SYNOPSIS

Understanding the structural differences among M1-class aminopeptidases and development of novel inhibitors

Introduction

Proteins or enzymes are important biological machinery essential for the normal and healthy growth of every living cell. *In vivo*, every protein or enzyme has its specific function. Once the biological role of a particular enzyme is fulfilled, it undergoes proteolytic degradation and forms individual amino acids. These free amino acids are again recycled into a functional protein whenever required. In nature, different kingdoms of life have their own proteolytic machinery for protein degradation. But unique machinery is present among all kingdoms of life in the form of endopeptidases and exopeptidases which are essential in the final stages of the protein degradation pathway.

Aminopeptidases catalyze the hydrolysis of amino acids from the amino terminus of the small peptides. M1 family aminopeptidases play important role in the pathogenicity of bacteria and shown to be critical in many physiological processes such as regulation of peptide hormone levels, protein maturation, cell cycle regulation, memory, antigen presentation, and angiogenesis in humans. They are mono zinc-containing enzymes which are involved in the release of free amino acids from smaller peptides. Except for viruses these enzymes are present in all kingdoms of life. To date, there are twelve members of the M1 class represent in humans. Some of them are membrane-bound and others are found in the cytosol. Most of the M1 class aminopeptidases are reported to be drug targets in various diseases in addition to having aminopeptidase activity. Due to the essential roles played by these enzymes, much emphasis was given to their biochemical and structural characterization. In addition, M1 class proteases from the archaea and bacterial origin are also structurally characterized. These are F3 from *Thermoplasma acidophilum*, aminopeptidase N from *Neisseria Meningitides* (*Nm*APN), aminopeptidase N from *Plasmodium falciparum* (*Pf*A-M1), and aminopeptidase N from *Escherichia coli* (*e*PepN).

All the M1 class proteases exist as multi-domain proteins. Domains arrangement around the catalytic domain, zinc-binding motif (HEXXH motif), and exopeptidase motif (GXMEN motif) are conserved among M1 class proteases. All the members of

this family have the same catalytic principles. They all have unique amino terminus recognition pattern where the carboxy terminus of three glutamates co-ordinates with the zinc. Specificity was derived due to the natural variation in residues around the substrate-binding pocket or S1 pocket. All the M1 class proteases consist of an archetypical S1 pocket including *e*PepN. In the present study, *e*PepN has been used as a prototype in understanding the biochemical properties and structural basis of M1-class aminopeptidases. In this protein, the exopeptidase motif (GXMEN motif) was preceded by methionine (M260) which acts as a gatekeeper in the S1 pocket and accommodates the peptides with a diverse range of residues at the P1 position.

The work executed in this thesis entitled "Understanding the structural differences among M1-class aminopeptidases and development of novel inhibitors" was undertaken with an aim of understanding the role of specific residues on the activity, stability, substrate specificity, and protein-ligand interactions of M1 class proteases. In this study, we worked on two proteases from this family namely, aminopeptidase A from *Legionella pneumophila* (*Le*PepA) and aminopeptidase N from *Plasmodium falciparum* (*Pf*A-M1). We carried out site-directed mutagenesis on specific residues of *Le*PepA and different mutants were generated (*Le*PepA-K132E, *Le*PepA-K132L, and *Le*PepA-S276M). Their biochemical properties, kinetic parameters were determined and the structural basis was provided for the unique substrate specificity of *Le*PepA. In this study, we also have designed a library of hydroxamic acid-based leucine ureido derivatives to discover molecules that can specifically inhibit *Pf*A-M1 with a novel mode of binding. By analyzing the structural data obtained from different *Pf*A-M1-inhibitor complexes, we synthesized an inhibitor molecule that has a high degree of selectivity towards *Pf*A-M1 over its mammalian counterpart porcine aminopeptidase N (*Pm*APN).

This thesis has been divided into four chapters. Chapter 1 discusses the introduction, conservation pattern, biochemical, structural information, and pathophysiological roles of different M1 class proteases from humans. It provides the information about the S1 pocket of M1 class aminopeptidases. Chapter 2 describes the materials and methods employed to execute this work. Chapter 3 and chapter 4, each chapter is subdivided into three sections. Section 1 is an introduction, section 2 results and discussion, and section 3 is the conclusion. Chapter 3 discusses "Discovery,

biochemical, mechanistic and structural studies of a unique Glu/Asp specific M1 class aminopeptidase from *Legionella pneumophila*". Chapter 4 deals with "Design and synthesis of *Pf*A-M1 selective inhibitors over mammalian counterpart as an antimalarial agent and structural evaluation".

Statement of problem

Most of the M1 family aminopeptidases studied till date are displayed broad substrate specificity, mostly specific to hydrophobic and basic residues. This is explained by the presence of hydrophobic S1 pocket in the active site of *e*PepN along with polar residues at the rooftop of the active site. Methionine at gatekeeper position (M260) followed by methionine in the exopeptidase motif (GXM263EN) contributes to the hydrophobicity of the S1 pocket in *e*PepN. In the case of human aminopeptidase A (*h*APA) with additional substrate specificity towards acidic residues, a polar residue (T356) occupied the gatekeeper position and diluted the hydrophobicity of the S1 pocket. From this observation, we anticipated that if there could be a member from the M1 class aminopeptidase that has both the methionine residues, one at gatekeeper position (M260) and the other in the exopeptidase motif (M263) are substituted by polar residues, and may have higher substrate specificity towards acidic residues. To find out the M1-class aminopeptidase with two polar residues at both gatekeeper and exopeptidase methionine positions, we performed an unbiased search, by providing query sequence "X-G-X-X-E-X (2,100)-H-E-X (2)-H-X (18)-E" (Where X represents variable amino acid) which includes signature motifs (exopeptidase motif and zinc-binding motif) of M1 class against the PROSITE database using the ScanProsite tool. 100 sequences were identified out of ~10,000 hits to have polar residues (Z-G-X-Z-E-X, Where Z represents the polar residue) at both the methionine positions. 25 of these 100 sequences consists of serine at both gatekeeper (S273) and exopeptidase (GAS276EP) positions. Interestingly we found that all these 25 sequences belonged to only *Legionella* species.

In this study, to meet the first objective, we focused on the biochemical and structural properties of aminopeptidase A from the extensively studied pathogen *Legionella pneumophila* (*Le*PepA). In this protein, the gatekeeper residue (S273) and the residue in the exopeptidase motif (GAS276EP) are polar and the rooftop of the active site is occupied by positively charged (K132 and R375) residues in contrast to

hydrophobic residues at gatekeeper position (M260) and exopeptidase motif (GAM263EN) and polar but not charged residues at the rooftop of *e*PepN active site. Due to the increased polarity of the wall of the cylinder and positively charged rooftop of the S1 pocket, we speculated that *Le*PepA would have different substrate specificity, preferably toward acidic residues in contrast to basic and hydrophobic residues in the *e*PepN. *Le*PepA may have evolved to fulfill the specific functions in *Legionella* species. Recently published report on crystal structure of *h*APA revealed that a polar residue (T356) at the gatekeeper position is involved in stabilization of the substrate with glutamate at P1 position by hydrogen bonds and also positively charged residue (R887) at the rooftop of the active site forms a salt bridge with the glutamate at P1 position. These observations strongly support our reasoning of unique specificity of *Le*PepA towards only acidic residues. To confirm our predictions and for further studies we have cloned this protein, purified, and crystallized.

The second objective of this study is to develop selective inhibitors against *Pf*A-M1. *Plasmodium* parasites causing malaria has developed resistance to most of the antimalarials in use, including the artemisinin-based combinations, which are the last line of defense against malaria. This necessitates for the discovery of new targets and the development of new antimalarials. Aminopeptidase N from *Plasmodium falciparum* (*Pf*A-M1), an M1 class aminopeptidase, is essential for the asexual erythrocytic stage of development. This enzyme has been validated as a potential antimalarial drug target. Herein we describe the design, synthesis, in vitro and cell based evaluation of a library of compounds with leucine hydroxamic acid attached to urea-substituted groups. The design is to extend the binding of molecules beyond S2' pocket of the enzyme to achieve the much-needed specificity. Most of the 40 compounds described in this study displayed inhibition at sub-micromolar range against the recombinant *Pf*A-M1 with negligible cytotoxicity on human cell lines and strong inhibition of the malarial parasite growth. However, the selectivity index between the malarial and mammalian M1 aminopeptidases for all these compounds is very less except for compound **39**, which has an inclination towards parasite enzyme, by more than 700 fold. Crystal structures of *Pf*A-M1 at atomic resolution in complex with at least nine different compounds including compound **39** establish the structural basis for the observed inhibition. Binding of molecule **39** extends beyond the S2' pocket where differences between the

malarial and mammalian enzymes are apparent. Together, our data provide important insights for the rational and structure-based design of potent and selective inhibitors of *Pf*A-M1 that will likely lead to novel chemotherapeutics for the treatment of malaria.

Objectives

1. Discovery, biochemical, mechanistic and structural studies of a unique Glu/Asp specific M1 class aminopeptidase from *Legionella pneumophila.*
2. Design and synthesis of *Pf*A-M1 selective inhibitors over mammalian counterpart as an antimalarial agent and structural evaluation.

Methodologies used and sample results:

Discovery, biochemical, mechanistic and structural studies of a unique Glu/Asp specific M1 class aminopeptidase from *Legionella pneumophila* (*Le*PepA)

Activity and substrate specificity of LePepA:

The full-length gene of aminopeptidase A from *Legionella pneumophila* (*Le*PepA) was cloned, expressed and purified to homogeneity (900 amino acids protein). Activity assays were performed with 15 aminoacyl-*p*NAs and 4 aminoacyl-AMCs as substrates against the purified enzyme. Interestingly *Le*PepA preferred only acidic residues as we predicted from our earlier bioinformatics studies.

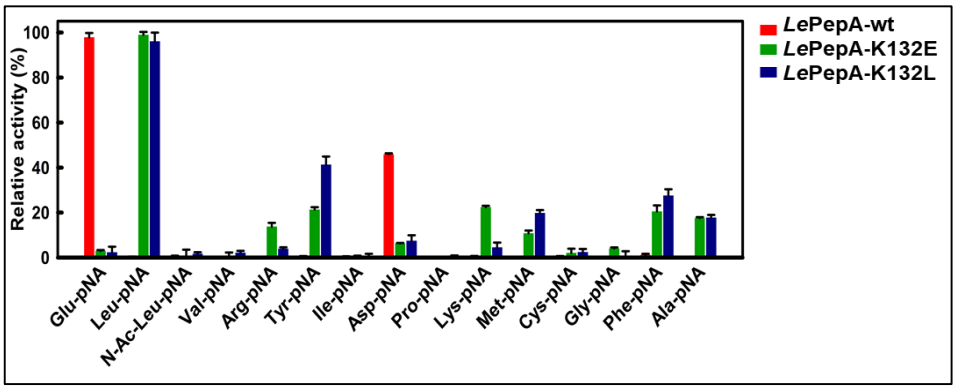

Figure. Enzymatic activity of *Le*PepA with different substrates.

The activity of *Le*PepA against the Glu-*p*NA was set to 100 % and compared relative activities of other substrates. Results presented are the means ± standard deviation of three independent measurements.

Residues in the S1 pocket and rooftop regions are critical for unique specificity toward acidic residues:

To provide the structural basis for the biochemical properties of the *Le*PepA, we have crystallized the enzyme and solved the crystal structures in the apo form and in complex with glutamate and aspartate. The overall structure of the *Le*PepA is folded into four-domains with thermolysin like catalytic domain. Similar to *e*PepN structure, *Le*PepA adopted the closed conformation. A single zinc ion is observed in the active site of the apo structure with conserved H310EFFHX(18)E333 zinc-binding motif. Strong electron density was observed for E311 and Y383 which are critical for peptide hydrolysis. E311 activates nucleophilic water and Y383 stabilizes oxyanion that results from the nucleophilic addition of water on the peptide bond. Representation of S1 pocket as a cylinder could be reconstructed for the *Le*PepA. Two serine residues present in the S273GAS276EP motif are locked by a weak hydrogen bond (3.8 Å). K132 and R375 occupied in the rooftop of the S1 pocket and positioned to interact with the incoming glutamate or aspartate containing substrates. K132 involved in hydrogen bonding with the main chain carbonyl of the E134 (2.6 Å) and S276 (3.1 Å). Similar to *e*PepN, three glutamates (E134, E277, and E333) involved in the formation of an anion hole that would accommodate the α-amino group of N-termini of the peptide substrate.

Design and synthesis of *Pf*A-M1 selective inhibitors over mammalian counterpart as antimalarial agents and structural evaluation

Library screening against PfA-M1 and PmAPN:

Fresh stocks of all 40 compounds and control molecule actinonin were made in the dimethyl sulfoxide (DMSO) to a stock concentration of 10 mM. IC_{50} was determined against *Pf*A-M1, *Pm*APN using seven different concentrations of inhibitor in the range between 0.016 µM and 100 µM at 200 µM of substrate (Leu-*p*NA) concentration. IC_{50} data was determined using SigmaPlot 13. The K_i value for all compounds against *Pf*A-M1 was determined by the Dixon method using Leu-*p*NA as a substrate. The assay consists of reaction buffer, enzyme (0.1 µM) and Leu-*p*NA (30 µM, 80 µM, 200 µM, and 340 µM concentrations). At each substrate concentration, inhibitor concentration

was varied in an appropriate range. Data were plotted as 1/rate vs. inhibitor concentration for each substrate concentration and a linear fit was calculated by non-linear regression using SigmaPlot 13. All the experiments were carried out in triplicates and standard deviation (S.D) values were reported. All inhibitors tested were exhibited competitive inhibition.

Determination of inhibition constants (K_i) against *Pf*A-M1:

Compound	K_i against *Pf*A-M1 (μM±S.D)
1	0.58±0.03
2	0.42±0.02
3	0.44±0.02
4	0.35±0.01
5	>500
6	0.74±0.04
7	0.87±0.04
8	0.44±0.02
9	0.14±0.01
10	0.22±0.01
11	0.24±0.03
12	0.43±0.05
13	0.26±0.01
14	0.53±0.02
15	0.09±0.008
16	0.18±0.01
17	0.24±0.01
18	0.28±0.01
19	0.09±0.006
20	0.40±0.02
21	0.16±0.02
22	0.13±0.007
23	0.19±0.01
24	0.99±0.05
25	1.06±0.05
26	1.60±0.10
27	2.61±0.13
28	0.65±0.04
29	0.76±0.04
30	0.15±0.01
31	0.84±0.05
32	59.15±4.82
33	113.04±11.86
34	112.42±5.25
35	209.66±12.91
36	125.47±6.72
37	75.11±4.98
38	119.20±15.23
39	0.05±0.005
40	66.31±9.70
Actinonin	9.49±0.69
Bestatin	0.86±0.06

K_i values were determined using Dixon plots of 1/rate versus inhibitor concentration for each substrate concentration. Substrate concentration was maintained lower than that of the K_m of the enzyme.