacted with Leu202 and His231.

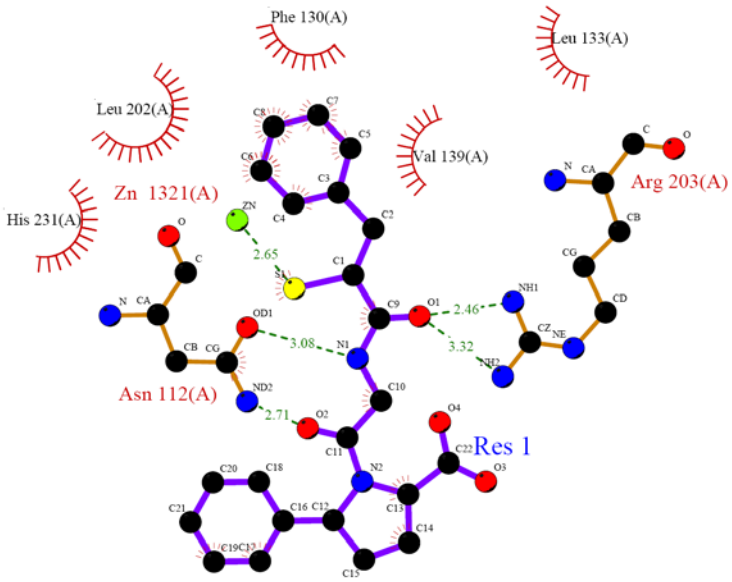


Figure 4.8: Schematic diagram of the interactions of compound **4** with the active site of thermolysin (Target PDB code 1gxw).

LigPlot and stereoscopic representations of compound **5** are shown in Figure 4.9.1 and 4.9.2. For compound 5 atom S1 showed hydrophobic interactions with Phe114 and His146. C1, C3, C5, C7, and C8 interacted with Phe114; C4 interacted with Asn112 (CG); C15 interacted with Leu133, Phe130, Leu202 and Ile188; C16 and C17 interacted with Leu202, Val139, and Ile188. C17 also interacted with Leu133. C13, C22 and C23 interacted with Phe130, Leu133 and Leu202. C19 and C27 interacted with His231 and C19 also interacted with Asn112 (CG); and C26 interacted with Phe130. The active site metal ion Zn showed ligand bonding (purple line in Figure 4.9.1) with O1 atom of compound **5**.

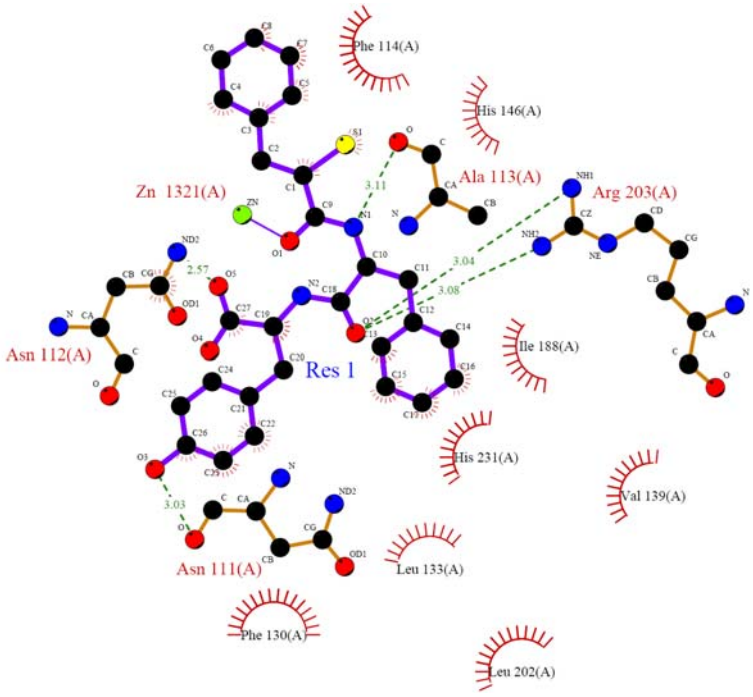


Figure 4.9.1: Schematic diagram of the interactions of compound **5** with the active site of thermolysin (Target PDB code 1gxw).

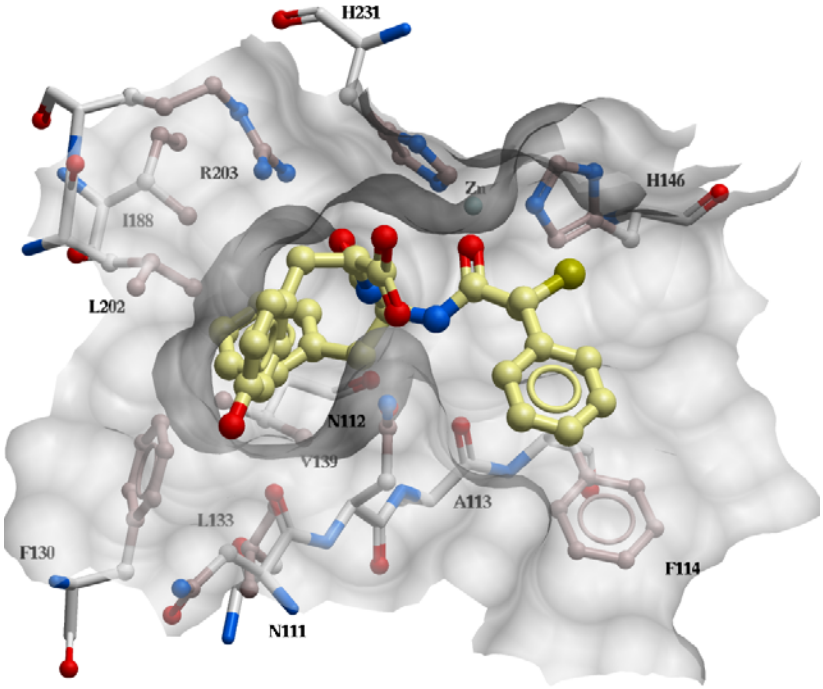


Figure 4.9.2: Corresponding (to Figure 4.9.1) balls and sticks view of compound 5 at the active site of thermolysin.

Figure 4.10.1 and 4.10.2 show the detailed interaction of compound **6** with thermolysin. In this compound, atoms C1, 4 and 5 showed hydrophobic interactions with Phe114. C11 and C13 interacted with Phe130, Leu133, Val139 and Leu202; C14 interacted

with Val139, Leu133 and Ile188; C15 interacted with Phe130, Leu133, Val139, Ile188 and Leu202; C19 interacted with His231.

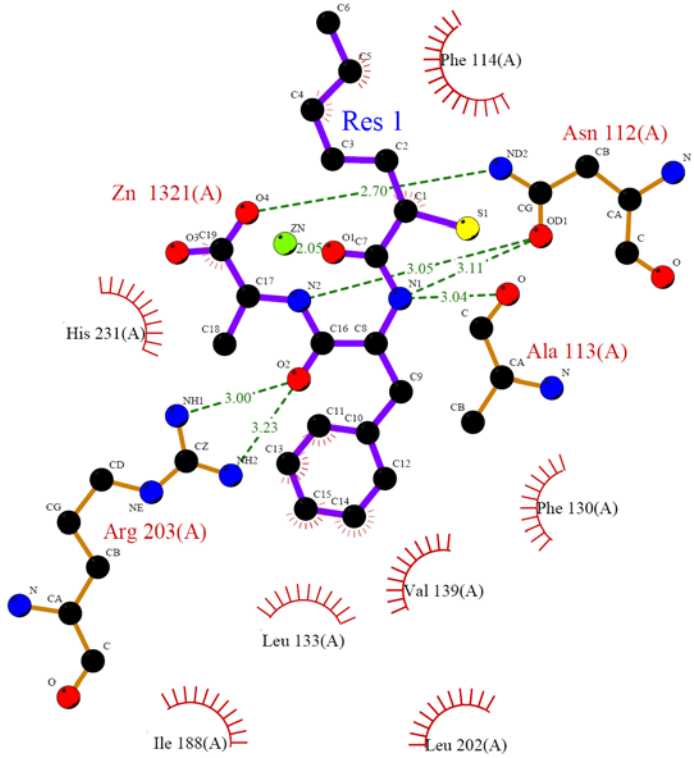


Figure 4.10.1: Schematic diagram of the interactions of compound **6** with the active site of thermolysin (Target PDB code 1gxw).

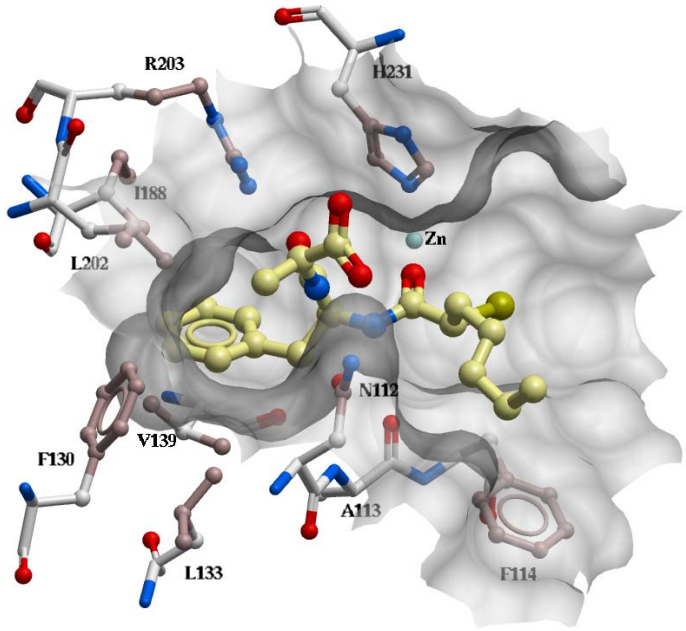


Figure 4.10.2: Corresponding (to Figure 4.10.1) stereoview of compound 6 at the active site of thermolysin.

For compound **7**, atoms S1, C20 and C21 showed hydrophobic interactions with Trp115, Phe114, and Glu166, but C21 did not show interaction with Glu166 (Figure 4.11.1

and 4.11.2). C3 interacted with His231; C4 interacted with Asn112 (CG) and C8 interacted with Leu202; C9 and C10 interacted with Leu202, Ile188 and Val139; C10 also interacted with His142 (CB); C11 and C12 interacted with His142 (CD2, CG, CB) and Val139; while C15, C16, C17 and C18 interacted with Phe114.

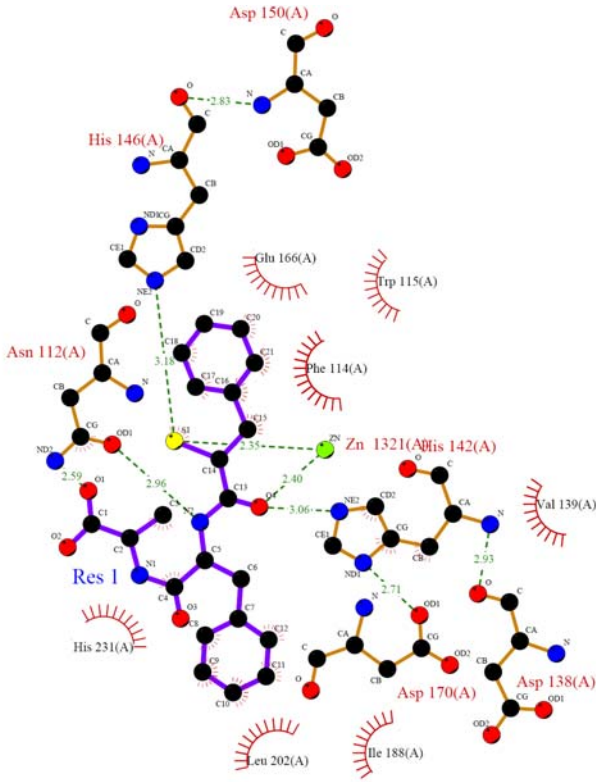


Figure 4.11.1: Molecular interactions of compound **7** with the active site of thermolysin (Target PDB code 1gxw).

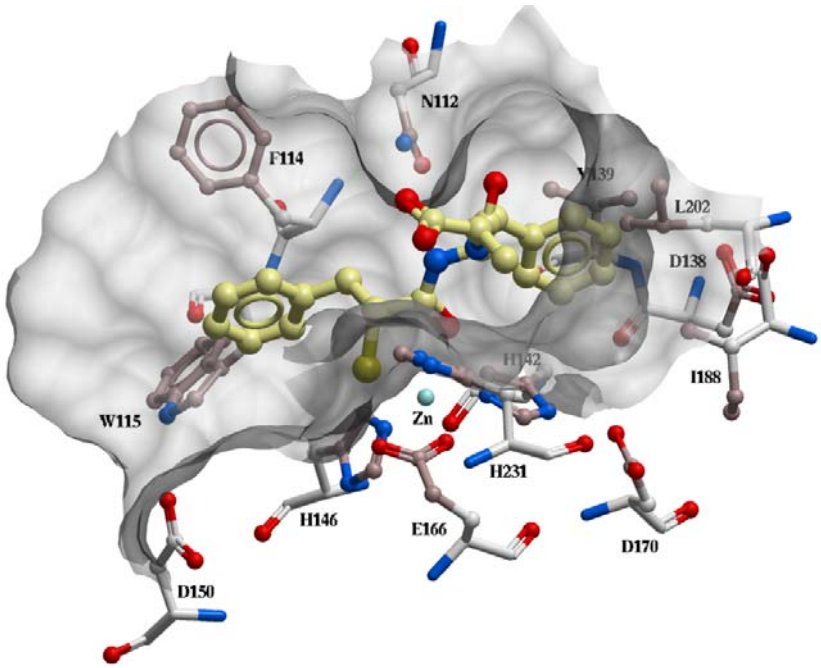


Figure 4.11.2: Corresponding (to Figure 4.11.1) stereoview of compound 7 at the active site binding pocket of thermolysin.

The detailed interactions of compound **8** with thermolysin are shown in Figure 4.12. Atom S1 of the ligand showed hydrophobic interactions with Phe114 of thermolysin, and C3 interacted with His231. C10 and C11 interacted with Phe130, Val139, Leu133 and Leu202. C12 interacted with Phe130, Val139, Leu133, Leu202 and Ile188. C13 interacted with Val139 and Ile188 residues of the active site of thermolysin.

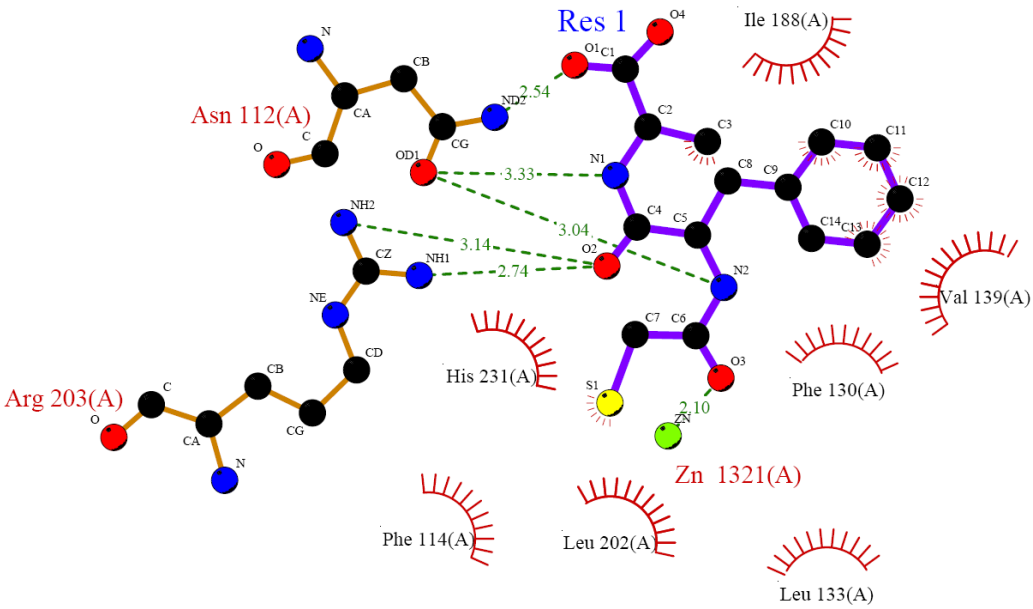


Figure 4.12: Molecular interactions of compound **8** with the active site of thermolysin (Target PDB code 1gxw).

For compound **9** (Figure 4.13), atom S1 showed hydrophobic interaction with His231 (CD2) residue, while C11 had hydrophobic interaction with the same residue but with different atoms (CD2, CG and CB) of the residue. C14 also interacted with His231, but with atoms CD2 and CG. Atom C16 interacted with His142. C18 and C19 interacted with Val139 and Leu202. C20 interacted with Val139, Leu202 and Phe130, and C21 interacted with Phe130, Val139, Leu133 and Leu202.

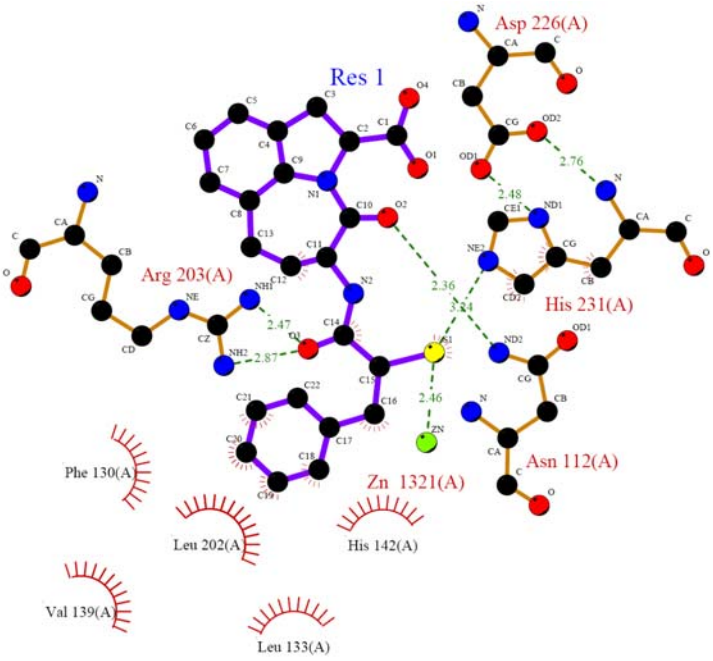


Figure 4.13: Molecular interactions of compound **9** with the active site of thermolysin (Target PDB code 1gxw).

Figure 4.14 shows that S1 of compound **10** had hydrophobic interactions with His142, His146, His231 and Glu143 (long distant interaction). These residues are mostly the catalytic residues of the active site of thermolysin and the hydrophobic interactions with S1 seem important for the binding of the ligand within the active site. C3 also showed hydrophobic interactions with His231 and Leu202. C5 interacted with His231; C6 with His142 and His146; C8 with Leu202 and Ile188; C9 with Val139, Leu202, Ile188 and Leu133; atom C10 interacted with Phe130, Leu202, Val139, Leu133 and Ile188. C11 interacted with Phe130, Leu133, Val139 and Leu202.

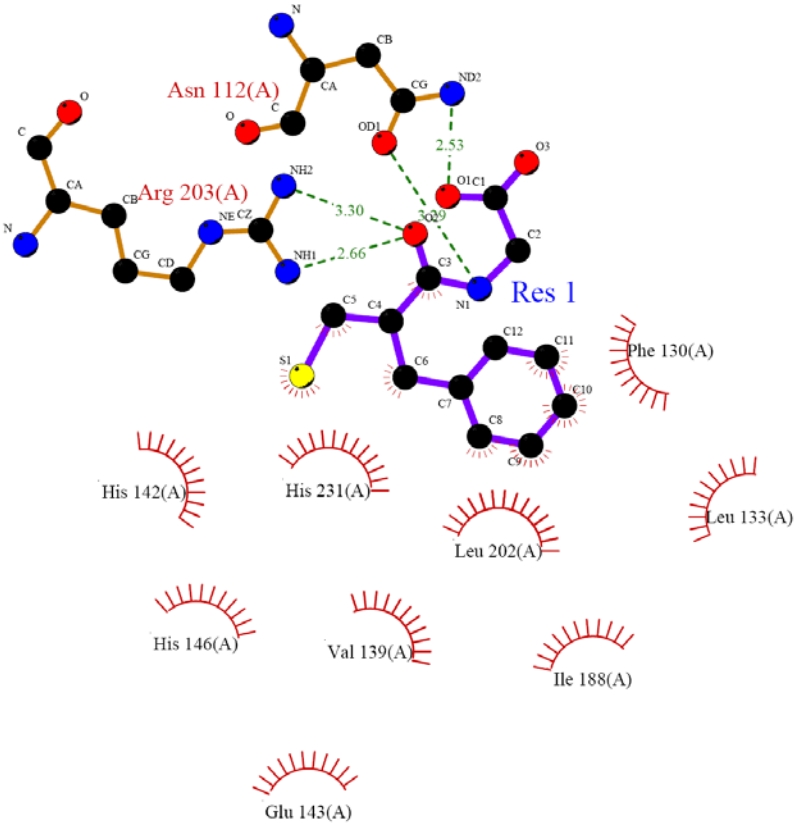


Figure 4.14: Molecular interactions of compound **10** with the active site of thermolysin (Target PDB code 1gxw).

For compound **11**, all the detailed interactions with thermolysin are shown in Figure 4.15. Atom S1 showed hydrophobic interactions with His231 and His142 (CD2). Atom C3 interacted with Asn112 (CG); C6 with His142 (CD2); C8 and C9 interacted with Ile188, Leu202 and Val139, while C10 and C11 interacted with Leu202, Val139, Leu133 and Phe130.

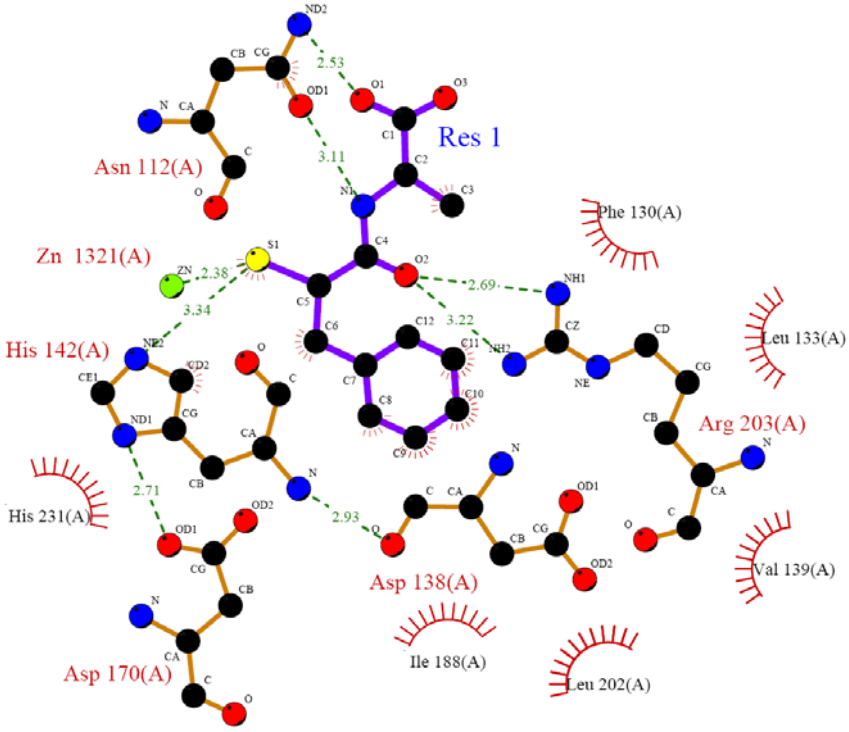


Figure 4.15: Molecular interactions of compound **11** with the active site of thermolysin (Target PDB code 1gxw).

Figure 4.16 is showing the LigPlot of compound **12** in the active site of thermolysin. Atoms C1, C3, C4, C5, C6, C7 and C8 of **12** showed hydrophobic interactions with Phe114 of thermolysin. Atoms C13, C15 and C17 interacted with Leu133, Leu202 Val139 and Phe130. C16 interacted with Leu133, Leu202 and Val139. C18 interacted with Leu202 and His231 (CD2). C21 interacted with Asn112 (CG) and C24 interacted with Leu202, Phe130 and His231 (CD2).

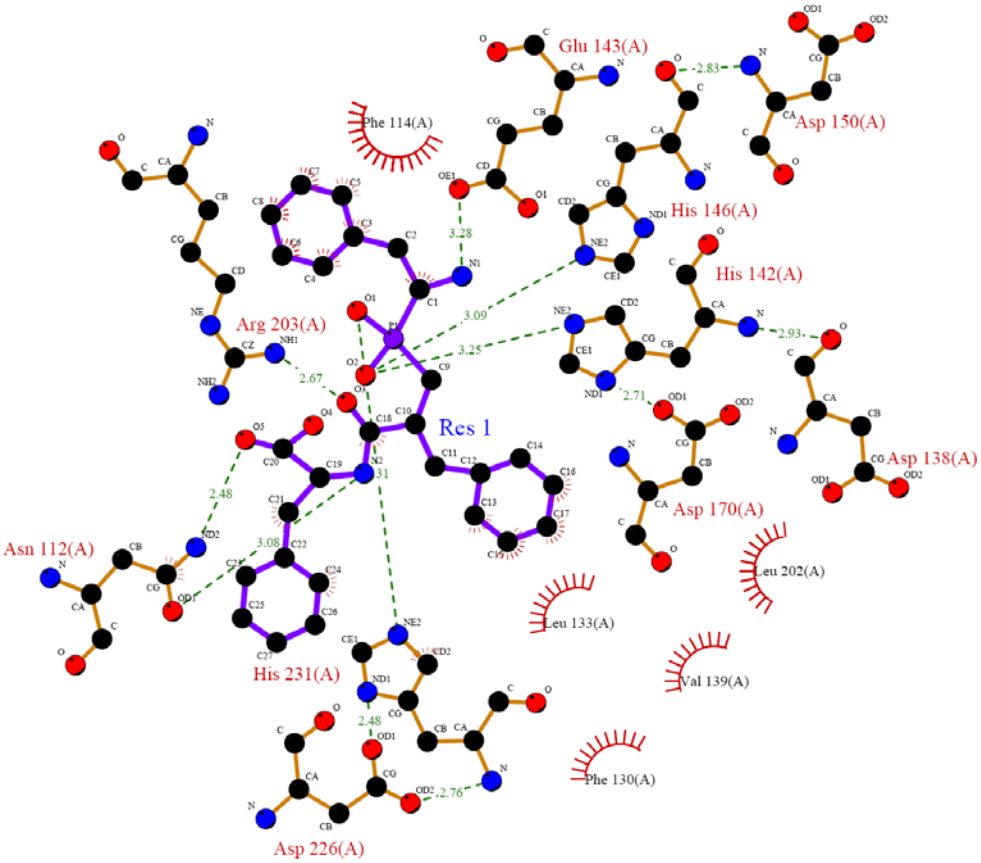


Figure 4.16: Molecular interactions of compound **12** with the active site of thermolysin (Target PDB code 1gxw).

The detailed interactions of compound **13** are shown in Figure 4.17. For this compound atoms C1, C2, C3, C4, C6 and C7 had hydrophobic interactions with Tyr157, Phe114 and His146. C9 showed hydrophobic interactions with His231 (CD2). C14 and C15 interacted with Leu202 and Val139. C16 interacted with Leu202, His142, Val139 and Ile188; atom C17 interacted with Ile188; C18 interacted with Leu202, His142 and Val139; and C21 interacted with Asn112 (CG). In this complex, hydrogen bond between Asp226 and His231 (2.48 and 2.76 Å) are important for the active site architecture.

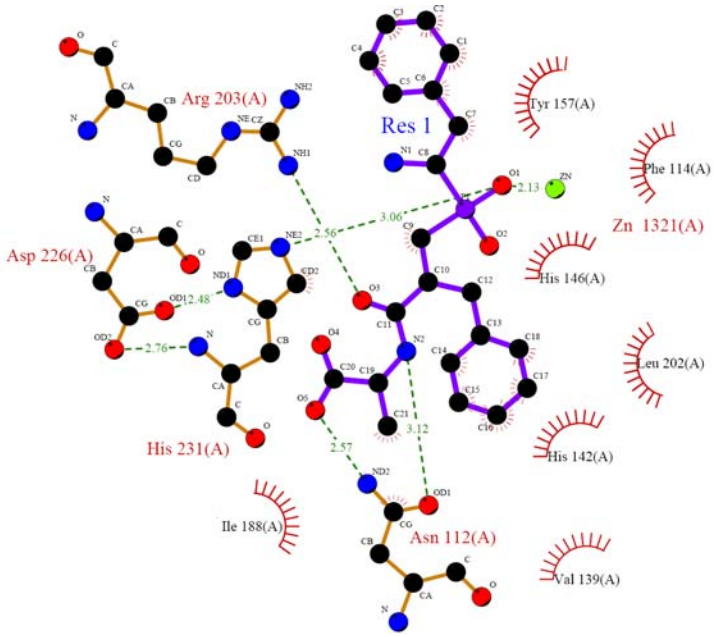


Figure 4.17: Molecular interactions of compound **13** with the active site of thermolysin (Target PDB code 1gxw).

The detailed interactions of compound **14** are shown in Figure 4.18.1 and 4.18.2. For compound **14** atoms C2 and C3 showed hydrophobic interactions with Asn112 (CG). Atom C4 interacted with Phe114 and Trp115. C13, C14 and C15 interacted with Trp115, Asn116, Gly117 and Tyr157. C16 interacted with Phe114, Trp115 Asn116 and Tyr157; C17 interacted with His142 and His231; C19 interacted with His142 and Leu202. C21 interacted with Leu202, His142 and Ile188. C22 and C23 interacted with Leu202, His142, Ile188 and Val139; and C24 interacted with Leu202, Leu133 and Val139.

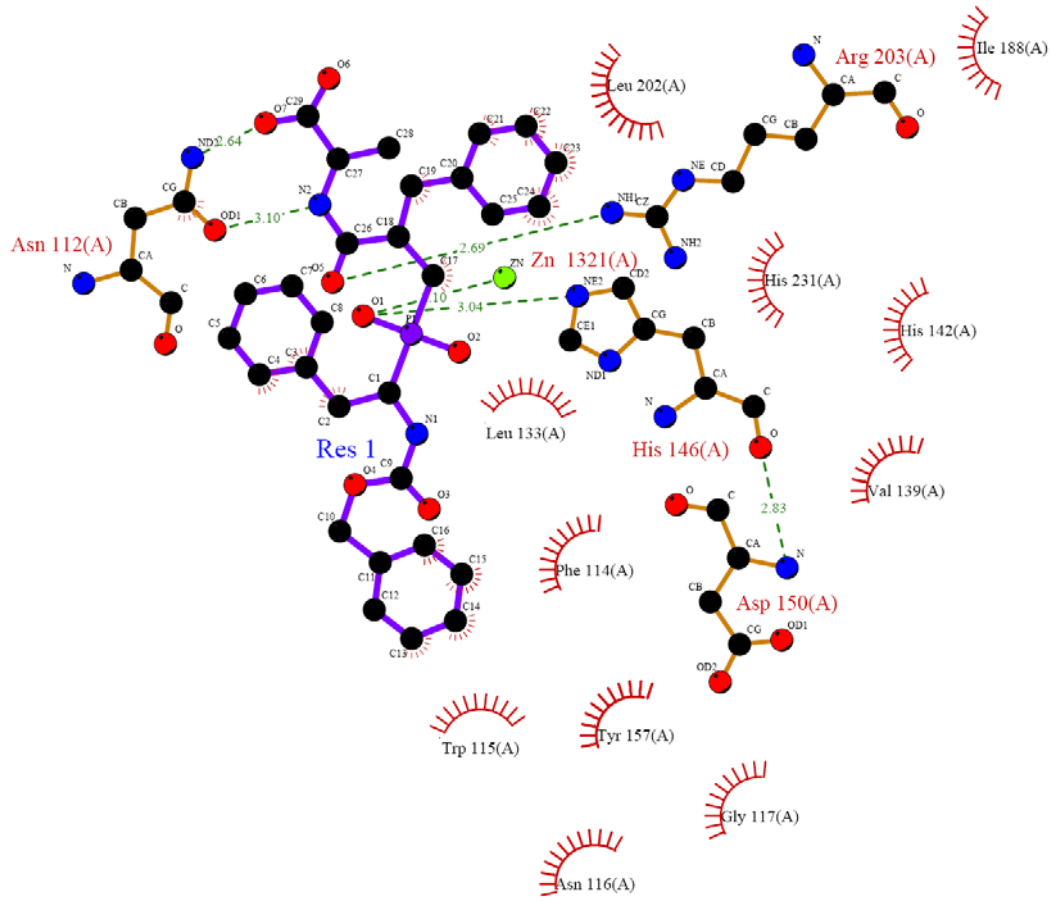


Figure 4.18.1: Molecular interactions of compound **14** with the active site of thermolysin (Target PDB code 1gxw).

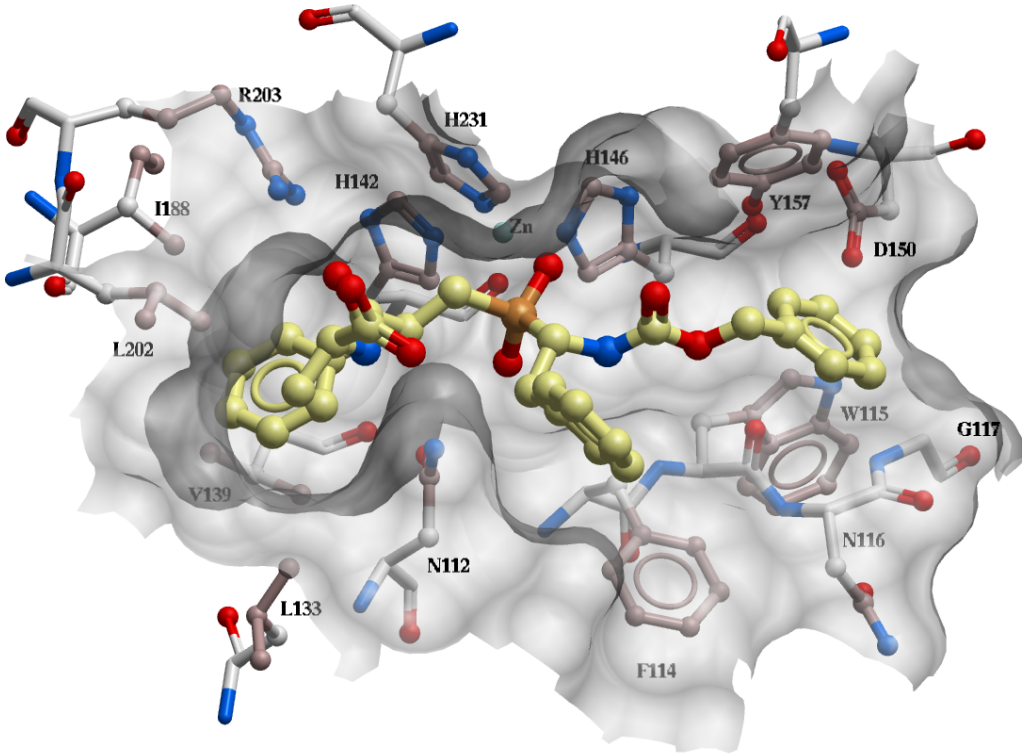


Figure 4.18.2: Corresponding (to Figure 4.18.1) stereoscopic view of compound 14 at the active site of thermolysin.

Detailed interactions of compound **15** are shown in Figure 4.19.1 and 4.19.2. Atoms C2, C3 and C4 showed hydrophobic interactions with Asn112 (CG). C10, C11, C12, C13, C14, C15 and C16 interacted with Tyr157. C18 and C21 interacted with His142.

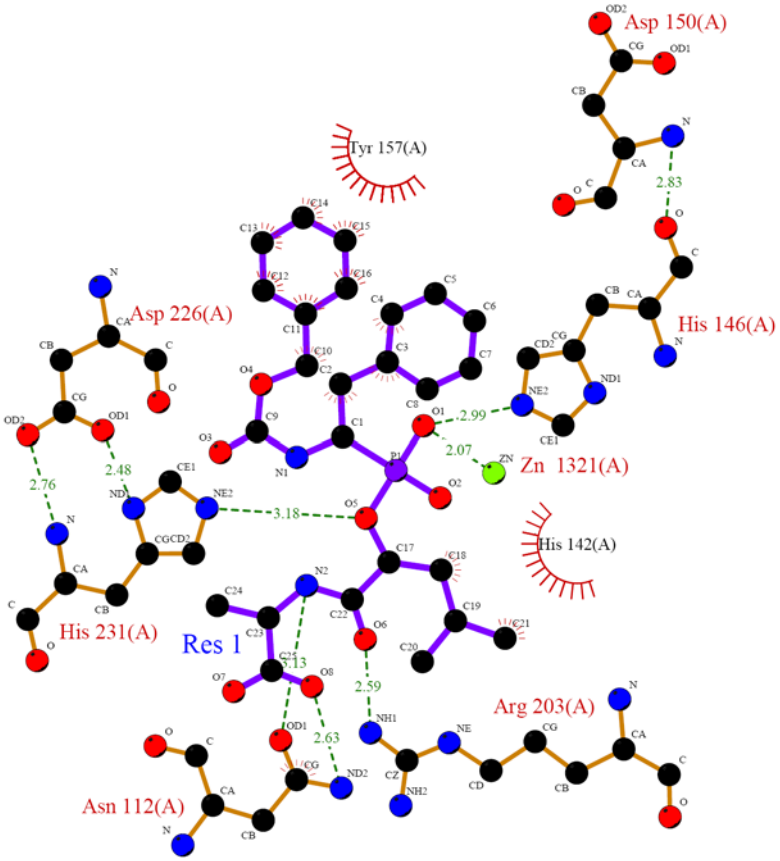


Figure 4.19.1: Molecular interactions of compound **15** with the active site of thermolysin (Target PDB code 1gxw).

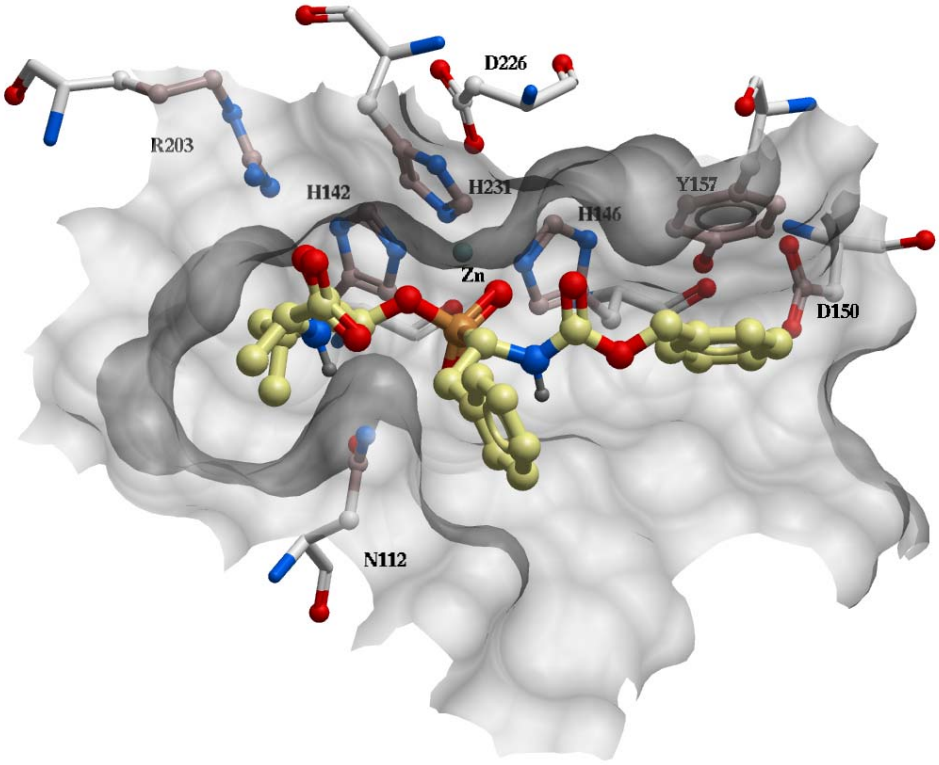


Figure 4.19.2: Corresponding (to Figure 4.19.1) stereoscopic view of compound 15 at the active site of thermolysin.

Figure 4.20.1 and 4.20.2 show the molecular interactions of compound **16** at the active site of thermolysin. For compound **16** atoms C2 and C3 exhibited hydrophobic interactions with Asn112 (CG) and C4 interacted with Phe114. Atoms C10, C12, C14 and C15 interacted with Tyr157. C14 and C15 also interacted with Asn116. C18 and C21 interacted with His142.

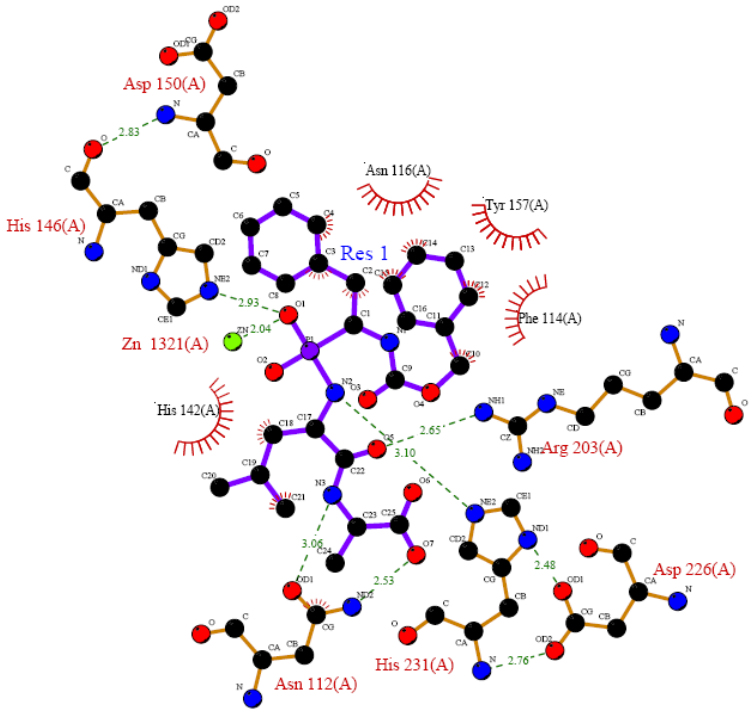


Figure 4.20.1: Molecular interactions of the compound **16**, the most potent inhibitor (K_i = 6.8×10^{-11} mM) of the 25 ligands, within the active site of thermolysin (Target PDB code 1gxw).

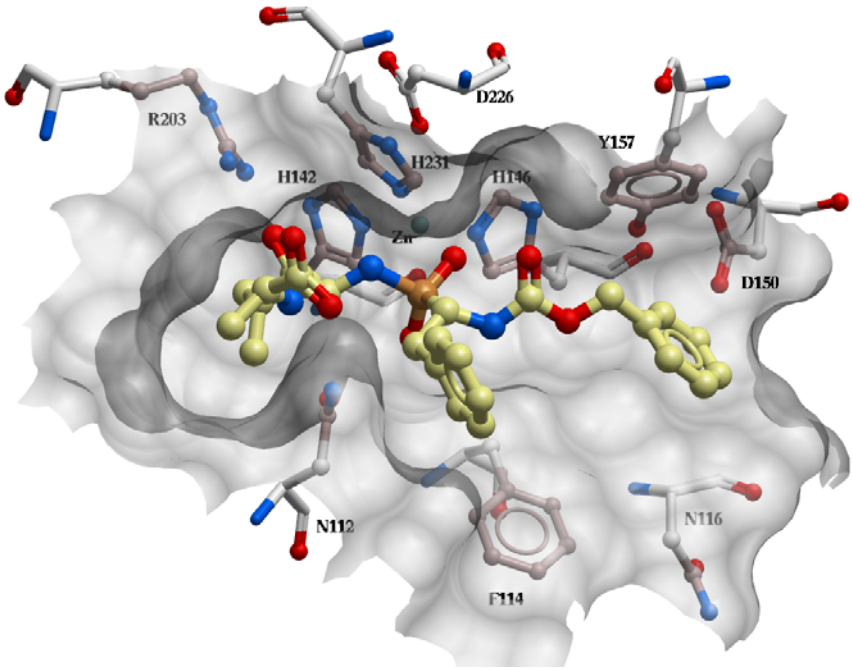


Figure 4.20.2: Corresponding (to Figure 4.20.1) stereoscopic view of compound 16 at the active site binding pocket of thermolysin.

For compound **17**, all detailed interactions are shown in Figure 4.21.1 and 4.21.2. For this compound, atoms C1, C6, C7 and C8 demonstrated hydrophobic interactions with Leu202; C7 and C8 also interacted with Val139 and Ile188. Atom C2 interacted with the His231 (CG, CD2 and CE1) residue. Both C5 and C8 interacted with the atoms of His142

(CG, CD2, CB); C11, C12 and C13 interacted with Trp115 and C12 and 13 also interacted with Tyr157. Finally C17 and C18 interacted with His231 (CG, CE1 and CD2).

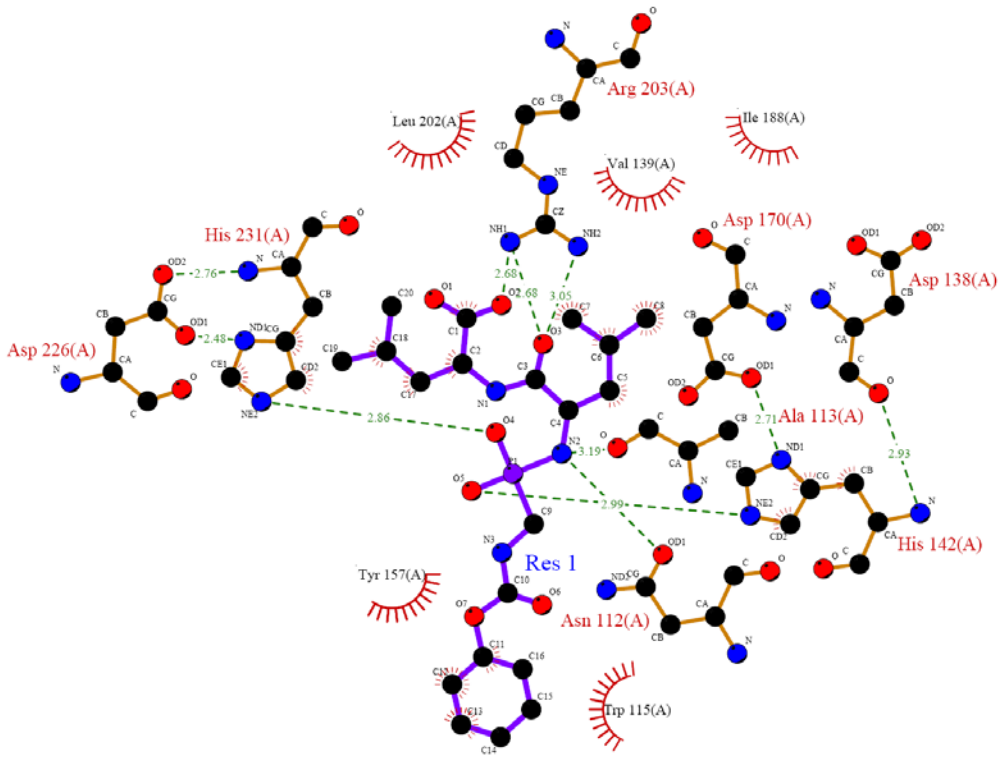


Figure 4.21.1: Molecular interactions of compound **17** with the active site of thermolysin (Target PDB code 1gxw).

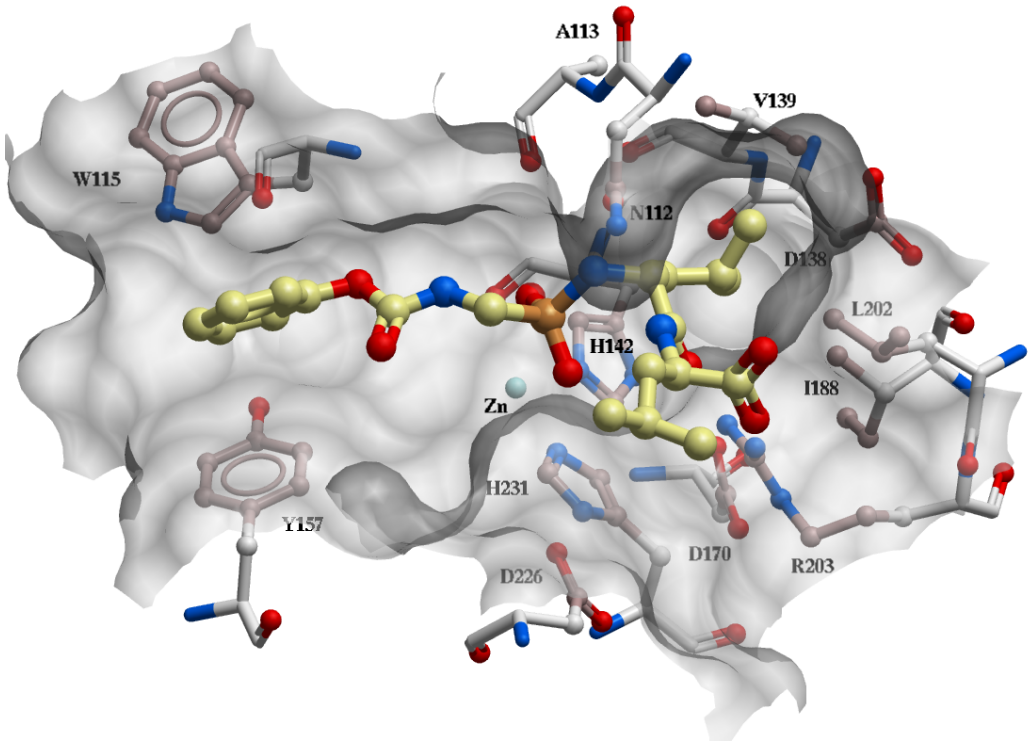


Figure 4.21.2: Corresponding (to Figure 4.21.1) stereoscopic view of compound 17 at the active site binding pocket of thermolysin.

For compound **18**, all detailed interactions are shown in Figure 4.22.1. and 4.22.2. Atoms C1 and C2 had hydrophobic interactions with His146 (CD2); C2 and 3 interacted with Tyr157. C8 and C19 interacted with Phe114 and His231 (CD2 and CG), respectively. Atoms C15, C16 and C17 showed interactions with Val139, Leu202, Leu133 and Phe130. C21 interacted with Asn112 and C22 with Leu202 and Phe130.

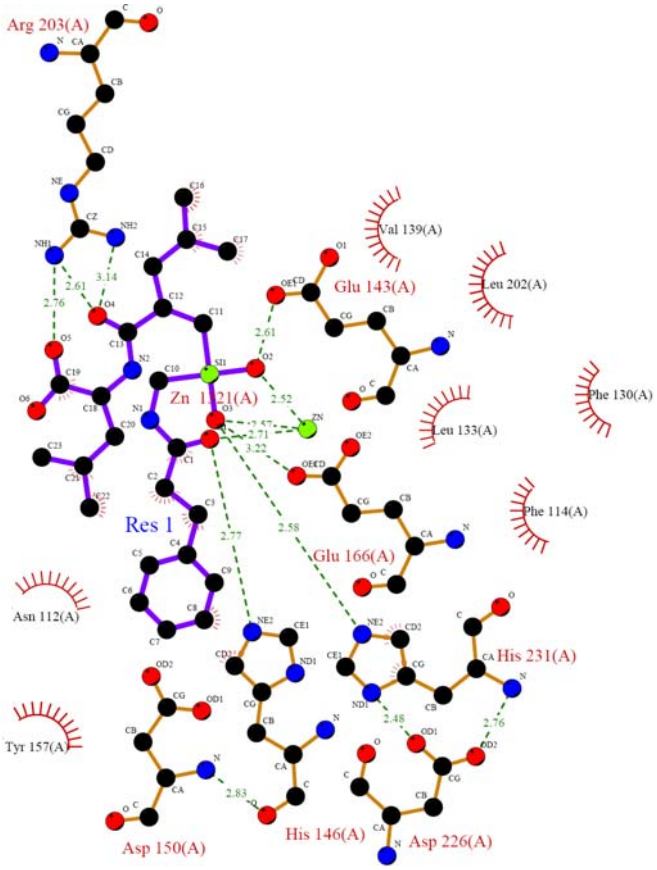


Figure 4.22.1: Molecular interactions of compound **18** with the active site of thermolysin (Target PDB code 1gxw).

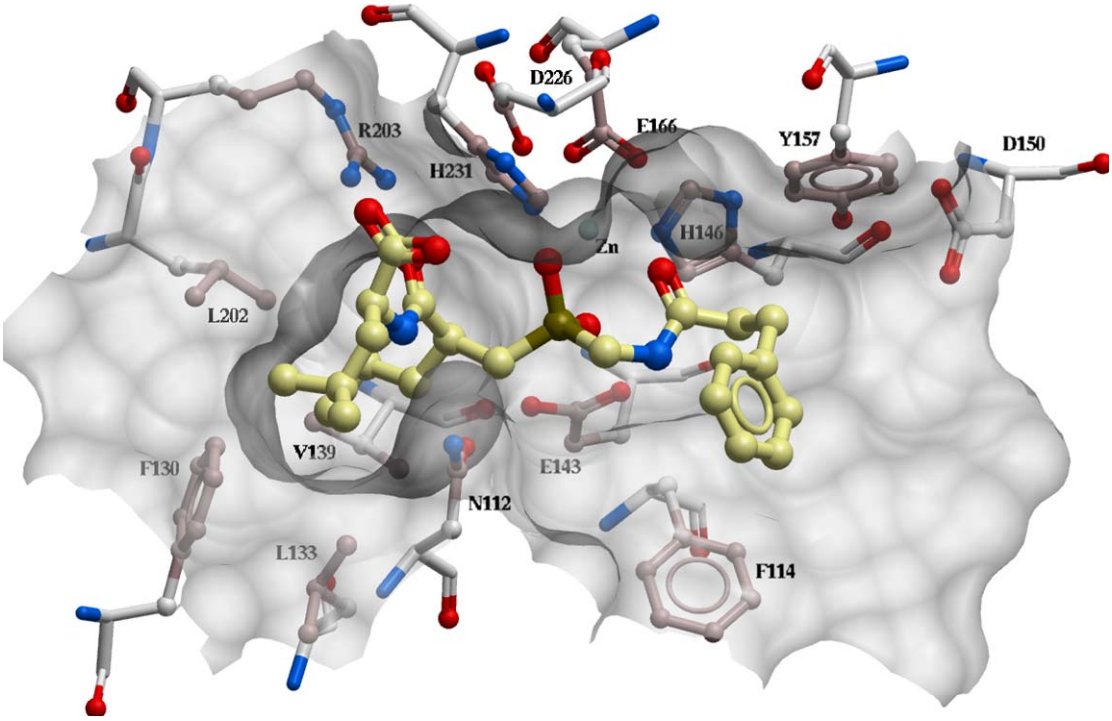


Figure 4.22.2: Corresponding (to Figure 4.22.1) stereoscopic view of compound 18 at the active site binding pocket of thermolysin.

For compound **19**, all the detailed interactions with thermolysin are shown in Figure 4.23.1 and 4.23.2. Atom C1 revealed hydrophobic interaction with His231 (CE1). Atoms C7 and C10 interacted with Leu202. C11 and 12 interacted with Leu202, Ile188 and Val139. Atom C11 also interacted with Asn112 (CG) and C12 also interacted with His142. Atom C13 interacted with Asn112 (CG). C16 interacted with His146 and atoms C18 and C19 with Tyr157.

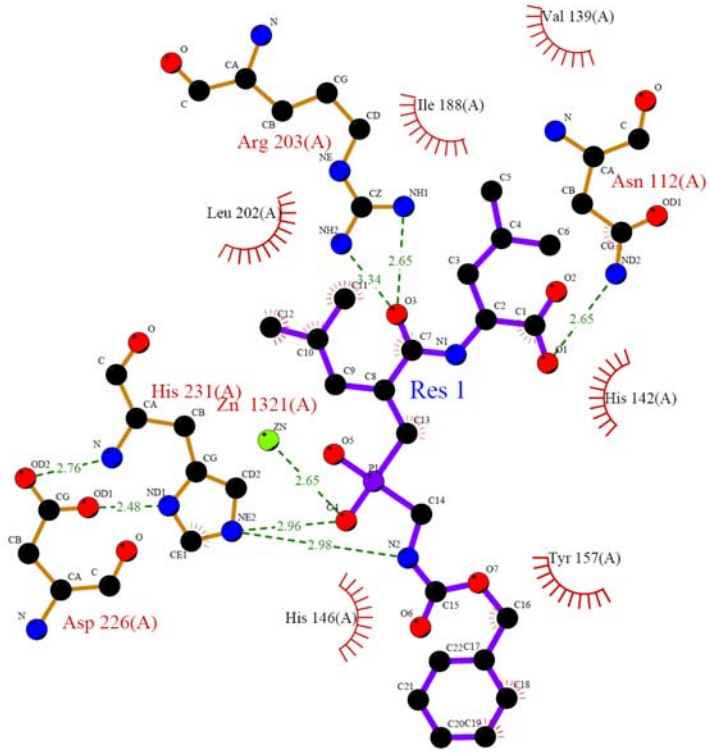


Figure 4.23.1: Molecular interactions of compound **19** with the active site of thermolysin (Target PDB code 1gxw).

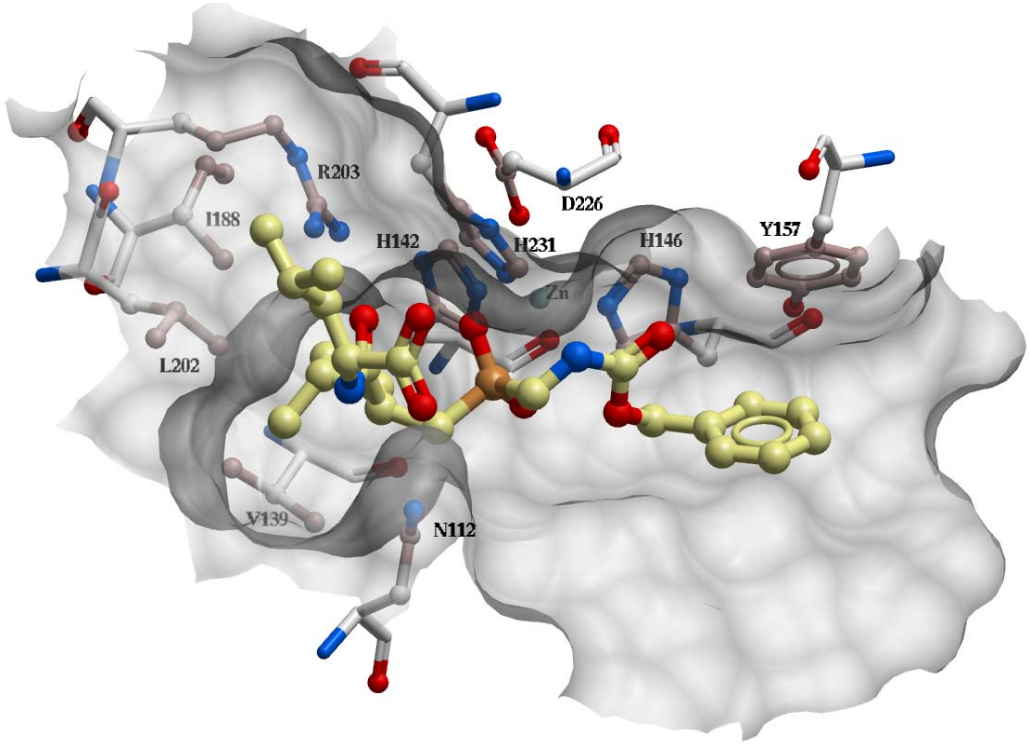


Figure 4.23.2: Corresponding (to Figure 4.23.1) stereoscopic view of compound 19 at the active site binding pocket of thermolysin.

For compound **20**, all the detailed interactions are shown in Figure 4.24.1 and 4.24.2. Atoms C2 and C4 exhibited hydrophobic interactions with Tyr157 and His146. Atom C13 interacted with Val139, Phe130 and Leu202; atoms C14 and 15 interacted with Val139 and Leu202; C14 also interacted with Leu133 and Phe130; and atoms C18 and C21 interacted with His231 (CE1, CG and CD2).

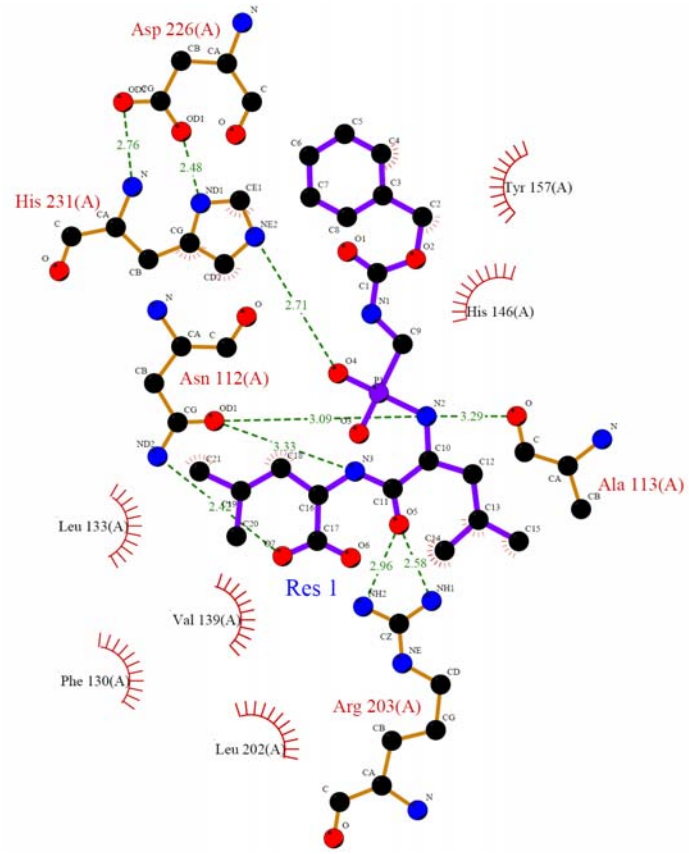


Figure 4.24.1: Molecular interactions of compound **20** with the active site of thermolysin (Target PDB code 1gxw).

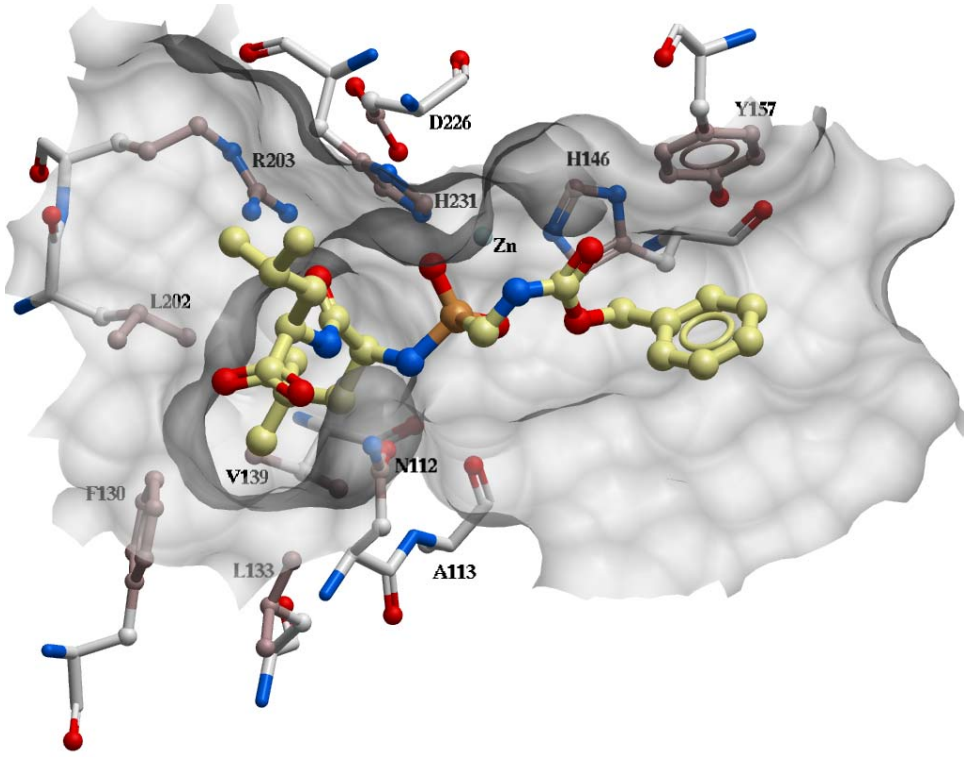


Figure 4.24.2: Corresponding (to Figure 4.24.1) stereoscopic view of compound 20 at the active site binding pocket of thermolysin.

Detailed interactions of compound **21** are shown in Figure 4.25. Atom C1 showed hydrophobic interactions with Tyr157, His146, Phe114 and Asn116. C2 interacted with Tyr157 and His146. C5 and C6 interacted with Phe114 and Asn116. C9 interacted with Phe114 and His146. C13 and C14 interacted with Leu202 and Phe130. C15 and C17 interacted with His231 (CB and CG). C17 also interacted with Leu202, while C20 interacted with Asn112 (CG), Leu202 and Phe130.

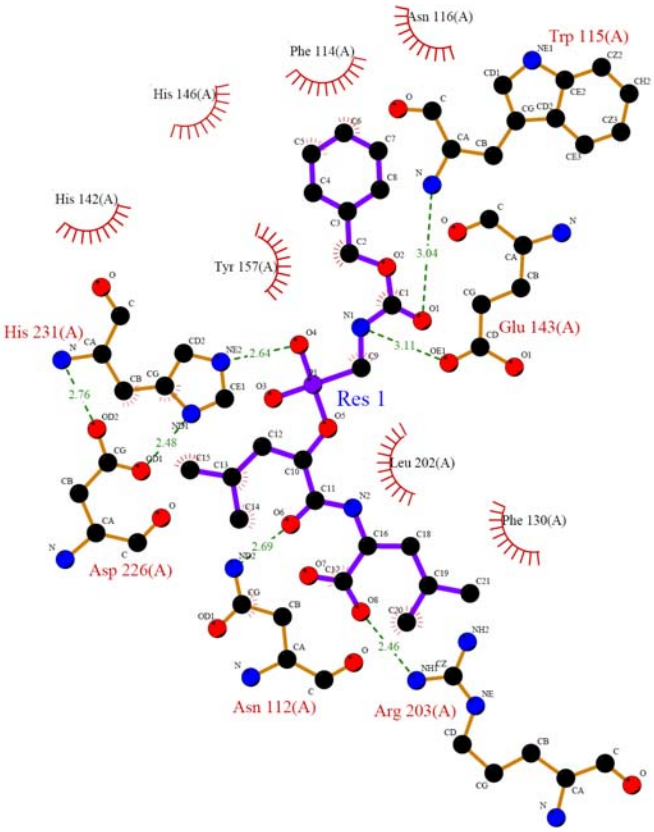


Figure 4.25: Molecular interactions of compound **21** with the active site of thermolysin (Target PDB code 1gxw).

Detailed interactions of compound **22** are shown in Figure 4.26.1 and 4.26.2. Atom C1 exhibited hydrophobic interaction with Phe114, C3 with His231 (CE1), C4 interacted with Tyr157 and His146 (CD2), C5 with Tyr157 and Phe114, C6 interacted with Phe114 and His146 (CD2). C9 interacted with Leu202 and Phe130. C10 interacted with Val139, Leu133, Phe130, Leu202 and Asn112 (CG). C15 and C16 interacted with Leu202 and Phe130. C17, C18, C19 and C20 interacted with Leu202 and Phe130, while C21 interacted with Asn112 (CG).

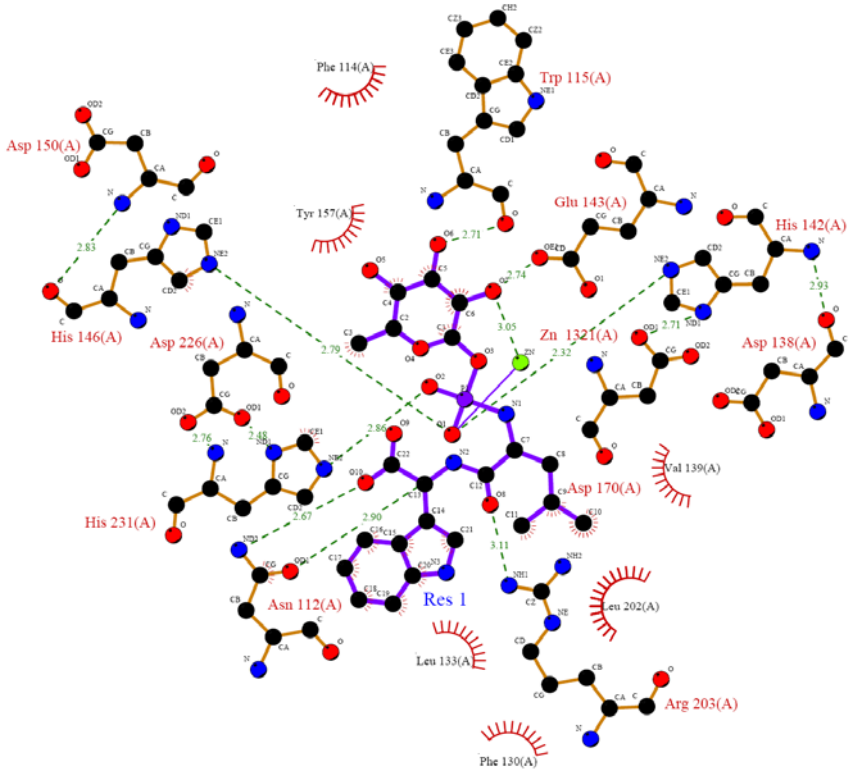


Figure 4.26.1: The molecular interactions of compound **22** with the active site of thermolysin (Target PDB code 1gxw).

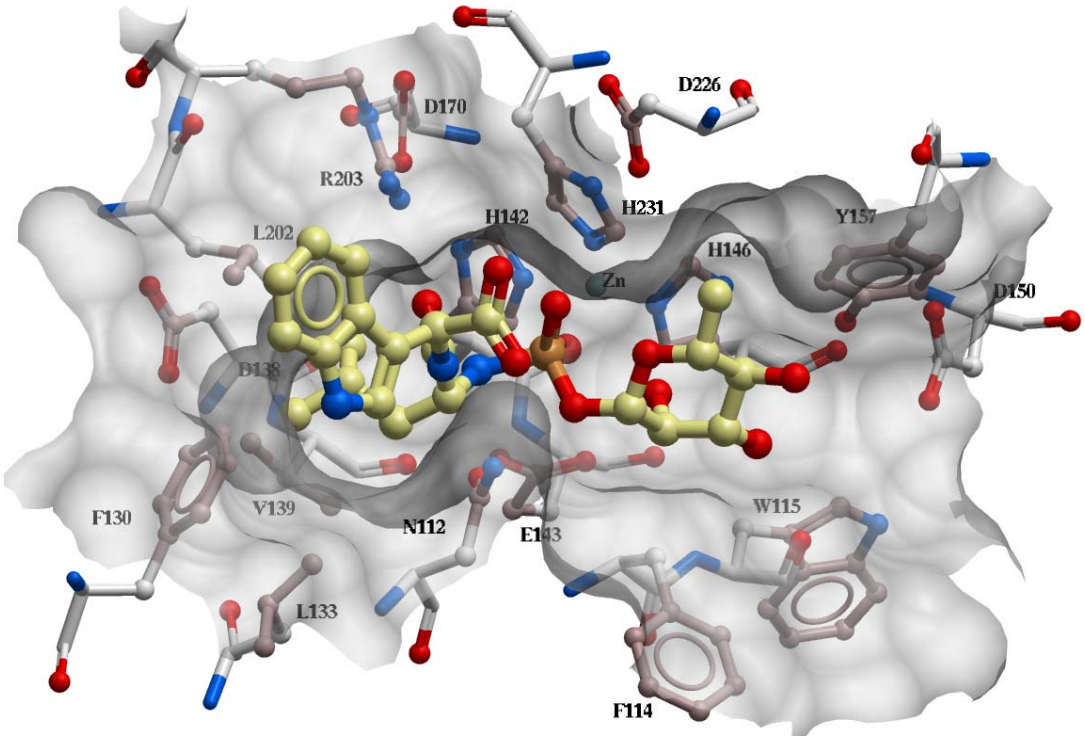


Figure 4.26.2: Corresponding (to Figure 4.26.1) stereoscopic view of compound 22 at the active site binding pocket of thermolysin.

Detailed interactions of compound **23** are shown in Figure 4.27. For this compound atom C1 exhibited hydrophobic interactions with His146, His142 and Phe114. These interactions are very important as residues His146 and His142 are playing vital role as active site residue of thermolysin. C4, C5 and C6 interacted with Phe114 and His146. C8 and C9 interacted with Tyr157 and His146. C16 and C17 interacted with His231 (CG and CD2). C12 interacted with His142. C24 and C25 interacted with Phe130. C24 also interacted with Asn111 (CB), and C26 interacted with Asn112 (CG).

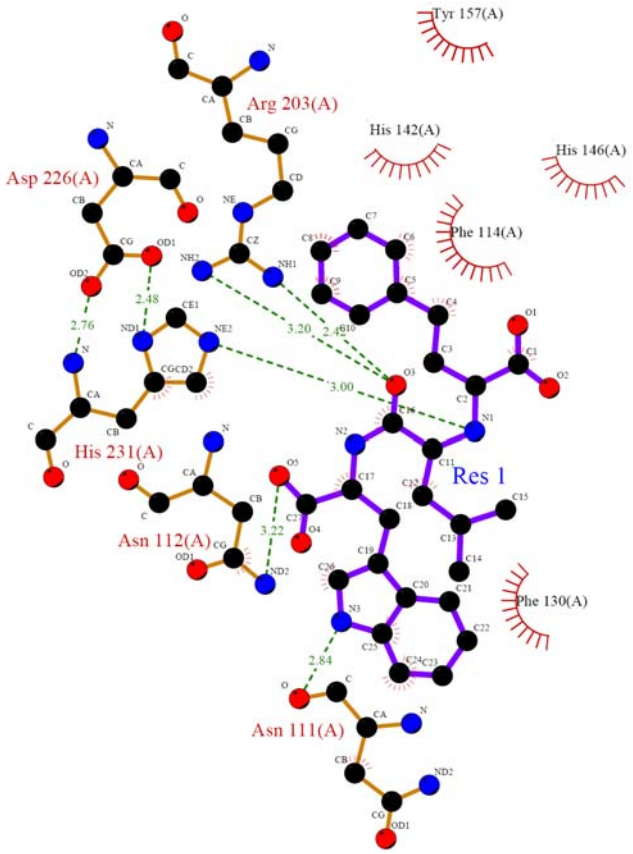


Figure 4.27: The molecular interactions of compound **23** with the active site of thermolysin (Target PDB code 1gxw).

For compound **24**, all detailed interactions are shown in Figure 4.28. Atoms C5 and C19 had hydrophobic interaction with Asn112, and C19 also interacted with His146 (CD2). C8 interacted with His142, while C9, C10, C11 and C12 interacted with Leu202; C9, C11 and C12 also interacted with His231 (CB, CG and CD2), and C18 interacted with Phe114.

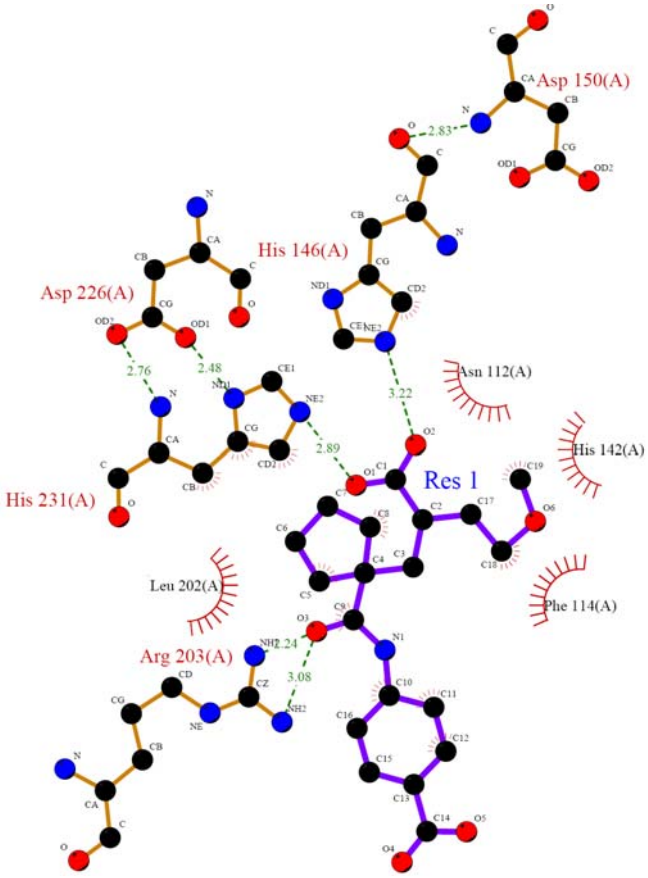


Figure 4.28: The molecular interactions of compound **24** with the active site of thermolysin (Target PDB code 1gxw).

For compound **25**, all detailed interactions are shown in Figure 4.29.1 and 4.29.2. For this compound, atoms C3 and C11 had hydrophobic interactions with Asn112 (CG). C4, C5, C6, C7, C8 and C9 interacted all with Phe114. C13 and C14 interacted with Leu202; C14 also interacted with Phe130 and Asn112 (CG). C16 showed interaction with Leu133, C18 and 19 with His231 (CE1), while C27 and C28 interacted with Leu202 and Phe130.

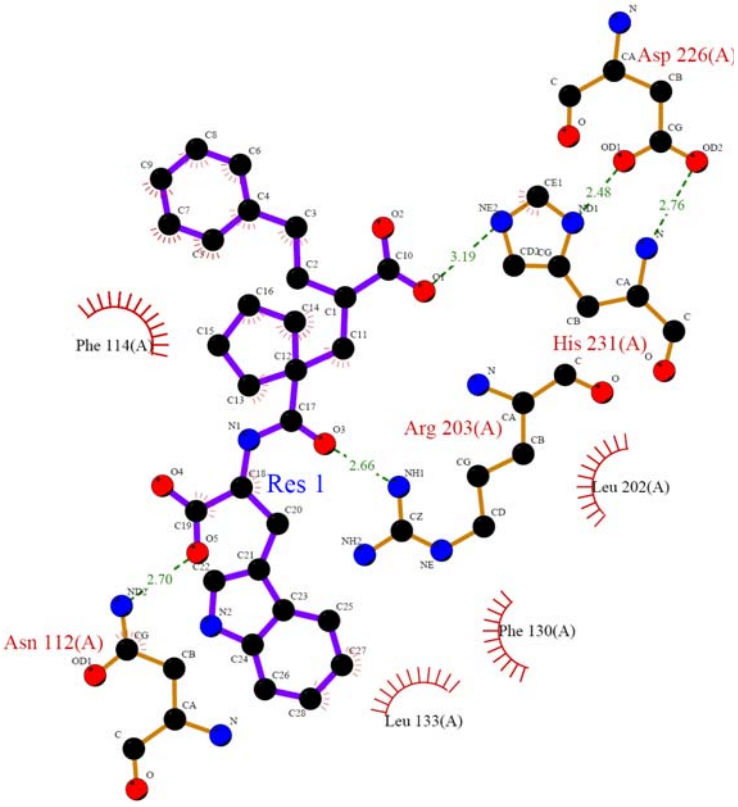


Figure 4.29: Molecular interactions of compound **25** with the active site of thermolysin (Target PDB code 1gxw).

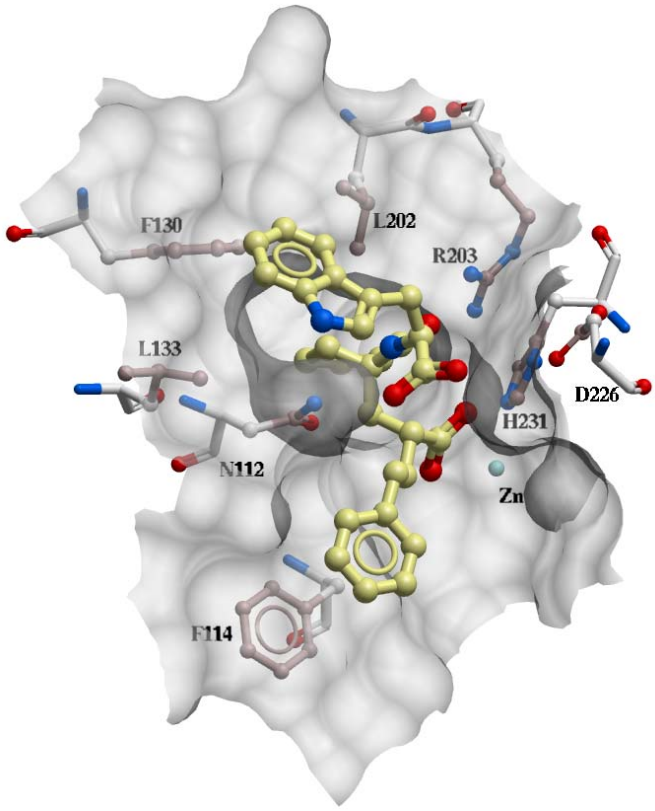


Figure 4.29.2: Corresponding (to Figure 4.29.1) stereoscopic view of compound 25 at the active site binding pocket of thermolysin.

Table 4.9: Docking Interactions of 25 thermolysin inhibitors from literature based on the second docking.

Comp.	Interactions with Zn Distance in Å	H-Bond (Atom no, Length in Å)
1	Zn-O4 (2.57) Zn-O5 (2.28)	N1-Asn112(OD1 3.24)
		O3-Arg203(NH1 2.79)
		O4-Glu143(OE1 2.57)
		O5-His146(NE2 2.97)
		O5-Glu166(OE2 2.64)
		O6-Trp115(N 3.28)
2	Zn-O2 (2.27)	O2-His146(NE2 3.32)
		O4-Arg203(NH1 2.57)
		N2-Asn112(OD1 3.22)
3	Zn-S1 (2.69)	S1-His231(NE2 3.24)
		O2-Asn112(ND2 2.50)
		O3-Arg203(NH1 2.29)
		O3-Arg203(NH2 3.18)
		N2-Asn112(OD1 3.31)
4	Zn-S1 (2.65)	O1-Arg203(NH1 2.46)
		O1-Arg203(NH2 3.32)
		N1-Asn112(OD1 3.08)
		O2-Asn112(ND2 2.71)
5		N1-Ala113(O 3.11)
		O2-Arg203(NH1 3.04)
		O2-Arg203(NH2 3.08)
		O3-Asn111(O 3.03)
		O5-Asn112(ND2 2.57)
6	Zn-O1 (2.05)	N1-Asn112(OD1 3.11)
		N1-Ala113(O 3.04)
		N2-Asn112(OD1 3.05)
		O2-Arg203(NH1 3.00)
		O2-Arg203(NH2 3.23)
		O4-Asn112(ND2 2.70)
7	Zn-S1 (2.35) Zn-O4 (2.40)	S1-His146(NE2 3.18)
		N2-Asn112(OD1 2.96)
		O4-His142(NE2 3.06)
8	Zn-O3 (2.10)	N1-Asn112(OD1 3.33)
		N2-Asn112(OD1 3.04)
		O1-Asn112(ND2 2.54)
		O2-Arg203(NH1 2.74)
		O2-Arg203(NH2 3.14)

Table 4.9 *Continued*

9	Zn-S1 (2.46)	S1-His231(NE2 3.24)
		O2-Asn112(ND2 2.36)
		O3-Arg203(NH1 2.47)
		O3-Arg203(NH2 2.87)
10	Zn-S1 (2.51)	N1-Asn112(OD1 3.29)
		O1-Asn112(ND2 2.53)
		O2-Arg203(NH1 2.66)
		O2-Arg203(NH2 3.30)
11	Zn-S1(2.38)	S1-His142(NE2 3.34)
		N1-Asn112(OD1 3.11)
		O1-Asn112(ND2 2.53)
		O2-Arg203(NH1 2.69)
		O2-Arg203(NH2 3.22)
12	Zn-O1(2.46) Zn-O2(2.21)	N1-Glu143(OE1 3.28)
		N2-Asn112(OD1 3.08)
		O1-His231(NE2 3.31)
		O2-His146(NE2 3.09)
		O2-His142(NE2 3.25)
		O3-Arg203(NH1 2.67)
		O5-Asn112(ND2 2.48)
13	Zn-O1(2.13)	N2-Asn112(OD1 3.12)
		O1-His231(NE2 3.06)
		O3-Arg203(NH1 2.56)
		O5-Asn112(ND2 2.57)
14	Zn-O)1(2.10)	O1-His146(NE2 3.04)
		N2-Asn112(OD1 3.10)
		O5-Arg203(NH1 2.69)
		O7-Asn112(ND2 2.64)
15	Zn-O1(2.07)	O1-His146(NE2 2.99)
		N2-Asn112(OD1 3.13)
		O5-His231(NE2 3.18)
		O6-Arg203(NH1 2.59)
		O8-Asn112(ND2 2.63)
16	Zn-O1(2.04)	O1-His146(NE2 2.93)
		N2-His231(NE2 3.10)
		N3-Asn112(OD1 3.06)
		O5-Arg203(NH1 2.65)
		O7-Asn112(ND2 2.53)

Table 4.9 ***Continued***

17	Zn-O4(2.51) Zn-O5(2.36)	O2-Arg203(NH1 2.68)
		O3-Arg203(NH1 2.68)
		O3-Arg203(NH2 3.05)
		O4-His231(NE2 2.86)
		O5-His142(NE2 2.96)
		N2-Ala113(O 3.19)
		N2-Asn112(OD1 2.99)
18	Zn-O1(2.71) Zn-O2(2.52) Zn-O3(2.57)	O1-His146(NE2 2.77)
		O2-Glu143(OE1 2.61)
		O3-Glu166(OE1 3.22)
		O3-His231(NE2 2.58)
		O4-Arg203(NH1 2.61)
		O4-Arg203(NH2 3.14)
		O5-Arg203(NH1 2.76)
19	Zn-O4(2.65) Zn-O5(2.77)	O1-Asn112(ND2 2.65)
		O3-Arg203(NH1 2.65)
		O3-Arg203(NH2 3.34)
		O4-His231(NE2 2.96)
		N2-His231(NE2 2.98)
20	Zn-O3(2.52) Zn-O4(2.50)	N2-Ala113(O 3.29)
		N2-Asn112(OD1 3.09)
		N3-Asn112(OD1 3.33)
		O4-His231(NE2 2.71)
		O5-Arg203(NH1 2.58)
		O5-Arg203(NH2 2.96)
		O7-Asn112(ND2 2.42)
21	Zn-O3(2.51) Zn-O4(2.59)	N1-Glu143(OE1 3.11)
		O1-Trp115(N 3.04)
		O4-His231(NE2 2.64)
		O6-Asn112(ND2 2.69)
		O8-Arg203(NH1 2.46)
22	Zn-O2(2.51) Zn-O7(3.05)	O1-His142(NE2 2.32)
		O1-His146(NE2 2.79)
		O2-His231(NE2 2.86)
		N2-Asn112(OD1 2.90)
		O6-Trp157(O 2.71)
		O7-Glu143(OE1 2.74)
		O8-Arg203(NH1 3.11)
		O10-Asn112(ND2 2.67)

Table 4.9 ***Continued***

23	Zn-O1(2.32) Zn-O2(2.27)	N1-His231(NE2 3.00)
		N3-Asn111(O 2.84)
		O3-Arg203(NH1 2.42)
		O3-Arg203(NH2 3.20)
		O5-Asn112(ND2 3.22)
24	Zn-O1(2.09) Zn-O2(2.29)	O1-His231(NE2 2.89)
		O2-His146(NE2 3.22)
		O3-Arg203(NH1 2.24)
		O3-Arg203(NH2 3.08)
25	Zn-O1(2.37) Zn-O2(2.23)	O1-His231(NE2 3.19)
		O3-Arg203(NH1 2.66)
		O5-Asn112(ND2 2.70)

4.2 MS Compounds

4.2.1 Experimental studies of MS compounds

Thermolysin activity was assayed with the chromophoric substrate FAGLA by monitoring the decrease in absorbance at 346 nm due to the hydrolysis of the Gly-Leu bond [149]. The inhibition characteristics of MS compounds for thermolysin using FAGLA as the substrate, is given in Table 4.10. Two (MS9 and MS12) out of 13 diterpenoid alkaloids were showed active, having IC_{50} values of 0.00991 and17.674 mM, respectively. Eight compounds were completely inactive, and three compounds showed very low activity and the IC_{50} value could not be detected. The inhibition of MS12 appears far poorer than MS9. M9 and M12 were subjected to kinetic and other experimental as well as molecular modeling studies in order to predict their mode of binding at the active site of thermolysin. The kinetic studies are described in a manuscript from our group [150].

Table 4.10: Experimental result of MS compounds.

Comp.	IC_{50} (in mM)
MS1	Low [a]
MS2	IN [b]
MS3	IN [b]
MS4	IN [b]
MS6	IN [b]
MS8	IN [b]
MS9	9.91×10^{-3}
MS10	IN [b]
MS11	IN [b]
MS12	17.674
MS13	Low [a]
MS14	Low [a]
MS15	IN [b]

NOTES: [a] due to the low activities the IC_{50} values were not possible to calculate; [b] completely inactive.

4.2.2 Docking of MS compounds

Before experimental assay, all the ligands were docked to target protein using regular docking method. Due to lack of crystal structure as a guideline to evaluate the poses, only the highest scored conformation (best docking energy) was selected. Calculated ICM binding energy indicated varying degree of binding for all 13 ligands (Table 4.11), while experimental result implied only 2 real binders (MS09 and MS12). The modified docking method used for 25 ligands was applied for re-docking of compounds MS09 and MS12 (Table 4.12 and 4.13).

Table 4.11: First time docking result of all 15 MS compounds.

Comp.	Cal.ΔG	Docking energy (in kcal/mol)
MS1	-3.2	-94.3
MS2	-7.0	-118.5
MS3	-6.2	-100.0
MS4	-5.9	-78.9
MS6	-5.8	-99.2
MS8	-3.9	-79.5
MS9	-3.6	-92.6
MS10	-3.3	-105.1
MS11	-5.2	-96.6
MS12	-5.1	-82.5
MS13	-7.4	-79.5
MS14	-5.0	-57.9
MS15	-4.5	-93.2

Table 4.12: Docking energy value (kcal/mol) of top poses (the 1st pose in each round). Reference docking run stands for the best ranked pose in each of 5 runs with the docking energy closest to the average value.

Complex	Rotatable bond	Dock 1	Dock 2	Dock 3	Dock 4	Dock 5	Reference docking run	Average energy
MS9	**15**	-108.3	-105.6	-107.8	-108.6	-108.9	1	-107.8
MS12	**6**	-72.6	-72.6	-73.5	-71.2	-73.0	2	-72.6

Table 4.13: sRMSD values of the best pose from each run relative to the top scored reference docking run.

Complex	Rotatable bond	Dock 1	Dock 2	Dock 3	Dock 4	Dock 5	Reference ligand No.	Average sRMSD
MS9	**15**	/	0.32	0.07	0.42	0.3	1	0.28
MS12	**6**	4.53	/	2.98	4.46	0.67	2	3.16

For compound **MS9** (Figure 4.30.1 and 4.30.2), the O4 atom formed a hydrogen bond with the active site metal cation Zn at an atomic distance of 2.57 Å. The same atom also formed hydrogen bonding with the NE2 atom of His146, which is an important active site amino acid residue. The same compounds also had several hydrophobic interactions with different atoms. Atoms C7, C9, C10, and C12 showed hydrophobic interactions with Asn112

and Phe114. C14 interacted with Tyr157. C23, C24, C25, C26 and C35 had hydrophobic interactions with His231, another important active site residue. Atom C35 interacted with Val230, while C36 exhibited several hydrophobic interactions with Asn112, Phe114 and Asn116.

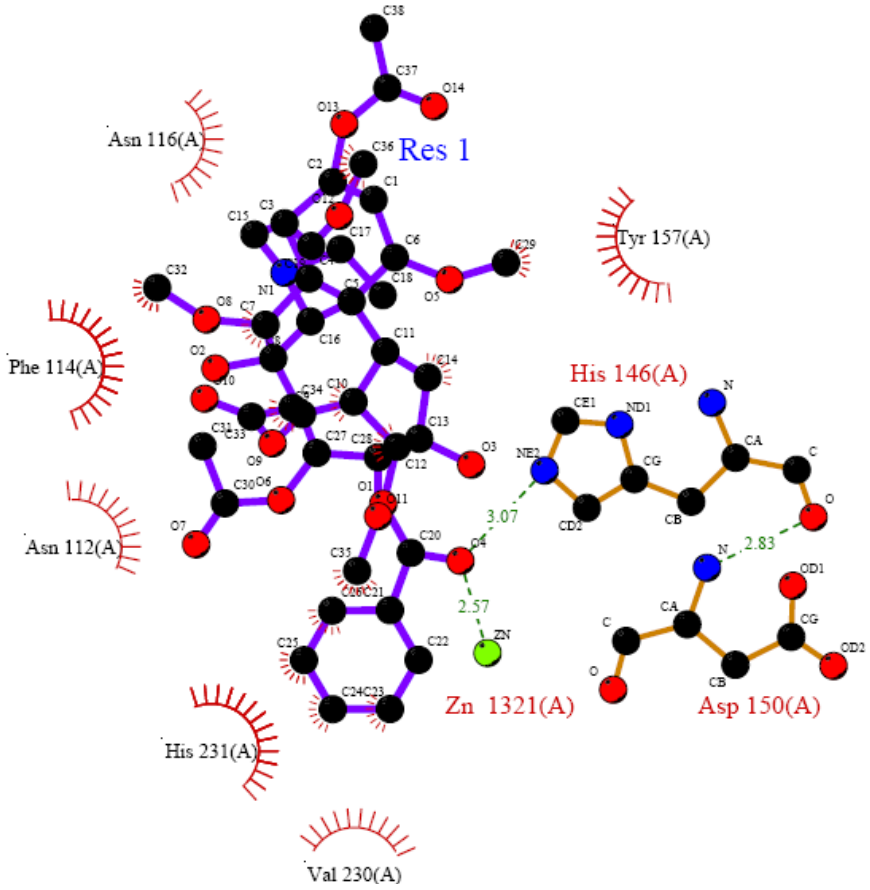


Figure 4.30.1: Molecular interactions of the **MS9** within the active site of thermolysin (Target PDB code 1gxw).

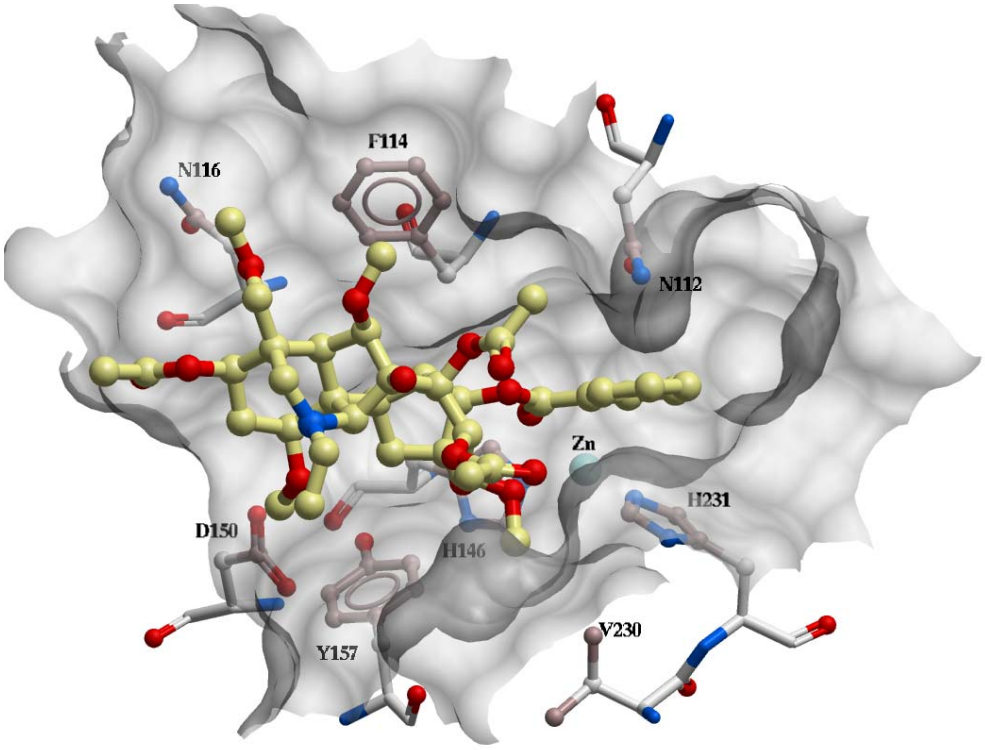


Figure 4.30.2: Corresponding (to Figure 4.30.1) stereoscopic view of compound MS9 at the active site binding pocket (grey transparent mode) of thermolysin.

For compound **MS12** (Figure 4.31.1 and 4.31.2), the two oxygen atoms O2 and O4 formed hydrogen bonds with Zn at atomic distance of 2.98 and 2.42 Å, respectively. The hydrogen bond between O4 and Zn was similar but shorter than that of O4 of compound MS12. With zinc, O2 also formed hydrogen bonds with the NE2 and OE2 atoms of His146 and Glu166, respectively. Their interatomic lengths were calculated using LigPlot and found to be 3.1 and 2.63 Å, respectively. Atoms C7, C8, C9, C10 and C17 showed hydrophobic interactions with Phe114; and atoms C11, 12, C13 and C14 interacted with His231.

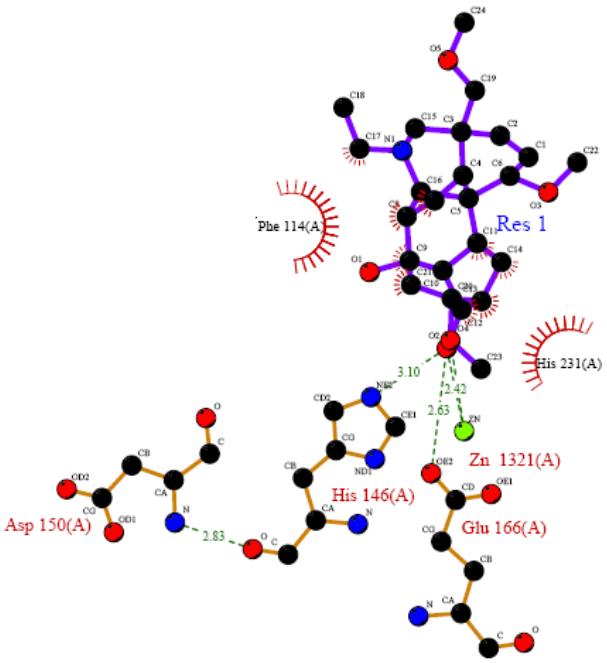


Figure 4.31.1: Molecular interactions of the **MS12** within the active site of thermolysin (Target PDB code 1gxw). Hydrogen bonding of Zn-O2 and Zn-O4 were overlapped in this figure.

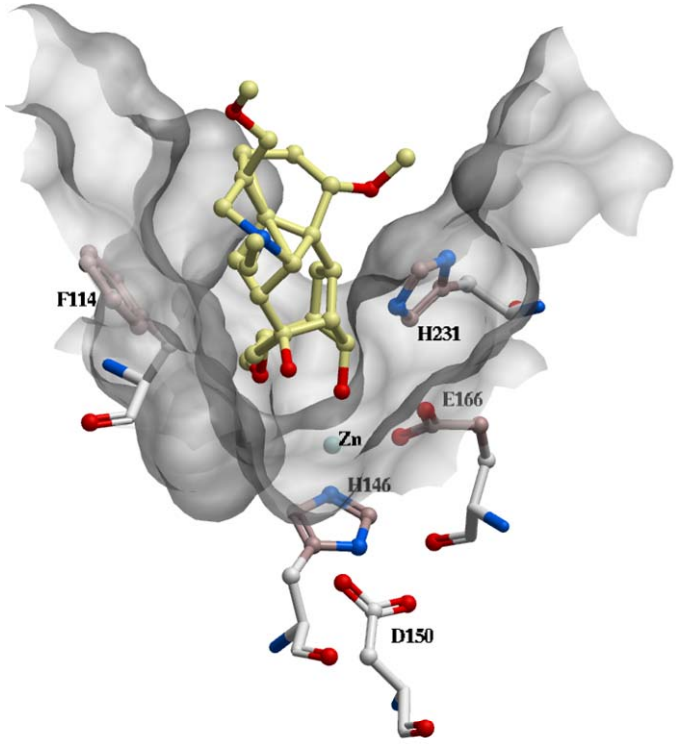


Figure 4.31.2: Corresponding (to Figure 4.31.1) stereoscopic view of compound MS12 at the active site binding pocket (grey transparent mode) of thermolysin.

Table 4.14: Docking Interactions of MS compounds.

Comp.	Docking Energy (Kcal/mol)	Interactions with Zn (Distance, in Å)	H-Bond (Atom no, Length, in Å)
MS9	-107.8	Zn-O4 (2.57)	O4-His146(NE2 3.07)
MS12	-72.61	Zn-O2 (2.98)	O2-His146(NE2 3.10)
		Zn-O4 (2.42)	O2-Glu166(OE2 2.63)

4.3 R Compounds

4.3.1 Experimental studies of R compounds

The IC_{50} values of the R compounds for thermolysin using FAGLA as the substrate are given in Table 4.15. The studies indicated that 12 out of 39 compounds were active. Four compounds were completely inactive; 22 compounds showed low activities. The binding of RS17, RSH57 and RSH78 (f) appeared stronger than the rest of the active compounds. Enzyme kinetic studies were described in our manuscript [151].

Table 4.15: Experimental activity profiles (IC_{50}) of R compounds.

Comp.	IC_{50} (in mM)	Comp.	IC_{50} (in mM)
RK2-16	IN [b]	**RSH30-d**	Low [a]
RSH11	0,648	**RS 25**	Low [a]
RS10	IN [b]	**RSH-1**	Low [a]
RSH22(b)	IN [b]	**RSH 42**	Low [a]
RS17	1.1461×10^{-5}	**RK1-4**	Low [a]
RK10(2)	Low [a]	**RSH77(c)**	Low [a]
RSH79(a)	Low [a]	**RSH28(b)**	Low [a]
RSH78(f)	1.25×10^{-3}	**RSH 16**	Low [a]
J 11	Low [a]	**RS 14**	Low [a]
RSH 10	4.002	**RSH 41**	12.743
RS 12	Low [a]	**RSH 23**	Low [a]
RSH 57	2.4774×10^{-4}	**RSH78(g)**	Low [a]
RS 7	Low [a]	**RSH39(d)**	Low [a]
RSH 66	Low [a]	**RS-1**	IN [b]
RSH30d-2	Low [a]	**J1**	122.637

Table 4.15 *Continued*

RK2-10	Low [a]	**RM19**	0.037862
RS20(R)	Low [a]	**RS3**	0.042035
RSH19(b)	Low [a]	**RSH12**	1.8321
RSH35-A	3.118	**RSH44a**	0.076851

NOTES: [a] due to the low activities the IC_{50} values were not possible to calculate; [b] completely inactive.

4.3.2 Docking studies of R compounds

Only the R-compounds found to bind thermolysin in experimental studies were docked into thermolysin. The ligands were charged and Glu143 was neutral in the docking. A unique Quinazoline group distinguished R compounds from all the other ligands in this study (Figure 3.1). Different sources (PDB database, Cambridge structural database) were searched for potential crystal structure of the experimentally active R compounds, but structures were not found. Models of the compounds were therefore constructed using CSChemOffice. Lack of flexibility (most of them have only 1 to 2 rotatable bonds) made it less important to sort native geometries for the R compounds. In general feed batch docking with 2D structures (in SD file converted from mol file) and all necessary minimizations were performed by default settings. A modified docking approach and evaluation of the docking poses were performed as described previously (paragraph 3.2.2 and 3.2.3). The results are shown in Table 4.16 and 4.17. Except RK1-4, all docked compounds showed negative binding free energies, indicating that a realistic binding scenario was generated (Table 4.18). However, the correlation between binding energies and IC_{50} was not linear, which indicates that the docked conformations are not yet 'perfect'.

Table 4.16: Docking energy values (kcal/mol) of 5 runs for R compounds. Reference docking run stands for the best ranked pose in each of 5 runs with the docking energy closest to the average value.

Compound	Rotatable bond	Dock 1	Dock 2	Dock 3	Dock 4	Dock 5	Reference docking run	Average energy
J1	4	-48.3	-47.5	-47.5	-48.3	-48.3	1	-48.0
RK1_4	2	-49.2	-48.7	-48.1	-49.8	-48.1	2	-48.8
RM19	1	-35.9	-36.2	-35.9	-36.1	-35.9	3	-36.0
RS17	2	-49.3	-49.4	-49.3	-46.6	-46.6	3	-48.2
RS3	4	-65.9	-66.1	-66.1	-65.9	-66.2	3	-66.1
RSH10	3	-61.2	-57.5	-57.6	-57.5	-57.6	3	-58.3
RSH11	3	-61.5	-56.1	-61.6	-61.6	-61.6	4	-60.5
RSH12	2	-61.1	-60.7	-55.0	-54.7	-60.7	3	-58.4
RSH35a	2	-51.8	-51.8	-51.8	-51.8	-51.8	5	-51.8
RSH41	1	-48.5	-49.9	-48.6	-49.3	-47.8	3	-48.8
RSH44a	2	-54.5	-53.4	-53.3	-54.4	-54.4	1	-54.0
RSH57	1	-41.2	-41.2	-41.2	-41.1	-41.2	1	-41.2
RSH78f	1	-51.2	-51.1	-51.1	-51.2	-51.1	3	-51.1

Table 4.17: sRMSD between the best pose of reference docking run and the best poses of the other docking runs.

Compound	Rotatable bond	Dock 1	Dock 2	Dock 3	Dock 4	Dock 5	Reference docking run	Average sRMSD
J1	4	/	0.36	0.34	0.01	0.00	1	0.18
RK1_4	2	1.21	/	0.05	0.19	0.05	2	0.38
RM19	1	0.00	0.01	/	0.02	0.00	3	0.01
RS17	2	0.00	0.00	/	0.26	0.26	3	0.13
RS3	4	0.86	0.86	/	0.02	0.21	3	0.49
RSH10	3	1.42	0.00	/	0.00	0.03	3	0.36
RSH11	3	0.00	0.27	0.00	/	0.00	4	0.07
RSH12	2	1.51	1.51	/	0.03	1.51	3	1.14
RSH35a	2	0.02	0.00	0.00	0.00	/	5	0.00
RSH41	1	0.01	0.19	/	1.38	1.47	3	0.76
RSH44a	2	/	0.30	0.30	0.00	0.00	1	0.15
RSH57	1	/	0.01	0.01	0.00	0.11	1	0.03
RSH78f	1	0.01	0.00	/	0.01	0.00	3	0.01

Table 4.18: Final results of R compounds.

Comp.	Rotatable bond	Docking energy (kcal/mol)	Cal.ΔG	IC_{50} (in mM)
J1	4	-48.3	-1.8	122.637
RK1-4	2	-48.7	1.0	low activity
RM19	1	-36.2	-1.5	0.037862
RS17	2	-49.3	-1.0	0.00001146
RS3	4	-66.2	-3.9	0.042035
RSH10	3	-57.6	-3.5	4.002
RSH11	3	-61.6	-4.1	0.648
RSH12	2	-55.0	-2.4	1.8321
RSH35a	2	-51.8	-3.5	3.118
RSH41	1	-48.6	-4.8	12.743
RSH44a	2	-54.5	-3.7	0.076851
RSH57	1	-41.2	-9.7	0.00024774
RSH78f	1	-51.1	-2.8	0.00125

The different interactions (docking energy, inter atomic interactions with Zn, hydrogen bondings and hydrophobic interaction) of thermolysin inhibitors are comparatively shown in Table 4.19. The hydrophobic interactions depicted on 2D LigPlot figures were confirmed by visually analysing the 3D structure complexes in ICM environment.

All the detailed interactions of compound **J1** are shown in Figure 4.32. The atom C2 showed hydrophobic interaction with His142 (CD2). C5 interacted with Asn112. C7 and C8 interacted with Leu202. C9 interacted with Asn112 and His231 (CE1). C10 interacted with Asn112. C13, C14 and C17 interacted with His231 (CE1). C16 interacted with Phe114, while C18 interacted with Tyr157.

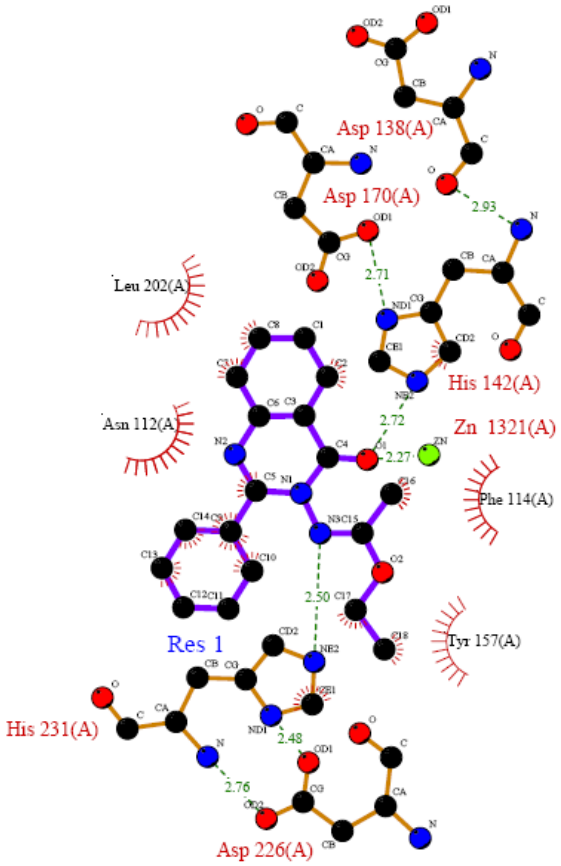


Figure 4.32: Schematic illustration of the interactions of **J1** at the active site of thermolysin (Target PDB code 1gxw). The schematic 2D interaction plot between thermolysin active site and the **J1** was produced using the program LigPlot [141].

For compound **RK1-4**, all detailed interactions are shown in Figure 4.33. Atom C1 had hydrophobic interaction with Leu202, while atom C5 and C7 interacted with Asn112 (CG). C6 interacted with His231 (CE1). C8 and C9 interacted with Leu202, Phe130 and Asn112 (CG).

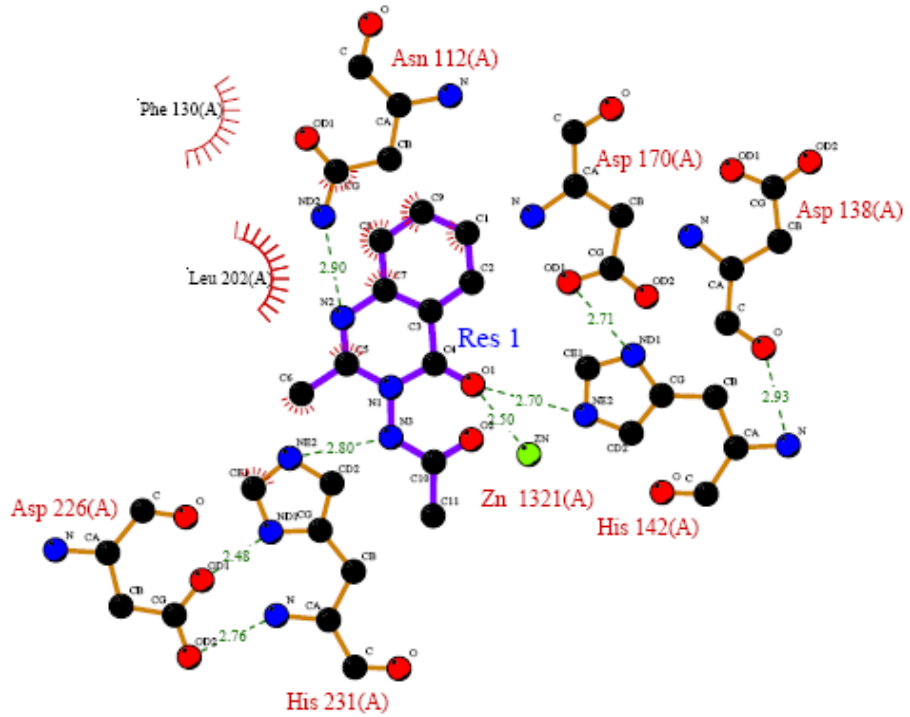


Figure 4.33: Schematic illustration of the interactions of **RK1-4** with the active site of thermolysin (Target PDB code 1gxw).

All detailed interactions of compound **RM19** are shown in Figure 4.34. Atoms C1 and C8 had hydrophobic interaction with Leu202. C6 and C7 interacted with Asn112 (CG), while C9 interacted with His231.

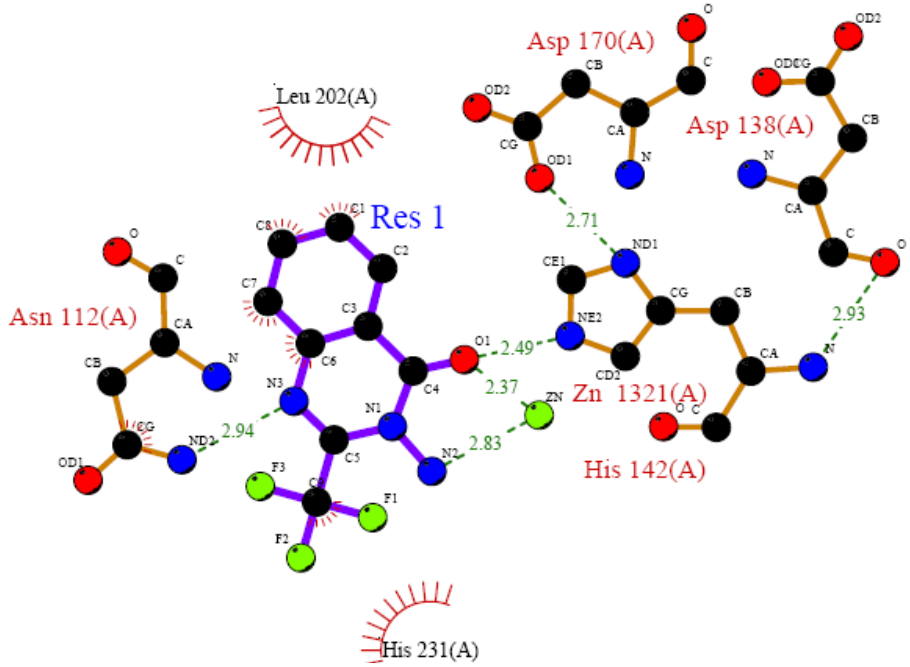


Figure 4.34: Schematic illustration of **RM19** at the active site of thermolysin (Target PDB code 1gxw).

For **RS3**, all the detailed interactions with thermolysin are shown in Figure 4.35. C1 and C8 had hydrophobic interaction with Leu202, while C2 and C3 interacted with His231. C6 and C7 interacted with Asn112 (CG). C16 interacted with Trp115, Tyr157 and His146. C17 and C18 interacted with Trp115 and Tyr157, and C19 interacted with Tyr157. C21 and C22 interacted with Trp115 and Tyr157.

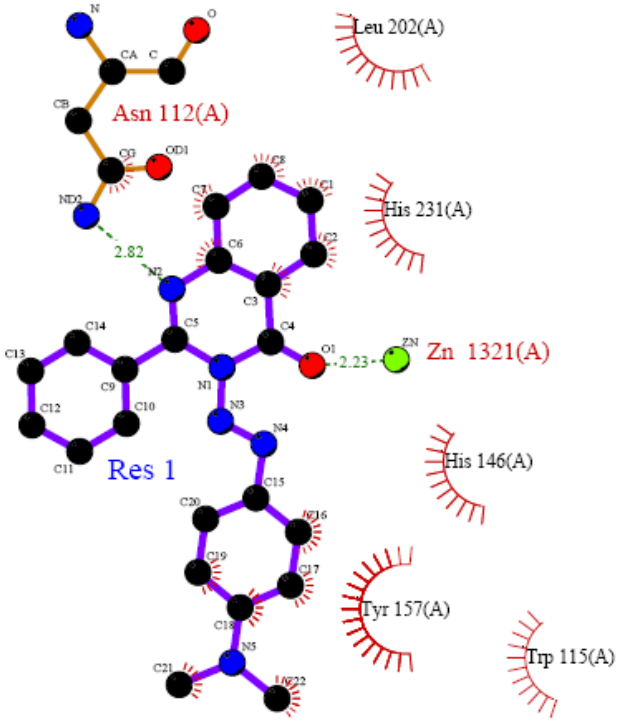


Figure 4.35: Schematic illustration of the interactions of **RS3** at the active site of thermolysin (Target PDB code 1gxw).

All detailed interactions of compound **RS17** are shown in Figure 4.36.1 and 4.36.2. C1 and C2 had hydrophobic interaction with Tyr157 and His146 (CG, CD2). C7 interacted with Asn112, while C9 and C10 interacted with Asn112 and His231. C11 and C12 interacted with Phe114.

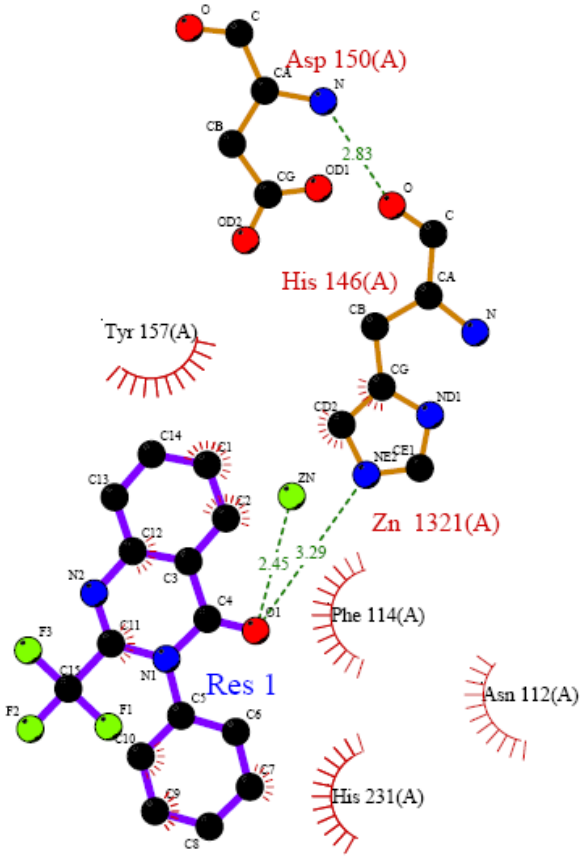


Figure 4.36.1: Schematic illustration of **RS17** at the active site of thermolysin (Target PDB code 1gxw).

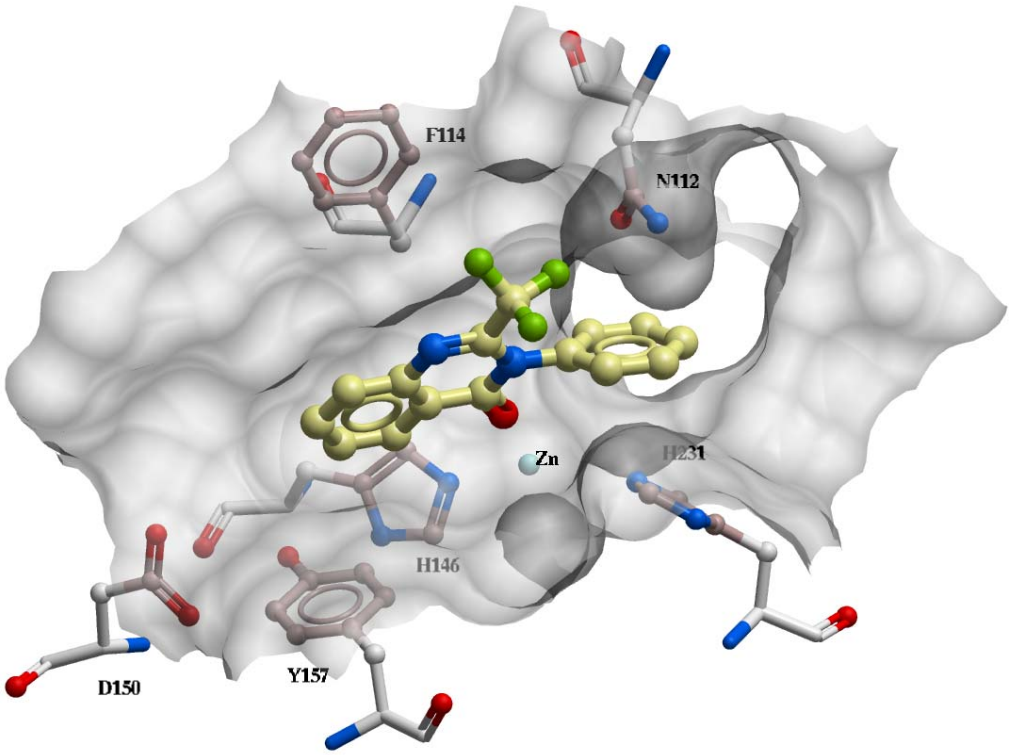


Figure 4.36.2: Corresponding (to Figure 4.36.1) stereoscopic view of compound RS17 at the active site binding pocket (grey transparent mode) of thermolysin.

All detailed interactions of compound **RSH10** are shown in Figure 4.37. Atom C1 had hydrophobic interaction with Val139. C2 interacted with His142 (CG, CD2). C4 interacted with His142 (CD2). C8, C9 and C10 interacted with Tyr157, while C11 interacted with His146. C12, C16, C17 and C18 interacted with Asn112. C14 and C15 interacted with Leu202. C20 and C21 interacted with His231 (CE1).

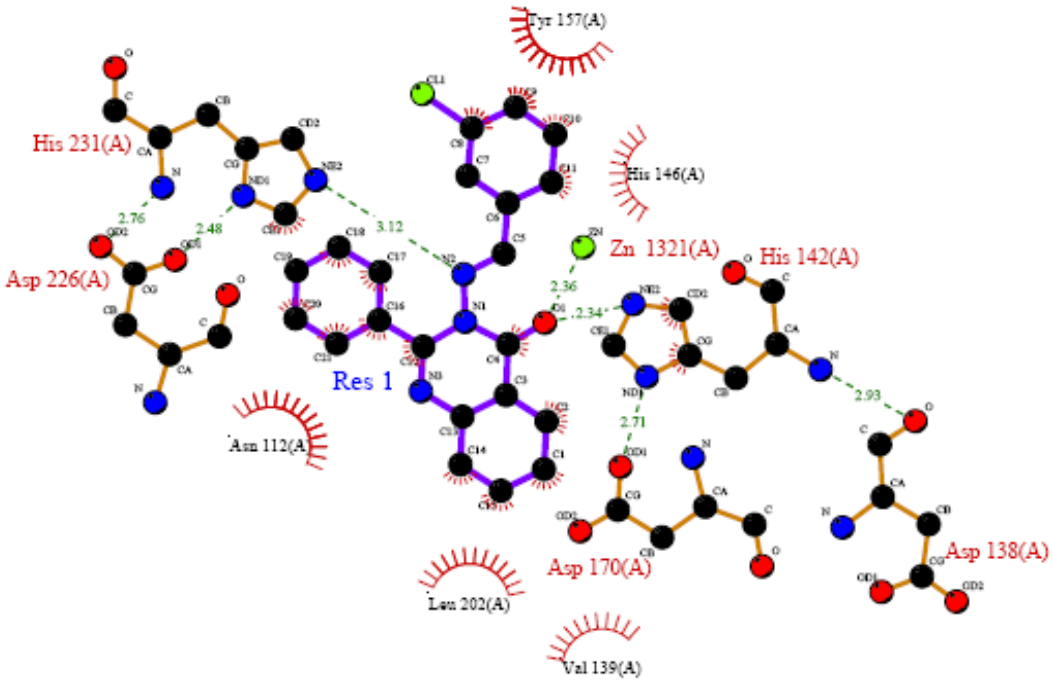


Figure 4.37: Schematic illustration of **RSH10** in the active site of thermolysin (Target PDB code 1gxw).

Figure 4.38 shows the detailed interactions of **RSH11** with thermolysin. Atom C2 had hydrophobic interaction with His142 (CG, CD2), while C4 interacted with His142 (CD2). C5, C9 and C10 interacted with Asn112. C7 and C8 interacted with Leu202. C13 and C14 interacted with His231, and C17 interacted with His146. C18, C19 and C20 interacted with Tyr157.

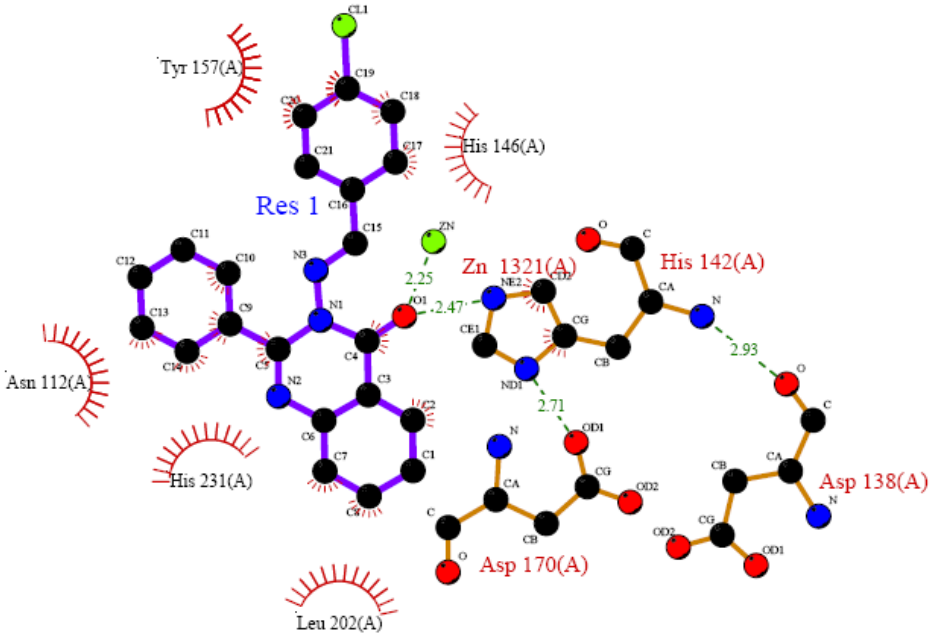


Figure 4.38: Schematic illustration of **RSH11** in the active site of thermolysin (Target PDB code 1gxw).

Detailed interactions for **RSH12** are shown in Figure 4.39. The figure shows that C2 had hydrophobic interaction with His142 (CG, CD2). C4 interacted with His142 (CD2), and C5 interacted with Asn112. C6 interacted with Asn112 and His231, while C8 and C9 interacted with Leu202. C13 and C14 interacted with Tyr157. C15 and C16 interacted with His146 and Tyr157.

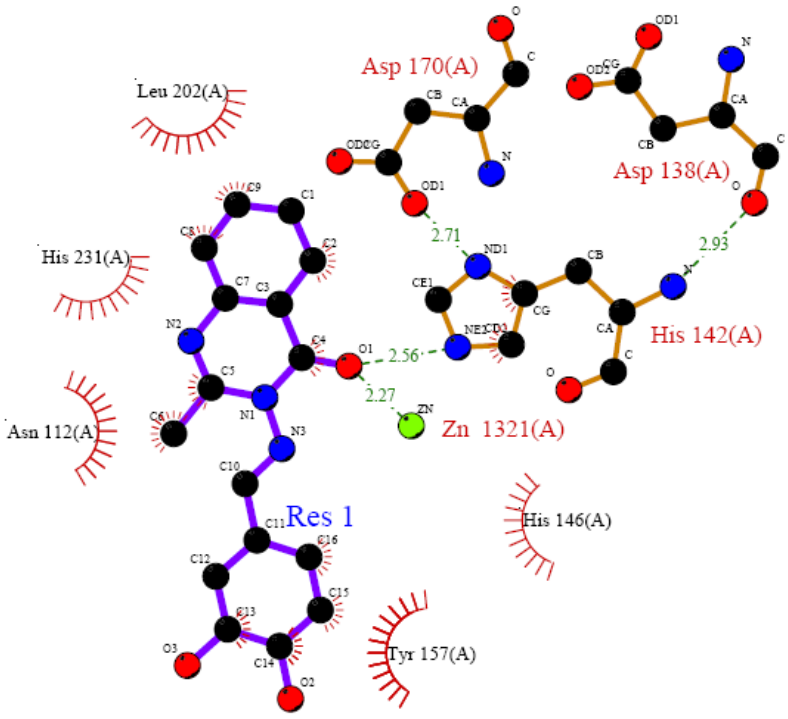


Figure 4.39: Schematic illustration of the **RSH12** in the active site of thermolysin (Target PDB code 1gxw).

Detailed interactions for compound **RSH35a** are shown in Figure 4.40. The figure shows that atom C1 had hydrophobic interaction with His231 and Asn112. C2, C3, C4, C6 and C7 interacted with His231. C8 interacted with Asn112, while C12, C13 and C14 interacted with Tyr157.

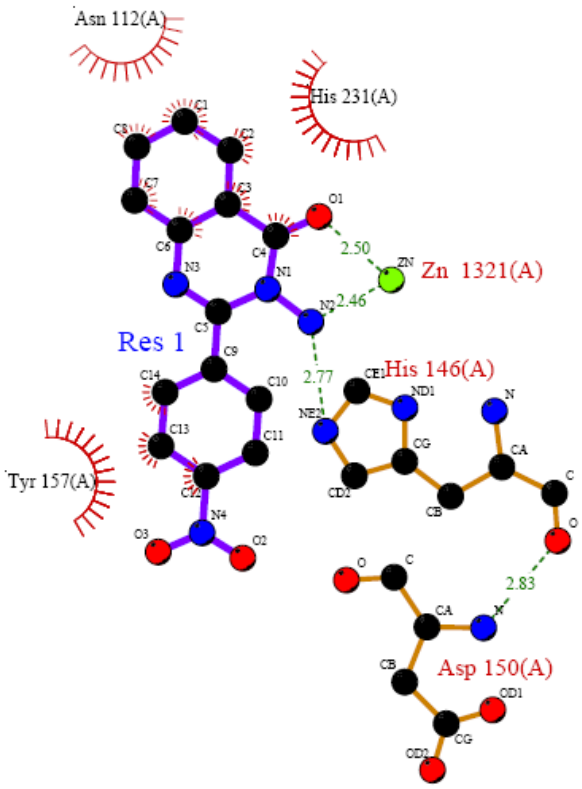


Figure 4.40: Schematic illustration of **RSH35a** at the active site of thermolysin (Target PDB code 1gxw).

For compound **RSH41**, all detailed interactions are shown in Figure 4.41. The atoms C2, C3, C4, C5, C6 and C7 showed hydrophobic interaction with His231. C1 and C8 interacted with Asn112, while C12 and 13 interacted with Tyr157.

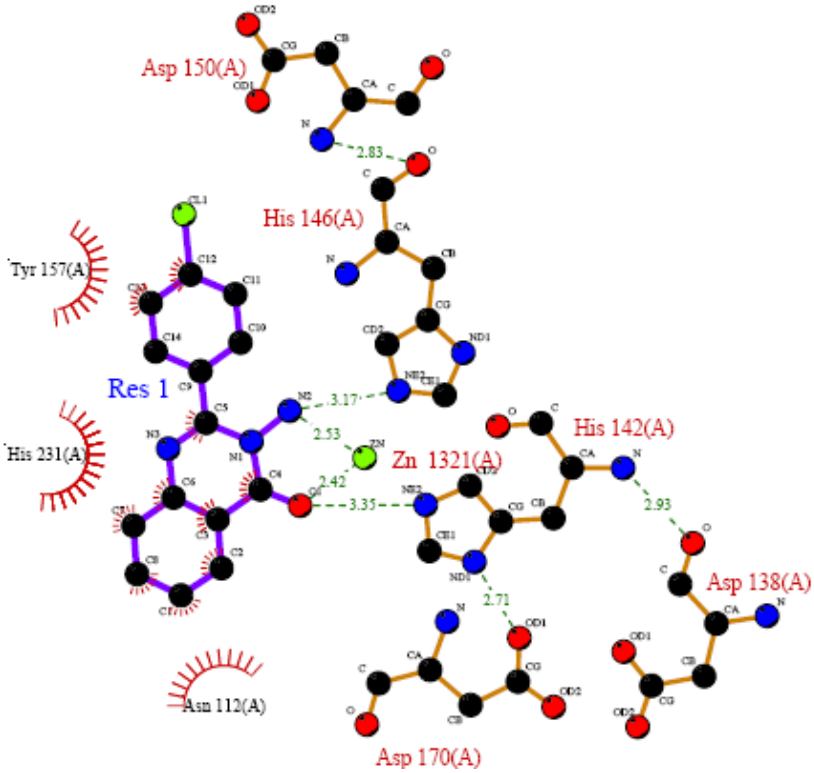


Figure 4.41: Schematic illustration of **RSH41** at the active site of thermolysin (Target PDB code 1gxw).

For compound **RSH44a**, detailed interactions are shown in Figure 4.42. The figure shows that atoms S1, C11 and C18 had hydrophobic interaction with Phe114. C1, C2, C3 and C4 interacted with His231, while C5 interacted with Asn112 (CG). C6 interacted with Asn112 (CG) and His231. C10 interacted with Asn112 (CG) and Phe114. C17 interacted with His146.

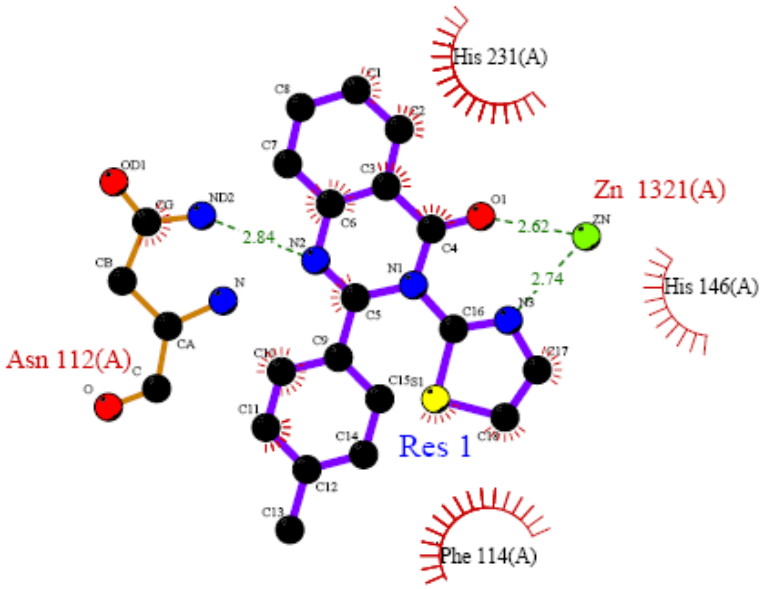


Figure 4.42: Schematic illustration of **RSH44a** at the active site of thermolysin (Target PDB code 1gxw).

All the detailed interactions of compound **RSH57** are shown in Figure 4.43.1 and 4.43.2. C2 had hydrophobic interaction with His142 (CG, CD2 and CE1). C3 interacted with His142 (CG, CD2), and C4 interacted with His142 (CG). C6 interacted with Asn112 (CG) and C7 interacted with His231 (CE1, CD2). C12 interacted with His146, Glu143 and Phe114, while C13 interacted with Tyr157 and His146. Atom O1 formed covalent bonds (purple line in Figure) with active site zinc ion and His142 (NE2).

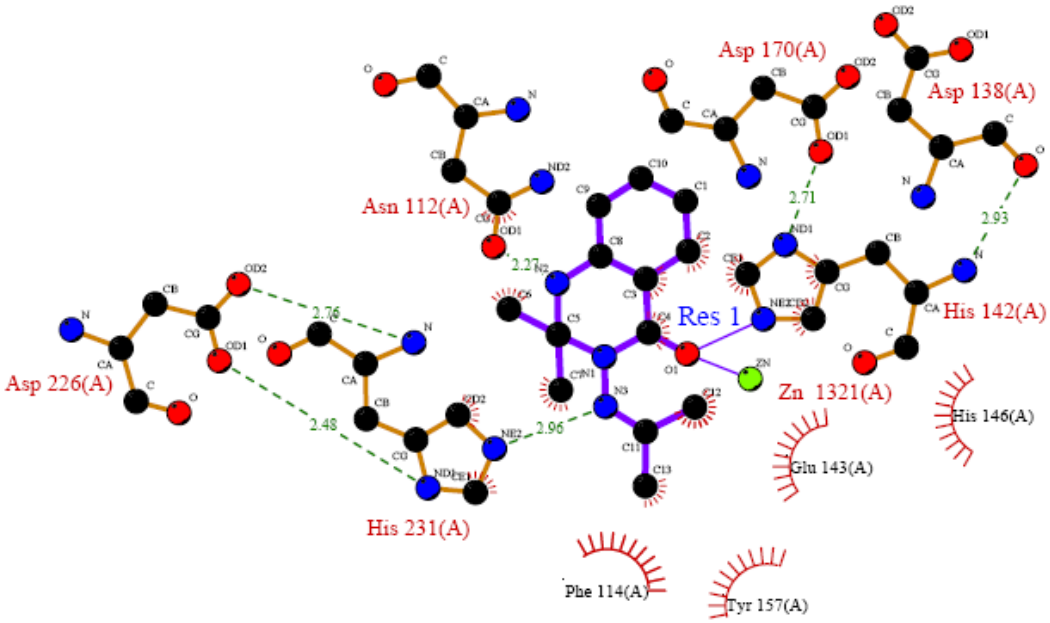


Figure 4.43.1: Schematic illustration of **RSH57** at the active site of thermolysin (Target PDB code 1gxw).

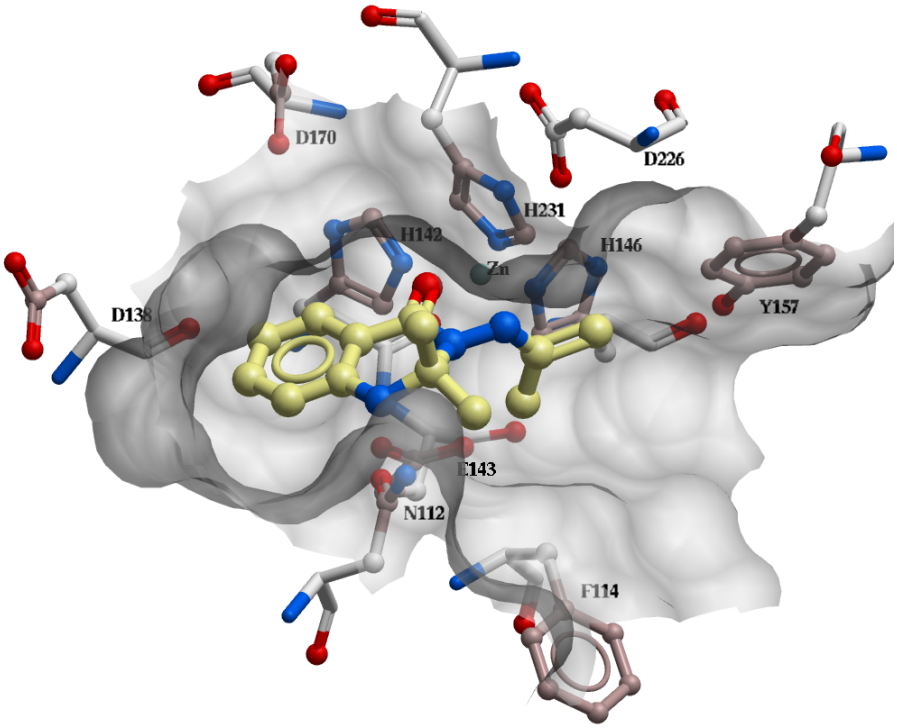


Figure 4.43.2: Corresponding (to Figure 4.43.1) stereoscopic view of compound RSH57 at the active site binding pocket (grey transparent mode) of thermolysin

For compound **RSH78f**, detailed interactions with thermolysin are shown in Figure 4.44.1 and 4.44.2. C1 had hydrophobic interaction with Val139, and C2 interacted with His142 (CD2). C4 interacted with His231, while C6 and C7 interacted with Leu202. C8 interacted with Leu202 and Val139. C10 interacted with His231.

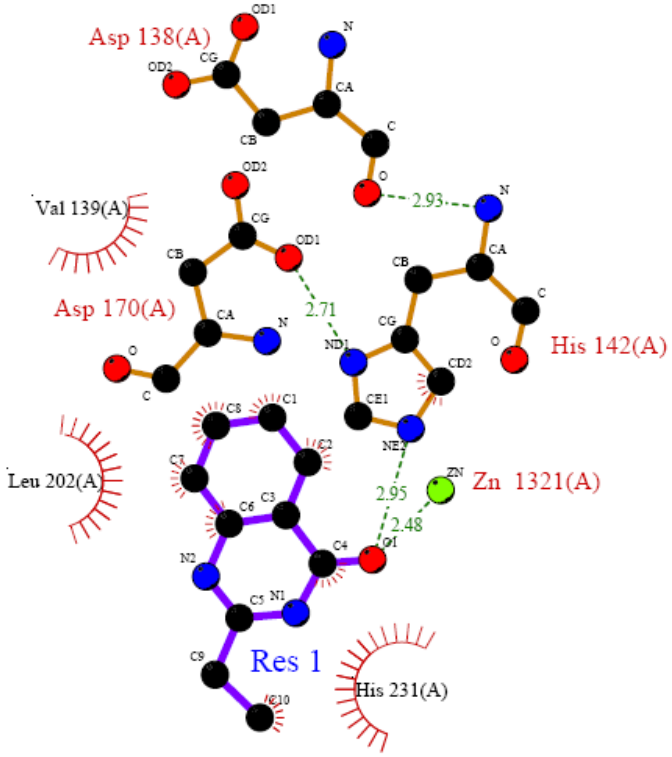


Figure 4.44.1: Schematic illustration of **RSH78f** at the active site of thermolysin (Target PDB code 1gxw).

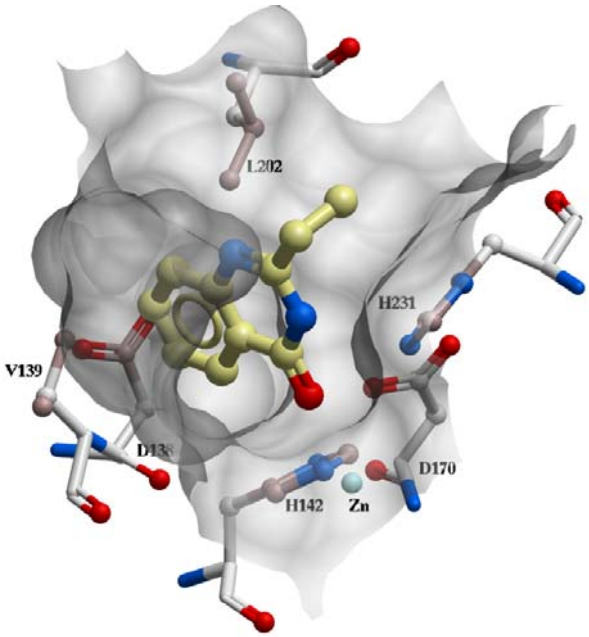


Figure 4.44.2: Corresponding (to Figure 4.44.1) stereoscopic view of compound RSH78f at the active site binding pocket (grey transparent mode) of thermolysin.

Table 4.19: Docking Interactions of R compounds.

Comp.	Interactions with Zn Distance in Å	H-Bond (Atom no, Length, in Å)
J1	Zn-O1(2.27)	O1-His142(NE2 2.72)
		N3-His231(NE2 2.50)
RK1-4	Zn-O1(2.50)	O1-His142(NE2 2.70)
		N2-Asn112(ND2 2.90)
		N3-His231(NE2 2.80)
RM19	Zn-O1(2.37) Zn-N2(2.83)	O1-His142(NE2 2.49)
		N3-Asn112(ND2 2.94)
RS3	Zn-O1(2.23)	N2-Asn112(ND2 2.82)
RS17	Zn-O1(2.45)	O1-His146(NE2 3.29)
RSH10	Zn-O1(2.36)	O1-His142(NE2 2.34)
		N2-His231(NE2 3.12)
RSH11	Zn-O1(2.25)	O1-His142(NE2 2.47)
RSH12	Zn-O1(2.27)	O1-His142(NE2 2.56)
RSH35a	Zn-O1(2.50) Zn-N2(2.46)	N2-His146(NE2 2.77)
RSH41	Zn-O1(2.42) Zn-N2(2.53)	O1-His142(NE2 3.35)
		N2-His146(NE2 3.17)
RSH44a	Zn-O1(2.62) Zn-N3(2.74)	N2-Asn112(ND2 2.84)
RSH57		N3-His231(NE2 2.96)
RSH78f	Zn-O1(2.48)	O1-His142(NE2 2.95)

4.4 Experimental result of M compounds

The IC_{50} values of another series of compound (M compounds) were also performed in this thesis. The inhibition characteristics of M compounds for thermolysin using FAGLA as the substrate, is given in Table 4.18. Two out of seven compounds were active. Compound MTA8 showed extraordinarily strong inhibition with an IC_{50} value of $4{,}411 \times 10^{-11}$ mM. Four compounds showed low active and one compound very low activity. The inhibition of MTA8AC appeared far poorer than that of MTA8. Due to shortage of time, the docking analysis of this group was not performed.

Table 4.20: Experimental result of MS compounds.

Comp.	IC_{50} (in mM)
MEA2	Low [a]
MEA3	Low [a]
MTA8	$4.411x10^{-11}$
MTA8AC	0.0324
SZA1	Low [a]
SZA4	Low [a]
MTA9	Low [a]

NOTES: [a] due to the low activities the IC_{50} values were not possible to calculate; [b] completely inactive.

5 Discussion

A large number of reported thermolysin inhibitors are dipeptides [52, 145, 152-158]. Thermolysin is very much susceptible towards peptide bonds, and therefore dipeptides can easily occupy the catalytic site. A number of small organic molecules are also found to inhibit thermolysin. These molecules are most often derivatives of phosphate [145, 147, 159-161], silicon [136, 162], or sulfate [145]. Several surveys have suggested that, the major and a large number of the thermolysin inhibitors contain some common fragments,that bind the catalytic zinc ion. Nishino and Powers suggested that the oxygen atoms present in these fragments are forming bidentate bonds with the catalytic zinc ion [163-165]. The present docking studies indicate that, other atoms like sulfate (ligand 7) and nitrogen (ligand RM19) also may form part of bidentate bonds. Geometries of potential bidentate arrangments at zinc binding site are depicted in Figure 5.1. Other single interactions like Zn-O (ligand 2, 6, MS09, MS12, almost all RS compounds) and Zn-S (ligand 3, 4) are also common. The multiple zinc coordination geometries play a critical role in the stability of the metalloproteinase inhibitor complex [166].

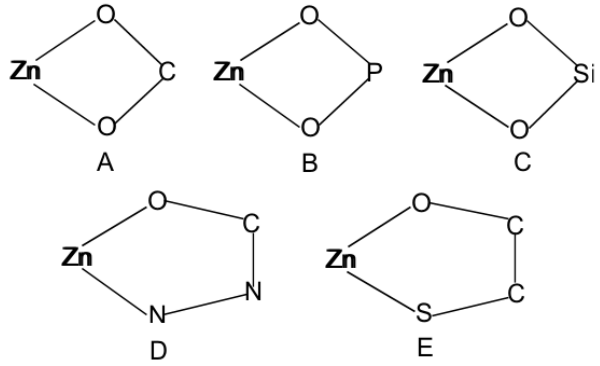


Figure 5.1: Geometry of bidentate arrangements of zinc binding sites. A: Bidentate arrangements of carboxylate groups give rise to a four membered chelate ring (such as ligands 22, 23, 24, 25). B: Bidentate arrangements of phosphate oxygens groups giving a four membered chelate ring (such as ligands 12, 17, 19, 20, 21).. C: Bidentate arrangements of silanediols groups giving a four membered chelate ring (such as ligand 1). D: Bidentate arrangements of OCNN group (ligand RM19) giving a five membered chelate ring. E: Bidentate arrangements of OCCS group (ligand 7) giving a five membered chelate ring.

Extensive studies of flexible ligand docking accuracy using different programs have been reported [101, 122, 123, 144, 167, 168]. Features like, crystallographic resolution of docking target, starting conformation of ligand, number of rotatable bond, conformational space of active site topologies water accessibility, placement of polar hydrogen, correct form of histidine, charged or uncharged form of ligand(s) and amino acid residues within the active

site have been comprehensively mentioned as factors influencing docking accuracy, and need to be taken account during docking. In this section we have endeavored to discover more features that might lead to incorrect predictions. According to Hu [144] who docked 40 zinc-dependent metalloproteinase/ligand complexes with known crystal structures, the actually delocalized charges (+2e) of zinc on the coordinating atoms was unintentionally used as full charge in every docking method, so that the total energy of zinc-ligand binding was overestimated due to overestimated electrostatic interactions. Not having restraints between the zinc ion and the inhibitor atom replacing the water molecule may be another problem. Adekoya mentioned in one of his studies that the inhibitor atom coordinating the zinc ion drifts away from the active site in the course of molecular dynamic simulations when restrains are not used [45].

Table 5.1: Docking successfulness measured by RMSD from the X-ray crystallographic structure of the complex.

Literatures and docking program		**RMSD from corresponding X-ray structure**		**Partition of metal complexes**		**RMSD from corresponding metal complexes (X-ray)**	
	Program	**<2.0 Å**	**<1.0 Å**	**Metal**	**Zinc**	**<2.0 Å**	**<1.0 Å**
Bursulaya et al.,2003	Autodock	46 %	24 %	9/37	9/37	1/9	/
	Dock	30 %	8 %	9/37	9/37	1/9	/
	FlexX	35 %	11 %	9/37	9/37	2/9	/
	ICM	76 %	41 %	9/37	9/37	5/9	2/9
	GOLD	46 %	22 %	9/37	9/37	3/9	/
Chen et al., 2006	FlexX	43 %	26 %	42/164	34/164	/	/
	GOLD	55 %	39 %	42/164	34/164	/	/
	GLIDE	63 %	49 %	42/164	34/164	/	/
	ICM	91 %	56 %	42/164	34/164	/	/
Xin Hu et al.,2004	Autodock	40 %	/	40	40	16/40	/
	Dock	20 %	/	40	40	8/40	/
	FlexX	35 %	/	40	40	14/40	/
	DrugScore	40 %	/	40	40	16/40	/
	GOLD	53 %	/	40	40	21/40	/

The ability of different docking programs to reproduce the X-ray crystallographic conformation of native ligand within a metal ion containing binding pocket was reassessed in three comparative studies (Table 5.1). Table 5.1 shows that 9 and 40 metal containing targets were evaluated by Bursulaya [122] and Hu, respectively. The docked complexes with RMSD values <2Å showed relatively poor performance ranging from 1 to 5 out of 9 by Bursulaya,

and 8 to 21 out of 40 by Hu. These rates were surprisingly low, indicating that docking programs often failed to find the correct binding mode in the presence of metal ion within active site. However, both Bursulaya and Chen [123] found relative high overall docking accuracy for the ICM program (76% and 91% respectively) with RMSD <2 Å. However, more than 70% of the tested targets did not have metal ions at the active site, indicating that the metal ion at the active site is a problem for docking studies. Based on these facts, we can say that Chen and Bursulay's relatively high docking accuracy was a direct result of a low number of metal containing targets. Upon this careful examination, we can conclude that a possible source of discrepancy that might rationalize the observed difference in success rates of different docking program could, arguably, be ascribed both to the lack of experience in how to get the best out of the program and the management of metal ion within the binding site. This would serve to exaggerate the difference in ICM performance between other docking programs. But clearly, there are potentially many other reasons: versions of the code used, the definition of the binding site, the human element in the study, and so on. However, as Hu suggested, using constrains between zinc and ZBG, the correct binding mode could be predicted more easily.

It is believed that, when an inhibitor replaces the zinc bound water, particularly in thermolysin, the coordinated atom occupies a similar position as the water oxygen [65]. This may be utilized as another clue to predict binding poses. Traditionally, considering the chemical repulsion to small ligand in solution and irrelevant with binding, waters outside the binding site are deleted before docking. Examining the water molecules around bound ligands in crystal structures shows that waters molecules may appear in the region of the ligand. Moreover, these waters bare distance below 3 Å between ketone or carboxyl oxygen of the ligands, which indicating strong hydrogen bond interaction and that they may influence the ligand position. Table 5.2 shows the water molecules within Van der Waal distance around ligand atoms, and distance between ligand atoms and water molecules. The sRMSD relative to the corresponding water molecules in other X-ray structures are also given. Each of 5 complexes (1y3g, 1thl, 1os0, 1qf0 and 1qf2) had water molecule corresponding to water173 in 5tmn with low sRMSD. This suggests that the docking might be performed with water.

Table 5.2: Interaction of water with the ligand in the crystal structures of thermolysin.

PDB code of Complex	Ligand atoms	Water molecules within Van Der Waal distance around ligand atoms	Distance between ligand atoms and water molecules (in Å)	Corresponding water in other X-ray structure	sRMSD between two waters (in Å)
5tmn	a_5tmn.i/1/o1	a_5tmn.w173/173/o	2.6	a_6tmn.w167/167/o	0.22
	A_5tmn.i/^L4/o	a_5tmn.w173/173/o	2.6		
	a_5tmn.i/^L4/oxt	a_5tmn.w171/171/o	2.6		
	A_5tmn.i/2/o1	a_5tmn.w28/28/o	2.7	a_6tmn.w29/29/o	0.06
6tmn	A_6tmn.i/2/o1	a_6tmn.w29/29/o	2.9		
	a_6tmn.i/1/o1	a_6tmn.w167/167/o	2.6		
	a_6tmn.i/^L4/o	a_6tmn.w167/167/o	2.4		
1y3g	a_1y3g.i/1/o	a_1y3g.w108/606/o	2.8	a_5tmn.w173/173/o	0.84
	a_1y3g.i/^L3/o	a_1y3g.w108/606/o	2.5	a_6tmn.w167/167/o	0.71
1thl	a_1thl.m7/^W322/o	a_1thl.w137/500/o	2.7	a_5tmn.w173/173/o	2.02
	a_1thl.m7/^W322/oxt	a_1thl.w143/506/o	2.6	a_5tmn.w171/171/o	1.17
1os0	a_1os0.m/^F3/o	a_1os0.w186/586/o	2.8	a_1thl.w137/500/o	1.2
	a_1os0.m/^F3/oxt	a_1os0.w93/493/o	2.5	a_1thl.w143/506/o	0.89
	a_1os0.m/1/o2p	a_1os0.w70/470/o	2.9	a_6tmn.w29/29/o	0.13
				a_5tmn.w173/173/o	1.01
1qf0	a_1qf0.ti2/317/o3	a_1qf0.w67/393/o	2.8	a_5tmn.w173/173/o	1.11
	a_1qf0.ti2/317/oxt	a_1qf0.w131/457/o	2.9	a_5tmn.w171/171/o	2.47
	a_1qf0.ti2/317/oxt	a_1qf0.w90/416/o	2.7		
1qf1	a_1qf1.ti1/317/o3	a_1qf1.w87/413/o	2.8	a_1qf0.w67/393/o	0.68
	a_1qf1.ti1/317/oxt	a_1qf1.w152/478/o	2.8	a_1qf0.w131/457/o	0.51
	a_1qf1.ti1/317/oxt	a_1qf1.w133/459/o	2.9	a_1qf0.w90/416/o	0.4
1qf2	a_1qf2.ti3/317/o2	a_1qf2.w216/543/o	2.9	a_5tmn.w173/173/o	0.5
	a_1qf2.ti3/317/oxt	a_1qf2.w216/543/o	3		

5.1 25 ligands from literature

Despite the tedious calculating during the first docking of 25 ligands, the first time docking played major rule in exploring and seeking the ideal docking method for theromolysin. Due to some bad interaction between ligand and zinc ion (spotted atomic impact in cpk representation mode) were among the best ranked poses, these poses were subjected to further refinement method.

Although we found the elimination of those bad interactions after refinement, the deviation of Cal ΔG from Exp.ΔG were even more obvious after refinements for most of the ligands (Table 4.1). Regression analysis showed that the relationship between experimental and predicted inhibition (experimental and predicted ΔG values after refinement, shown in Table 4.1) of the molecules were not linear (R^2 = 0.1874), thus indicated that the refinement may have induced more errors, and therefore the docking approach was modified.

Eight X-ray crystal structure complexes of thermolysin with ligand are available in the PDB-database. These were analyzed to investigate the trends in geometrical and stereo-chemical binding patterns of the ligands, and especially the ZBG. The observations were helpful in improving the prediction of ligand binding poses, and provide valuable guideline for screening thermolysin inhibitors, and formed the basis of reclassifying the 25 ligands according to their similarity of to one of the crystal structures.

Ligands from the 8 X-ray crystal complexes were then docked into thermolysin in the conformation seen in the X-ray complexes. In this way we could also investigate the dependence of the starting geometry on the docked conformation. However the initial geometry of the ligand never explicitly enters as a docked conformation due to changing of bond angles during the adding of hydrogen atoms and geometry optimizing process before docking. Phosphorus based protease inhibitors are acidic, with pKa values of 1,4-3.1 [169], and therefore carry a negative charge at physiological pH, while dialkylsilanediols (SiO_2H_2) are less acidic, with pKa near 10-12, and neutral at pH 7-7.4 [136]. All ligands from the 8 X-ray complexes were, except 1y3g, therefore docked in a ionized form.

It is well known from the literature that small inhibitors could be accurately docked by most docking software, while docking of larger compounds (i.e., those with extended aromatic cycles or long aliphatic chains) was more problematic [101, 123, 167]. The results may also depend on the starting geometries of the ligand. The native ligands extracted directly from the X-ray complexes and optimized by MMFF generally had less RMSD from the ligand in the X-ray complexes than those manually designed (by comparing average RMSD with previous RMSD between Table 3 & 4). The calculated RMSDs in this docking result were based on heavy (non-hydrogen) atoms of the ligand; furthermore it is between two selected chemical (hetero) molecules according to the optimal chemical match but without 3D superposition of the ligands. In terms of flexibility, it is well accepted that the accuracy of any docking program decreases with the number of rotatable bonds of the ligand. As expected, RMSD from X-ray complexes increase with ligand flexibility (Table 4.4 and 4.5). However, it

is quite modest for sets of ligands having 8-13 rotatable bonds in this study (Fig 4.3 & 4.4). There is less RMSD from X-ray complex within 5 rounds of docking for each inhibitor with neturaled Glu143 group. Generally, both with charged and neutral Glu143, the docking successfulness seemed to be increased (comparing with previous RMSD value). Lignads with rotatable bonds >13 had slightly less average RMSD from the X-ray crystal structure with charged Glu143 than with neutral Glu143. On the contrary, ligands with number of rotatable bonds <13 had average RMSD more similar to X-ray complexes with neutral Glu143. For both approaches, docking energy inclined to decrease as the number of rotatable bond in ligands increased (Fig 4.1 & 4.2). Except a slightly decrease of average energy with neutral Glu143 group, there was not dramatic individual energy difference between each docking approaches. Five and six membered rings, as well as tryptophan group in 1thl are generally well predicted in both methods (Figure 4.29.2). Generally, neualizing the Glu143 seemed to result in slight improvement in poses prediction, but docking energies with both charged and neutral Glu143 groups were not correlating with experimental pK_i values.

The large number of interactions between the ligand and the target (Table 4.9), contribute to the specificity of an enzyme for different ligands and discriminate between small differences in the structure of the ligand leading to a range of Ki values of ligand enzyme interaction. Experimental Ki values of the ligand in 5tmn and 6tmn complexes tell us that TLN can discriminate between very small changes in ligand structure. Example of 5tmn and 6tmn also indicates that a nitrogen instead of an oxygen increased binding affinity. In general, bidentate arrangements of phosphate oxygens (compound 2, 14, 15, 16, 17, 19, 20 and 22), silanediols (compound 1 and 18) and OCCS (compound 5, 6 and 7) groups inhibit down to the nanomolar concentration range. These consistencies in affinities may arise from a better fit of those bidentate groups to the thermolysin active site zinc, compared with the fit of Zn-O (ketone) and Zn-S (sulfate). It is now well established that the combination of these moiety and a specific amino acid or peptide moiety in the same molecule may result in potent inhibitors of metallopeptidases [170].

Besides thermolysin, this strategy has also been successfully utilized in the design of inhibitors for a variety of enzymes such as angiotensin-converting enzyme (ACE) [58, 171], carboxypeptidase A and B [172, 173] and neutral endopeptidase (NEP) [52]. Unfortunately this properties was not well discriminated in modified docking method since the difference in docking energy (such as the docking energy differences of 5tmn and 6tmn in Table 4.8) was not as dramatic as experimental K_i values indicated .

We assume that correct geometry of the coordination bonds between the ZBG and the catalytic zinc is important for docking accuracy and scoring reliability [144]. If ICM could impose constraints on the ZBG (e.g., by requiring that ZBG group atoms are coordinated to Zinc at the position that most of ZBG repeatedly occur) may improve the satisfaction of targeted protein interactions and speeds up the calculations by reducing the size of the conformational space that needs to be examined [101]. The protein data bank does not store any information about covalent bond types and formal charges of the chemical compounds interacting with proteins. It is therefore impossible to automatically convert those molecules to anything sensible and manual interactive assignment of bond types and formal charges for each compound in a pdb-entry is required. Therefore, apply formal charges, MMFF and bond type, as well as optimize conformation is necessary before docking. There are also some pitfalls by using ICM: Inaccurate target model, or incorrectly converted ligands, or insufficient optimization effort may lead to incorrect predictions.

As we have seen in the present study, parameterization of metal ions is a problem for docking studies. Metal ions as zinc may have several coordination numbers, and this is not taken directly into account by the docking program. Special treatment is necessary. In this case, consideration must also be taken into account with respect to specific interactions with the enzyme concerning the active site and inhibitor properties. In short, rational design of the active site to generate changes in enzyme ligand interaction requires understanding and balancing of the basic properties of the amino acid side chains and bound ligand.

5.2 *In vitro* assay

Experimental assay is well recognized as golden standard to test computer predicted binding affinities. It provides high throughput screening of samples which enables us to determine IC_{50} values from low concentrations of putative inhibitors. Although in general, fluorescent assay is much more sensitive than spectrophotometric assays, our method in this study showed its economical, efficient and cost worthy side. Besides, fluorescent assay can suffer from interference caused by impurities and the instability of many fluorescent compounds when exposed to light. The modified spectrophotometric method of Feder and Schuck [135] offer several advantages over conventional spectrophotometric assays. The requirement for small sample volumes allows for repeated sample analysis, and multiple samples within 96-well microplate instead of single curette can be analyzed simultaneously in a short time period. But still, results using this method should be interpreted with caution due

to several influencing factors. These factors can be such as the degradation of DMSO solubilized compounds due to moisture absorbed from air [174], and changing of sample concentration due to evaporation of alcohol in buffer. Generally, fairly elaborate workup of sampling, timing and controlling of the temperature and buffer concentration may need to an undue loss of sensitivity and accuracy.

5.3 R compounds

The interaction of ligands with enzymes result in multiple contacts with the amino acids surrounding the binding pocket, so that only small space is left between the protein and the ligand and only part of the ligand is exposed to the solvent. Unlike the 25 ligands from literatures, most active R compounds bare monodenate Zn-O arrangement. The average Zn-O distance for the R compounds was 2.26Å which is very close to the Zn-O distance of 2.01 ± 0.11 Å of monodenate carboxylate [65]. This, together with hydrophobic interaction contributed to the close packing of TLN and most active R compounds. Especially for compound RSH57, ZBG binding may contribute significantly to the predicted binding energy (-9.69 kcal/mol). Hu also mentioned in his study that for cases with good ZBG binding, the energy contributions from the zinc/ligand interactions amounted up to 20% of the total energies. He also suggested that improved scoring function, may alleviate a generally poor binding mode prediction [144]. In fact, the packing is even closer than illustrated in these figures as the space filling structures were generated taking into account only the van der Waals radii of atoms other than hydrogen. This close packing leads to strong bonding arising from van der Waals contacts due to the complementary surfaces, the formation of hydrogen and ionic bonds, and the hydrophobic interactions (compound RS17 and RSH57 in Figure 5.2).

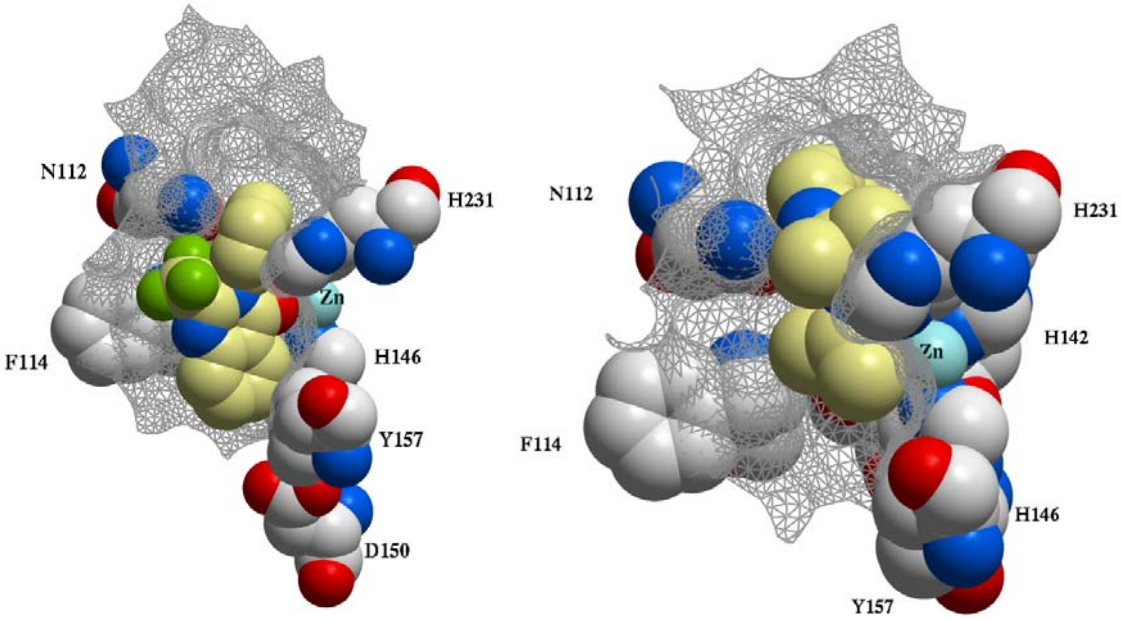


Figure 5.2: Space filling representation shows how the **RS17** (left) and **RSH57** (right) occupy the binding pockets (grey grid mode). Amino acids forming hydrophobic interactions with RS17 and RSH57 are shown.

The fluorinated inhibitors or bio-compounds have displayed a number of interesting physico-chemical and biological properties which made them very attractive in drug discovery or carrying, targeting and delivering devices [175, 176]. The rules of fluorines were summarized as below:

1. The fluorinated bio-compounds were found to enhance the hydrophobic interaction, exhibit anti-HIV activity and contribute to increased binding affinity [175, 177, 178]. In a study by E.J, Jacobsen, a similar nonfluorinated analog of inhibitor had led to 40-fold lower potency for stromelysin (a member of matrix metalloproteinase) [179]. This may be one of the reasons that caused very low IC_{50} value for compound RS17 in which contains hydrophobic group trifluoromethyl.
2. Combining fluorine showed dramatic improvement in oral activity in the pleural cavity assay, and increased protein binding as well as potency [180].

Thirty-eight fluorine or trifluoromethyl substructures containing EC3 (Hydrolase Nomenclature of enzyme) [181] enzyme-inhibitor complexes were retrieved from protein data bank and studied (Table 5.1). All complexes contained EC3 hydrolases. Eight of them have ligand-metal ion interaction. Fluorine group populated either outside binding pocket, or surrounded by several hydrophobic side chains of local residues. This indication was used as a guide to screen out unfavorable poses appeared during docking stack conformations. From

Figrue 5.2 we know that the trifluoromethyl group in RS17 is not essential for binding, so it point towards the solvent.

Table 5.1: Fluorine or trifluoromethyl functional group containing enzyme-inhibitor complexes

Ligands containing F	Metals in the target protein	Ligand action with metal	Interaction of F with metal
1ad8	Na	No	No
1b0e	Ca	No	No
1b0f	No	No	No
1bma	Ca	No	No
1eas	Na	No	No
1eat	Na	No	No
1eld	Ca	No	No
1fdk	Ca	yes, P atom	No
1fx9	Ca	yes, P atom	No
1fxf	Ca	yes, P atom	No
1gg6	No	No	No
1hbj	No	No	No
1hn4	Ca	No	No
1mmj	No	No	No
1vsn	No	No	No
1xe5	No	No	No
1xe6	No	No	No
1y6o	Ca	No	No
1yl1	Ca	No	No
2cng	Mg	No	No
2fj0	No	No	No
2fm0	Zn	yes, O atom	No
2fm5	Zn	yes, O atom	No
1hfs	Zn and Ca	yes, 2O atoms	No
1usn	Zn and Ca	yes, S atom	No
2fog	Ca	No	No
2foh	Ca	No	No
2fq9	No	No	No
2fra	No	No	No
2h9y	No	No	No
2i0u	Ca	No	No
2iit	No	No	No
2ow1	Zn and Ca	Yes, 2O atoms	No
2v96	No	No	No
2v97	No	No	No
2v98	No	No	No
6gch	No	No	No
7gch	No	No	No

5.4 MS compounds

A significant correlation between experimental binding affinities and computed score was found for this group of compounds (Table 4.14). The initial docking indicated that all compounds of this series could bind thermolysin. However, the experimental verification showed that only two compounds were binders. The cause of inactive among most MS compounds is not currently clear, but probably due to the difference in the reaction conditions and toxicity to thermolysin. Hashida suggested in his study that an inhibitor may inhibit thermolysin activity not only as a competitive inhibitor but also by promoting autolysis [153]. Calcium is considered to be crucial for the thermal stability of thermolysin [182], especially Calcium 4 is related to the autolysis [183]. Therefore, the toxicity of MS compounds to thermolysin may be induced by eliminating the Ca^{2+} by a conformational change induced by the interaction of MS compounds with thermolysin. We also speculate that the bulky structure itself may also keep zinc binding ketone group away from interacting distance. The kinetic and time dependent studies (data not shown in this study) showed that, M9 is a competitive and fast inhibitor; where as M12 is an uncompetitive and slow inhibitor of thermolysin [150]. After two hours still the reaction was going on.

6 Conclusion

The ability of ICM in this study to reproduce the X-ray crystallographic conformation of native ligand at the active site of thermolysin was 63% (5 best ranked docking poses out of 8 X-ray complexes had sRMSD values < 2Å). 12 out of 37 R-compounds inhibited thermolysin with varying IC_{50} values [151]. MS9 was found as competitive inhibitor of thermolysin, while MS12 is an uncompetitive and slow inhibitor of thermolysin. 2 out of 7 M compounds had the thermolysin inhibition property, and one of them was very potent inhibitor, with IC_{50} value of $4{,}411 \times 10^{-11}$ mM.

If an answer to a scientist's question does not exist in the literature, a computed answer is often better than no answer. Ultimately, docking program should be able to identify novel potential 'binders' very accurately. However, to reach this goal, many issues, including those discussed throughout this study, and scoring the binding to the metal have to be addressed. Specific metal binding term may need to be developed. Overall, this study indicates that a certain degree of improvement has been achieved both in docking and in experimental methodology and in both cases the technology developed for ICM and inhibition assay appears to provide the most consistent benefits. An interesting proposal would be to set up constrains between zinc and zinc-binding group of putative inhibitors (namely, template docking method). This may guides the binding of ligands to zinc and speeds up the calculations by reducing the size of the phase space that needs to be examined.

REFERENCES

1. Branden., C. and J. Tooze., *Introduction to Protein Structure*. Second ed. 1999: Garland Publishing, Inc. pp 12, 21-32, 205,.
2. Whitford, D., *Proteins structure and function*. 2005: John Wiley & Sons Ltd. pp189,224,225.
3. Rawlings, N.D., Morton, F.R. & Barrett, A.J. (2006) MEROPS: the peptidase database. Nucleic Acids Res 34, D270-D272, *http://merops.sanger.ac.uk/*. Jan-25-2008.
4. Lipscomb, W.N. and N. Strater, *Recent Advances in Zinc Enzymology.* Chem Rev, 1996. **96**(7): p. 2375-2434.
5. Roques, B.P., *Zinc metallopeptidases: active site structure and design of selective and mixed inhibitors: new approaches in the search for analgesics and anti-hypertensives.* Biochem Soc Trans, 1993. **21 (Pt 3)**(3): p. 678-85.
6. Blundell, T.L., *Metalloproteinase superfamilies and drug design.* Nat Struct Biol, 1994. **1**(2): p. 73-5.
7. H.D.Holtje, et al., *Molecular Modeling-Basic Principle and Appication*. 2nd ed. 2003: Wiley-VCH. pp98,115.
8. Fersht, A., *Enzyme structure and mechanism*. Second Edition ed. 1985: W.H.Freeman and company New York.
9. Holmes, M.A. and B.W. Matthews, *Structure of thermolysin refined at 1.6 A resolution.* J Mol Biol, 1982. **160**(4): p. 623-39.
10. Dideberg, O., et al., *Structure of a Zn2+-containing D-alanyl-D-alanine-cleaving carboxypeptidase at 2.5 A resolution.* Nature, 1982. **299**(5882): p. 469-70.
11. Holmes, M.A. and B.W. Matthews, *Binding of hydroxamic acid inhibitors to crystalline thermolysin suggests a pentacoordinate zinc intermediate in catalysis.* Biochemistry, 1981. **20**(24): p. 6912-20.
12. Orengo, C.A., et al., *CATH--a hierarchic classification of protein domain structures.* Structure, 1997. **5**(8): p. 1093-108.
13. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Res, 2000. **28**(1): p. 235-42.
14. Holland, D.R., et al., *Structural analysis of zinc substitutions in the active site of thermolysin.* Protein Sci, 1995. **4**(10): p. 1955-65.
15. Banbula, A., et al., *Amino-acid sequence and three-dimensional structure of the Staphylococcus aureus metalloproteinase at 1.72 A resolution.* Structure, 1998. **6**(9): p. 1185-93.
16. Beaumont, A., et al., *The role of histidine 231 in thermolysin-like enzymes. A site-directed mutagenesis study.* J Biol Chem, 1995. **270**(28): p. 16803-8.
17. Gaucher, J.F., et al., *The 2.2 A resolution structure of thermolysin (TLN) crystallized in the presence of potassium thiocyanate.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 12): p. 2198-200.
18. Gokhale, N.H., S. Bradford, and J.A. Cowan, *Stimulation and oxidative catalytic inactivation of thermolysin by copper.Cys-Gly-His-Lys.* J Biol Inorg Chem, 2007. **12**(7): p. 981-7.
19. Matsubara, H., et al., *Specific nature of hydrolysis of insulin and tobacco mosaic virus protein by thermolysin.* Arch Biochem Biophys, 1966. **115**(2): p. 324-31.
20. Morihara, K. and H. Tsuzuki, *Thermolysin: kinetic study with oligopeptides.* Eur J Biochem, 1970. **15**(2): p. 374-80.
21. Roques, B.P., et al., *Neutral endopeptidase 24.11: structure, inhibition, and experimental and clinical pharmacology.* Pharmacol Rev, 1993. **45**(1): p. 87-146.

22. Cheng, X.-M.A., Kyunghye; Haleen, Stephen J. , *Endothelin inhibitors* Annual Reports in Medicinal Chemistry, 1997. **32**: p. 61-70.
23. de Kreij, A., G. Venema, and B. van den Burg, *Substrate specificity in the highly heterogeneous M4 peptidase family is determined by a small subset of amino acids.* J Biol Chem, 2000. **275**(40): p. 31115-20.
24. Matthews, B.W., *Structural basis of the action of thermolysin and related zinc peptidases.* Acc Chem Res, 1988. **21**(333-340).
25. Bohacek R, D.L.S., McMartin C, Priestle J, Gruetter M, *Three-Dimensional Models of ACE and NEP Inhibitors and Their Use in the Design of Potent Dual ACE/NEP Inhibitors* J. Am. Chem. Soc., , 1996. **118 (35),** : p. 8231 -8249,.
26. Maeda, H., *Role of microbial proteases in pathogenesis.* Microbiol Immunol, 1996. **40**(10): p. 685-99.
27. Miyoshi, S. and S. Shinoda, *Microbial metalloproteases and pathogenesis.* Microbes Infect, 2000. **2**(1): p. 91-8.
28. Altincicek, B., et al., *Microbial metalloproteinases mediate sensing of invading pathogens and activate innate immune responses in the lepidopteran model host Galleria mellonella.* Infect Immun, 2007. **75**(1): p. 175-83.
29. Ohta, Y., Y. Ogura, and A. Wada, *Thermostable protease from thermophilic bacteria. I. Thermostability, physiocochemical properties, and amino acid composition.* J Biol Chem, 1966. **241**(24): p. 5919-25.
30. Matsuyama, S., et al., *Protease-mediated enhancement of severe acute respiratory syndrome coronavirus infection.* Proc Natl Acad Sci U S A, 2005. **102**(35): p. 12543-7.
31. Jin, F., et al., *Purification, characterization, and primary structure of Clostridium perfringens lambda-toxin, a thermolysin-like metalloprotease.* Infect Immun, 1996. **64**(1): p. 230-7.
32. Hobden, J.A., *Pseudomonas aeruginosa proteases and corneal virulence.* DNA Cell Biol, 2002. **21**(5-6): p. 391-6.
33. Mariencheck, W.I., et al., *Pseudomonas aeruginosa elastase degrades surfactant proteins A and D.* Am J Respir Cell Mol Biol, 2003. **28**(4): p. 528-37.
34. Naka, A., T. Honda, and T. Miwatani, *A simple purification method of Vibrio cholerae non-O1 hemagglutinin/protease by immunoaffinity column chromatography using a monoclonal antibody.* Microbiol Immunol, 1992. **36**(4): p. 419-23.
35. Schmidtchen, A., et al., *Elastase-producing Pseudomonas aeruginosa degrade plasma proteins and extracellular products of human skin and fibroblasts, and inhibit fibroblast growth.* Microb Pathog, 2003. **34**(1): p. 47-55.
36. Smith, A.W., B. Chahal, and G.L. French, *The human gastric pathogen Helicobacter pylori has a gene encoding an enzyme first classified as a mucinase in Vibrio cholerae.* Mol Microbiol, 1994. **13**(1): p. 153-60.
37. Clare, B.W., A. Scozzafava, and C.T. Supuran, *Protease inhibitors: synthesis of a series of bacterial collagenase inhibitors of the sulfonyl amino acyl hydroxamate type.* J Med Chem, 2001. **44**(13): p. 2253-8.
38. Travis, J. and J. Potempa, *Bacterial proteinases as targets for the development of second-generation antibiotics.* Biochim Biophys Acta, 2000. **1477**(1-2): p. 35-50.
39. Lee, L.M., et al., *Low molecular weight protamine as nontoxic heparin/low molecular weight heparin antidote (III): preliminary in vivo evaluation of efficacy and toxicity using a canine model.* AAPS PharmSci, 2001. **3**(3): p. E19.
40. Eichhorn, U., et al., *Synthesis of dipeptides by suspension-to-suspension conversion via thermolysin catalysis: from analytical to preparative scale.* J Pept Sci, 1997. **3**(4): p. 245-51.

41. G. Krix, U.E., H. -D. Jakubke and M. -R. Kula,, *Protease-catalyzed synthesis of new hydrophobic dipeptides containing non-proteinogenic amino acids* Enzyme and Microbial Technology,Copyright © 1997 Published by Elsevier Science Inc. , September 1997,. **21,** (4): p. 252-257
42. Takami, H., T. Akiba, and K. Horikoshi, *Characterization of an alkaline protease from Bacillus sp. no. AH-101.* Appl Microbiol Biotechnol, 1990. **33**(5): p. 519-23.
43. Takahashi, T., et al., *Molecular cloning of the gene encoding Vibrio metalloproteinase vimelysin and isolation of a mutant with high stability in organic solvents.* J Biochem (Tokyo), 2005. **138**(6): p. 701-10.
44. Houmard, J. and G.R. Drapeau, *Staphylococcal protease: a proteolytic enzyme specific for glutamoyl bonds.* Proc Natl Acad Sci U S A, 1972. **69**(12): p. 3506-9.
45. Adekoya, O.A., et al., *Comparative sequence and structure analysis reveal features of cold adaptation of an enzyme in the thermolysin family.* Proteins, 2006. **62**(2): p. 435-49.
46. Erbeldinger, M., X. Ni, and P.J. Halling, *Kinetics of enzymatic solid-to-solid peptide synthesis: synthesis of Z-aspartame and control of acid-base conditions by using inorganic salts.* Biotechnol Bioeng, 2001. **72**(1): p. 69-76.
47. Ulijn, R.V., M. Erbeldinger, and P.J. Halling, *Comparison of methods for thermolysin-catalyzed peptide synthesis including a novel more active catalyst.* Biotechnol Bioeng, 2000. **69**(6): p. 633-8.
48. Ulijn, R.V., et al., *Protease-catalyzed peptide synthesis on solid support.* J Am Chem Soc, 2002. **124**(37): p. 10988-9.
49. Kunugi, S., et al., *Kinetic characterization of the neutral protease vimelysin from Vibrio sp. T1800.* Eur J Biochem, 1996. **241**(2): p. 368-73.
50. Durham, D.R., *The unique stability of Vibrio proteolyticus neutral protease under alkaline conditions affords a selective step for purification and use in amino acid-coupling reactions.* Appl Environ Microbiol, 1990. **56**(8): p. 2277-81.
51. Tiraboschi, G., et al., *A three-dimensional construction of the active site (region 507-749) of human neutral endopeptidase (EC.3.4.24.11).* Protein Eng, 1999. **12**(2): p. 141-9.
52. Holland, D.R., et al., *Inhibition of thermolysin and neutral endopeptidase 24.11 by a novel glutaramide derivative: X-ray structure determination of the thermolysin-inhibitor complex.* Biochemistry, 1994. **33**(1): p. 51-6.
53. H.P.Rang, et al., *Pharmacology*. 5th ed. Vasoactive drugs. 2003: Churchil Livingstone. p296-298.
54. Yang, H.Y., E.G. Erdos, and Y. Levin, *A dipeptidyl carboxypeptidase that converts angiotensin I and inactivates bradykinin.* Biochim Biophys Acta, 1970. **214**(2): p. 374-6.
55. Williams, T.A., P. Corvol, and F. Soubrier, *Identification of two active site residues in human angiotensin I-converting enzyme.* J Biol Chem, 1994. **269**(47): p. 29430-4.
56. Natesh, R., et al., *Crystal structure of the human angiotensin-converting enzyme-lisinopril complex.* Nature, 2003. **421**(6922): p. 551-4.
57. Soubrier, F., et al., *Two putative active centers in human angiotensin I-converting enzyme revealed by molecular cloning.* Proc Natl Acad Sci U S A, 1988. **85**(24): p. 9386-90.
58. Waeber, B., J. Nussberger, and H.R. Brunner, *[Angiotensin-converting-enzyme inhibition in arterial hypertension].* Wien Med Wochenschr, 1990. **140**(1-2): p. 22-30.
59. Bland, N.D., et al., *Bioinformatic analysis of the neprilysin (M13) family of peptidases reveals complex evolutionary and functional relationships.* BMC Evol Biol, 2008. **8**(1): p. 16.

60. Moult, J., et al., *A large-scale experiment to assess protein structure prediction methods.* Proteins, 1995. **23**(3): p. ii-v.
61. Frigerio, F., et al., *Model building of a thermolysin-like protease by mutagenesis.* Protein Eng, 1997. **10**(3): p. 223-30.
62. Colman, P.M., J.N. Jansonius, and B.W. Matthews, *The structure of thermolysin: an electron density map at 2-3 A resolution.* J Mol Biol, 1972. **70**(3): p. 701-24.
63. Vallee, B.L. and D.S. Auld, *Active-site zinc ligands and activated H2O of zinc enzymes.* Proc Natl Acad Sci U S A, 1990. **87**(1): p. 220-4.
64. Rulisek, L. and J. Vondrasek, *Coordination geometries of selected transition metal ions (Co2+, Ni2+, Cu2+, Zn2+, Cd2+, and Hg2+) in metalloproteins.* J Inorg Biochem, 1998. **71**(3-4): p. 115-27.
65. Alberts, I.L., K. Nadassy, and S.J. Wodak, *Analysis of zinc binding sites in protein crystal structures.* Protein Sci, 1998. **7**(8): p. 1700-16.
66. Jongeneel, C.V., J. Bouvier, and A. Bairoch, *A unique signature identifies a family of zinc-dependent metallopeptidases.* FEBS Lett, 1989. **242**(2): p. 211-4.
67. Pelmenschikov, V., M.R. Blomberg, and P.E. Siegbahn, *A theoretical study of the mechanism for peptide hydrolysis by thermolysin.* J Biol Inorg Chem, 2002. **7**(3): p. 284-98.
68. Rawlings, N.D. and A.J. Barrett, *Evolutionary families of metallopeptidases.* Methods Enzymol, 1995. **248**: p. 183-228.
69. Mock, W.L. and D.J. Stanford, *Arazoformyl dipeptide substrates for thermolysin. Confirmation of a reverse protonation catalytic mechanism.* Biochemistry, 1996. **35**(23): p. 7369-77.
70. Argos, P., et al., *Similarities in active center geometries of zinc-containing enzymes, proteases and dehydrogenases.* J Mol Biol, 1978. **126**(2): p. 141-58.
71. Taylor, K.B., *Enzyme Kinetics and Mechanisms. CHAPTER 6 EFFECTS OF ANALOG INHIBITORS p 66-100.* 2002, Secaucus, NJ, USA: Kluwer Academic Publishers. http://site.ebrary.com/lib/tromsoub/Doc?id=10067367&ppg=81.
72. Savageau, M.A., *Michaelis-Menten mechanism reconsidered: implications of fractal kinetics.* J Theor Biol, 1995. **176**(1): p. 115-24.
73. Lineweaver, H. and D. Burk, *The Determination of Enzyme Dissociation Constants.* Journal of the American Chemical Society, 1934. **56**: p. 658-666.
74. NIH, *Assay Operations for SAR Support* 2008, Copyright © 2008, Eli Lilly and Company and the National Institutes of Health Chemical Genomics Center. http://www.ncgc.nih.gov/guidance/section3.html.
75. Cheng, H.C., *The power issue: determination of KB or Ki from IC50. A closer look at the Cheng-Prusoff equation, the Schild plot and related power equations.* J Pharmacol Toxicol Methods, 2001. **46**(2): p. 61-71.
76. Smith, H.J. and C. Simons, *Enzymes and Their Inhibition: Drug Development (Enzyme Inhibitors)*. 2005: Taylor & Francis Ltd. 185.
77. Sousa, S.F., P.A. Fernandes, and M.J. Ramos, *Protein-ligand docking: current status and future challenges.* Proteins, 2006. **65**(1): p. 15-26.
78. Kuntz, I.D., et al., *A geometric approach to macromolecule-ligand interactions.* J Mol Biol, 1982. **161**(2): p. 269-88.
79. Kitchen, D.B., et al., *Docking and scoring in virtual screening for drug discovery: methods and applications.* Nat Rev Drug Discov, 2004. **3**(11): p. 935-49.
80. Ha, S., et al., *Evaluation of docking/scoring approaches: a comparative study based on MMP3 inhibitors.* J Comput Aided Mol Des, 2000. **14**(5): p. 435-48.
81. Teague, S.J., *Implications of protein flexibility for drug discovery.* Nat Rev Drug Discov, 2003. **2**(7): p. 527-41.

82. Carlson, H.A., *Protein flexibility and drug design: how to hit a moving target.* Curr Opin Chem Biol, 2002. **6**(4): p. 447-52.
83. Carlson, H.A., *Protein flexibility is an important component of structure-based drug discovery.* Curr Pharm Des, 2002. **8**(17): p. 1571-8.
84. Morris GM, G.D., Halliday RS, Huey R, Hart WE, Belew RK et al. , *Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.* J Comput Chem, 1998. **19**: p. 1639-1662.
85. Osterberg F, M.G., Sanner MF, Olson AJ, Goodsell DS. , *Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in autodock.* Proteins Struct Funct Genet, 2002. **46**: p. 34-40.
86. Verdonk ML, C.J., Hartshorn MJ, Murray CW, Taylor RD . , *Improved protein-ligand docking using GOLD.* . Proteins Struct Funct Genet, 2003. **52**: p. 609-623.
87. Verdonk ML, C.G., Cole JC, Hartshorn MJ, Murray CW, Nissink JWM et al. , *Modeling water molecules in protein-ligand docking using GOLD.* . J Med Chem, 2005. **48**: p. 6504-6515.
88. Eldridge, M.D., et al., *Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes.* J Comput Aided Mol Des, 1997. **11**(5): p. 425-45.
89. Jones-Hertzog DK, J.W., *Binding affinities for sulfonamide inhibitors with human thrombin using Monte Carlo simulations with a linear response method.* J Med Chem, 1997. **40**: p. 1539-1549.
90. Rarey, M., et al., *A fast flexible docking method using an incremental construction algorithm.* J Mol Biol, 1996. **261**(3): p. 470-89.
91. Oshiro CM, K.I., Dixon JS . , *Flexible ligand docking using a genetic algorithm.* J Comput Aided Mol Des, 1995. **Des 9**: p. 113-130.
92. Knegtel RMA, K.I., Oshiro CM *Molecular docking to ensembles of protein structures.* J Mol Biol, 1997. **266**: p. 424-440.
93. Kang, X., R.H. Shafer, and I.D. Kuntz, *Calculation of ligand-nucleic acid binding free energies with the generalized-born model in DOCK.* Biopolymers, 2004. **73**(2): p. 192-204.
94. Moustakas DT, L.P., Pegg S, Pettersen E, Kuntz ID, Brooijmans N et al. , *Development and validation of a modular, extensible docking program: DOCK 5.* J Comput Aided Mol 2006. **Des 20**: p. 601-619.
95. Meng EC, S.B., Kuntz ID *Automated docking with grid-based energy evaluation.* J Comput Chem, (1992). **13**: p. 505-524.
96. Abagyan, R. and M. Totrov, *Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins.* J Mol Biol, 1994. **235**(3): p. 983-1002.
97. Totrov, M. and R. Abagyan, *Flexible protein-ligand docking by global energy optimization in internal coordinates.* Proteins, 1997. **Suppl 1**: p. 215-20.
98. Abagyan R, T.M., Kuznetsov D *ICM-a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation.* . J Comput Chem 1994b. **15**: p. 488-506.
99. Friesner RA, M.R., Repasky MP, Frye LL, Greenwood JR, Halgren TA et al. , *Extra precision Glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes.* . J Med Chem, 2006. **49**: p. 6177-6196.
100. Sherman, W., et al., *Novel procedure for modeling ligand/receptor induced fit effects.* J Med Chem, 2006. **49**(2): p. 534-53.
101. Friesner, R.A., et al., *Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.* J Med Chem, 2004. **47**(7): p. 1739-49.

102. McGann, M.R., et al., *Gaussian docking functions.* Biopolymers, 2003. **68**(1): p. 76-90.
103. Bernard R. Brooks, R.E.B., Barry D. Olafson, David J. States, S. Swaminathan, Martin Karplus, *CHARMM: A program for macromolecular energy, minimization, and dynamics calculations.* Journal of Computational Chemistry, 1983. **4**(2): p. 187-217.
104. Zsoldos, Z., et al., *eHiTS: an innovative approach to the docking and scoring function problems.* Curr Protein Pept Sci, 2006. **7**(5): p. 421-35.
105. Zsoldos, Z., et al., *eHiTS: a new fast, exhaustive flexible ligand docking system.* J Mol Graph Model, 2007. **26**(1): p. 198-212.
106. Schneidman-Duhovny, D., et al., *PatchDock and SymmDock: servers for rigid and symmetric docking.* Nucleic Acids Res, 2005. **33**(Web Server issue): p. W363-7.
107. Schneidman-Duhovny, D., et al., *Taking geometry to its edge: fast unbound rigid (and hinge-bent) docking.* Proteins, 2003. **52**(1): p. 107-12.
108. Jain, A.N., *Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine.* J Med Chem, 2003. **46**(4): p. 499-511.
109. Jain, A.N., *Scoring noncovalent protein-ligand interactions: a continuous differentiable function tuned to compute binding affinities.* J Comput Aided Mol Des, 1996. **10**(5): p. 427-40.
110. Bernstein, F.C., et al., *The Protein Data Bank: a computer-based archival file for macromolecular structures.* J Mol Biol, 1977. **112**(3): p. 535-42.
111. Miller, M.D., et al., *FLOG: a system to select 'quasi-flexible' ligands complementary to a receptor of known three-dimensional structure.* J Comput Aided Mol Des, 1994. **8**(2): p. 153-74.
112. von Itzstein, M., et al., *Rational design of potent sialidase-based inhibitors of influenza virus replication.* Nature, 1993. **363**(6428): p. 418-23.
113. Chiang, C.C., et al., *Product development of AG-331 lyophilized powder for injection.* J Pharm Sci Technol, 1994. **48**(1): p. 24-9.
114. Tame, J.R., *Scoring functions: a view from the bench.* J Comput Aided Mol Des, 1999. **13**(2): p. 99-108.
115. Mohan, V., et al., *Docking: successes and challenges.* Curr Pharm Des, 2005. **11**(3): p. 323-33.
116. Tame, J.R., *Scoring functions--the first 100 years.* J Comput Aided Mol Des, 2005. **19**(6): p. 445-51.
117. Jain, A.N., *Scoring functions for protein-ligand docking.* Curr Protein Pept Sci, 2006. **7**(5): p. 407-20.
118. Klebe, G., *Virtual ligand screening: strategies, perspectives and limitations.* Drug Discov Today, 2006. **11**(13-14): p. 580-94.
119. Gohlke, H. and G. Klebe, *Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors.* Angew Chem Int Ed Engl, 2002. **41**(15): p. 2644-76.
120. Wang, R., L. Lai, and S. Wang, *Further development and validation of empirical scoring functions for structure-based binding affinity prediction.* J Comput Aided Mol Des, 2002. **16**(1): p. 11-26.
121. Ruben Abagyan, M.T., Dmitry Kuznetsov, *ICM - A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation.* Journal of Computational Chemistry, 1994. **15**(5): p. 488-506.
122. Bursulaya, B.D., et al., *Comparative study of several algorithms for flexible ligand docking.* J Comput Aided Mol Des, 2003. **17**(11): p. 755-63.

123. Chen, H., et al., *On evaluating molecular-docking methods for pose prediction and enrichment factors.* J Chem Inf Model, 2006. **46**(1): p. 401-15.
124. Moitessier, N., et al., *Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go.* Br J Pharmacol, 2007.
125. Ben-Zeev, E., et al., *Docking to single-domain and multiple-domain proteins: old and new challenges.* Proteins, 2005. **60**(2): p. 195-201.
126. Jackson, R.M., H.A. Gabb, and M.J. Sternberg, *Rapid refinement of protein interfaces incorporating solvation: application to the docking problem.* J Mol Biol, 1998. **276**(1): p. 265-85.
127. Lynn, F.T.E., et al., *Surveying molecular interactions with DOT*, in *Proceedings of the 1995 ACM/IEEE conference on Supercomputing (CDROM)*. 1995, ACM: San Diego, California, United States.
128. Ritchie, D.W., *Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2.* Proteins, 2003. **52**(1): p. 98-106.
129. Chen, R. and Z. Weng, *Docking unbound proteins using shape complementarity, desolvation, and electrostatics.* Proteins, 2002. **47**(3): p. 281-94.
130. Fitzjohn, P.W. and P.A. Bates, *Guided docking: first step to locate potential binding sites.* Proteins, 2003. **52**(1): p. 28-32.
131. Palma, P.N., et al., *BiGGER: a new (soft) docking algorithm for predicting protein interactions.* Proteins, 2000. **39**(4): p. 372-84.
132. Gardiner, E.J., P. Willett, and P.J. Artymiuk, *GAPDOCK: a Genetic Algorithm Approach to Protein Docking in CAPRI round 1.* Proteins, 2003. **52**(1): p. 10-4.
133. Duncan B, O.A., *Applications of evolutionary programming for the prediction of protein-protein interactions.*, in *Evolutionary programming V. Proceedings of the 5th annual conference on evolutionary programming*. 1996, MIT Press: Boston, MA.
134. Tovchigrechko, A. and I.A. Vakser, *Development and testing of an automated approach to protein docking.* Proteins, 2005. **60**(2): p. 296-301.
135. Feder, J. and J.M. Schuck, *Studies on the Bacillus subtilis neutral-protease- and Bacillus thermoproteolyticus thermolysin-catalyzed hydrolysis of dipeptide substrates.* Biochemistry, 1970. **9**(14): p. 2784-91.
136. Kim, J. and S.M. Sieburth, *A silanediol inhibitor of the metalloprotease thermolysin: synthesis and comparison with a phosphinic acid inhibitor.* J Org Chem, 2004. **69**(9): p. 3008-14.
137. Kim, S.J., et al., *Origin of the stereospecificity in binding hydroxamates of alpha- and beta-phenylalanine methylamide to thermolysin revealed by the X-ray crystallographic study.* Bioorg Med Chem, 2003. **11**(11): p. 2421-6.
138. Halgren, T.A., *Potential energy functions.* Curr Opin Struct Biol, 1995. **5**(2): p. 205-10.
139. Hart, T.N. and R.J. Read, *A multiple-start Monte Carlo docking method.* Proteins, 1992. **13**(3): p. 206-22.
140. Schapira, M., M. Totrov, and R. Abagyan, *Prediction of the binding energy for small molecules, peptides and proteins.* J Mol Recog, 1999. **12**: p. 177-190.
141. Wallace, A.C., R.A. Laskowski, and J.M. Thornton, *LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions.* Protein Eng, 1995. **8**(2): p. 127-34.
142. Meltzer, H.Y., S. Matsubara, and J.C. Lee, *Classification of typical and atypical antipsychotic drugs on the basis of dopamine D-1, D-2 and serotonin2 pKi values.* J Pharmacol Exp Ther, 1989. **251**(1): p. 238-46.
143. F. H. Allen and V. J. Hoy. Vol. F, c., pp. 663-668 *The Cambridge Structural Database (CSD).* International Tables for Crystallography 2006. **F**(CH 24.3): p. 663-668.

144. Hu, X., S. Balaz, and W.H. Shelver, *A practical approach to docking of zinc metalloproteinase inhibitors.* J Mol Graph Model, 2004. **22**(4): p. 293-307.
145. Gaucher, J.F., et al., *Crystal structures of alpha-mercaptoacyldipeptides in the thermolysin active site: structural parameters for a Zn monodentation or bidentation in metalloendopeptidases.* Biochemistry, 1999. **38**(39): p. 12569-76.
146. Selkti, M., et al., *Interactions of a new alpha-aminophosphinic derivative inside the active site of TLN (thermolysin): a model for zinc-metalloendopeptidase inhibition.* Acta Crystallogr D Biol Crystallogr, 2003. **59**(Pt 7): p. 1200-5.
147. Bartlett, P.A. and C.K. Marlowe, *Possible role for water dissociation in the slow binding of phosphorus-containing transition-state-analogue inhibitors of thermolysin.* Biochemistry, 1987. **26**(26): p. 8553-61.
148. Morgan, B., Scholtz, J. M., Ballinger, M. D., Zipkin, I. D., and Bartlett, P. A., *Differential binding energy: a detailed evaluation of the influence of hydrogen-bonding and hydrophobic groups on the inhibition of thermolysin by phosphorus-containing inhibitors.* J. Am. Chem. Soc., 1991. **113**(1): p. 297 - 307.
149. Blumberg, S. and B.L. Vallee, *Superactivation of thermolysin by acylation with amino acid N-hydroxysuccinimide esters.* Biochemistry, 1975. **14**(11): p. 2410-9.
150. Khan, M.T.H., Yimingjiang, W., Sultankhudzaev, M., Sylte, I. , *Discovery of Two Diterpenoid Alkaloids as Thermolysin Inhibitors Using In vitro and Docking Techniques.* . Manuscript in preparation.
151. Khan, M.T.H., Khan, R., Yimingjiang, W., Ahmed, M., Sylte, I. and *Combinatorial Synthesis of Quinazolin-4(3H)-ones as a Novel Class of Thermolysin Inhibitors.* Manuscript in progress.
152. Hangauer, D.G., A.F. Monzingo, and B.W. Matthews, *An interactive computer graphics study of thermolysin-catalyzed peptide cleavage and inhibition by N-carboxymethyl dipeptides.* Biochemistry, 1984. **23**(24): p. 5730-41.
153. Hashida, Y. and K. Inouye, *Molecular mechanism of the inhibitory effect of cobalt ion on thermolysin activity and the suppressive effect of calcium ion on the cobalt ion-dependent inactivation of thermolysin.* J Biochem, 2007. **141**(6): p. 879-88.
154. Holden, H.M. and B.W. Matthews, *The binding of L-valyl-L-tryptophan to crystalline thermolysin illustrates the mode of interaction of a product of peptide hydrolysis.* J Biol Chem, 1988. **263**(7): p. 3256-60.
155. Kitagishi, K. and K. Hiromi, *Binding between thermolysin and its specific inhibitor, N-phosphoryl-L-leucyl-L-tryptophan (PLT).* J Biochem, 1986. **99**(1): p. 191-7.
156. Kuhn, D., et al., *Boilysin and thermolysin in dipeptide synthesis: a comparative study.* Biotechnol Appl Biochem, 2002. **36**(Pt 1): p. 71-6.
157. Monzingo, A.F. and B.W. Matthews, *Binding of N-carboxymethyl dipeptide inhibitors to thermolysin determined by X-ray crystallography: a novel class of transition-state analogues for zinc peptidases.* Biochemistry, 1984. **23**(24): p. 5724-9.
158. Yasukawa, K., et al., *Characterization of Gly-D-Phe, Gly-L-Leu, and D-Phe as affinity ligands to thermolysin.* Protein Expr Purif, 2006. **46**(2): p. 332-6.
159. Tronrud, D.E., A.F. Monzingo, and B.W. Matthews, *Crystallographic structural analysis of phosphoramidates as inhibitors and transition-state analogs of thermolysin.* Eur J Biochem, 1986. **157**(2): p. 261-8.
160. Bartlett, P.A. and C.K. Marlowe, *Phosphonamidates as transition-state analogue inhibitors of thermolysin.* Biochemistry, 1983. **22**(20): p. 4618-24.
161. Bradley Morgan, J.M.S., Marcus D. Ballinger, Ilan D. Zipkin, and Paul A. Bartlett, *Differential binding energy: a detailed evaluation of the influence of hydrogen-bonding and hydrophobic groups on the inhibition of thermolysin by phosphorus-containing inhibitors.* J. Am. Chem. Soc., 1991. **113**(1): p. 297-307.

162. Kim, J., A. Glekas, and N.S.S. Mc, *Silanediol-based inhibitor of thermolysin.* Bioorg Med Chem Lett, 2002. **12**(24): p. 3625-7.
163. Nishino, N. and J.C. Powers, *Design of potent reversible inhibitors for thermolysin. Peptides containing zinc coordinating ligands and their use in affinity chromatography.* Biochemistry, 1979. **18**(20): p. 4340-7.
164. Kam, C.M., N. Nishino, and J.C. Powers, *Inhibition of thermolysin and carboxypeptidase A by phosphoramidates.* Biochemistry, 1979. **18**(14): p. 3032-8.
165. Nishino, N. and J.C. Powers, *Peptide hydroxamic acids as inhibitors of thermolysin.* Biochemistry, 1978. **17**(14): p. 2846-50.
166. Feng Cheng, R.Z., Xiaomin Luo, Jianhua Shen, Xin Li, Jiande Gu, Weiliang Zhu, Jingkang Shen, Irit Sagi, Ruyun Ji, Kaixian Chen, and Hualiang Jiang* *Quantum Chemistry Study on the Interaction of the Exogenous Ligands and the Catalytic Zinc Ion in Matrix Metalloproteinases* . Phys. Chem. B 2002. **106**(17): p. 4552 -4559.
167. Perola, E., W.P. Walters, and P.S. Charifson, *A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance.* Proteins, 2004. **56**(2): p. 235-49.
168. Kroemer, R.T., et al., *Assessment of docking poses: interactions-based accuracy classification (IBAC) versus crystal structure deviations.* J Chem Inf Comput Sci, 2004. **44**(3): p. 871-81.
169. Grobelny, D., U.B. Goli, and R.E. Galardy, *Binding energetics of phosphorus-containing inhibitors of thermolysin.* Biochemistry, 1989. **28**(12): p. 4948-51.
170. Blumberg, S. and Z. Tauber, *Inhibition of metalloendopeptidases by 2-mercaptoacetyl-dipeptides.* Eur J Biochem, 1983. **136**(1): p. 151-4.
171. Galardy, R.E., *Inhibition of angiotensin converting enzyme with N alpha-phosphoryl-L-alanyl-L-proline and N alpha-L-valyl-L-tryptophan.* Biochem Biophys Res Commun, 1980. **97**(1): p. 94-9.
172. Byers, L.D. and R. Wolfenden, *A potent reversible inhibitor of carboxypeptidase A.* J Biol Chem, 1972. **247**(2): p. 606-8.
173. Ondetti, M.A., et al., *Design of potent and specific inhibitors of carboxypeptidases A and B.* Biochemistry, 1979. **18**(8): p. 1427-30.
174. Cheng, X., et al., *Studies on repository compound stability in DMSO under various conditions.* J Biomol Screen, 2003. **8**(3): p. 292-304.
175. J., V.P.S.C.G., *Highly fluorinated amphiphiles as drug and gene carrier and delivery systems.* Journal of Fluorine Chemistry, 2001. **107**(2): p. 337-354(18).
176. Gaucheron, J., et al., *In vitro cationic lipid-mediated gene delivery with fluorinated glycerophosphoethanolamine helper lipids.* Bioconjug Chem, 2001. **12**(6): p. 949-63.
177. Finzel, B.C., et al., *Structural characterizations of nonpeptidic thiadiazole inhibitors of matrix metalloproteinases reveal the basis for stromelysin selectivity.* Protein Sci, 1998. **7**(10): p. 2118-26.
178. Huai, Q., et al., *Enantiomer discrimination illustrated by the high resolution crystal structures of type 4 phosphodiesterase.* J Med Chem, 2006. **49**(6): p. 1867-73.
179. Jacobsen, E.J., et al., *Synthesis of a series of stromelysin-selective thiadiazole urea matrix metalloproteinase inhibitors.* J Med Chem, 1999. **42**(9): p. 1525-36.
180. Esser, C.K., et al., *Inhibition of stromelysin-1 (MMP-3) by P1'-biphenylylethyl carboxyalkyl dipeptides.* J Med Chem, 1997. **40**(6): p. 1026-40.
181. *Nomenclature committee of the international union of biochemistry and molecular biology (NC-IUBMB), Enzyme Supplement 5 (1999).* Eur J Biochem, 1999. **264**(2): p. 610-50.

182. Veltman, O.R., et al., *A single calcium binding site is crucial for the calcium-dependent thermal stability of thermolysin-like proteases.* Biochemistry, 1998. **37**(15): p. 5312-9.
183. Fassina, G., et al., *Autolysis of thermolysin. Isolation and characterization of a folded three-fragment complex.* Eur J Biochem, 1986. **156**(2): p. 221-8.