Investigations of Bacterial Methionine Aminopeptidase

by

Ronald Zahoruk

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

In

Chemistry

Waterloo, Ontario, Canada, 2009

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

ABSTRACT

The pathway representing methionine integration and excision is an increasingly important target in drug design. Methionine aminopeptidase (MetAP), a metalloprotease responsible for cleaving the N-terminal methionine from nascent peptides, has been the object of many studies aimed to produce potential anti-bacterial, anti-fungal and anti-angiogenic agents. Though clinical trials are underway for several of these compounds, like fumagillin and CKD-731, they are still flawed based on their relatively weak inhibition and their physiological side effects. Therefore, the search for novel and potent inhibitors continues. Previous work has utilized phosphinic and phosphonic acid derivatives of methionine in co-crystallization studies with *Escherichia coli* MetAP (eMetAP). The aim of the research presented in this work is to study and assay various phosphorus- and sulfur-containing compounds as inhibitors and substrates in an effort to learn more about the biochemical machinery underlying MetAP catalysis. As well, we outline a predictive molecular modeling approach to MetAP inhibitor design to assist in identifying lead candidates amongst a body of possible molecular inhibitors. Ultimately, we not only hope to have identified key functional properties of molecules potentially useful as MetAP inhibitors, but also to have contributed to the knowledge base of the mechanistic features involved in this enzyme's catalysis.

Acknowledgements

I would like to express my sincere gratitude to all involved in making this thesis and research behind it possible.

I would like to thank Cullen Myers, Uthaiwan Suttisansanee, Ignace Moya, Meijun Lu, Pei Hang, Lisa Pokrajac, Dr. Zhending Su, Dr. Elizabeth Daub and everyone else who so willingly offered me their assistance and expertise in solving any problem I encountered during my Master's.

Many thanks also to my committee members, Dr. G. Dmitrienko, and Dr. T. Dieckmann for their guidance and suggestions.

The technical assistance from Jan Venne and Dr. R. Smith is also greatly appreciated.

Cathy Van Esch also deserves many thanks for her always gracious and warm manner in handling any issue that came up during my Master's as well.

Most of all, I offer my most sincere gratitude and thanks to my supervisor, Dr. John Honek. Throughout the years, he constantly offered his advice, guidance, knowledge, and friendship generously and selflessly. It is because of him that I came to the University of Waterloo in the first place, and it is because of him I was able to undertake the research that I was interested in. I am forever indebted to him, and sincerely hope other graduate students in the future are as fortunate as I had been with him as my supervisor.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

3D	Three-dimensional
4NP	4-Nitrophenol-O-phosphate
AMC	7-amido-4-methyl coumarin
AMP	Adenosine-5'-monophosphate
ATP	Adenosine triphosphate
BNPP	bis(nitrophenyl)phosphate
bstMetAP	*Bacillus stearothermophilus*
CDK4	Cell-dependant kinase 4
DDQ	Dichloro-5,6-dicyanobenzoquinone
DFM	Difluoromethionine
DFM-AMC	Difluoromethionine-7-amido-4-methyl coumarin
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DPA	2,6-pyridinedicarboxylic acid
DPM	(diphenylmethyl)amine
E. coli	*Escherichia coli*
EDTA	Diaminoethanetetraacetic acid
eIF2	*E. Coli* translation initiation factor 2
eMetAP	*Escherichia coli* methionine aminopeptidase
EPR	Electron paramagnetic resonance
fTHF	N-10-formyltetrahydrofolate
FMOC-DFM-AMC	9-fluorenylmethyl-difluoromethionine-7-amido-4-methyl coumarin

FMOC-TFM-AMC	9-fluorenylmethyl-trifluoromethionine-7-amido-4-methyl Coumarin
HCl	Hydrochloric acid
HIV	Human immunodeficiency virus
hMetAP	Human methionine aminopeptidase
HSA	Human serum albumin
HPLC	High Performance Liquid Chromatography
IC_{50}	Half-maximal inhibitory concentration
LEP	Lowest energy pose
MetAP	Methionine aminopeptidase
MetRS	Methionyl-tRNA synthetase
Met-AMC	Met-7-amino-4-methylcoumarin
MTF	Methionyl-tRNA transformylase
mtMetAP	*Mycobacterium tuberculosis* methionine aminopeptidase
NME	N-terminal methionine excision
NNRTI	Non-nucleoside reverse transcriptase inhibitor
PDF	Peptide deformylase
pfMetAP	*Pyrococcus furiosus* methionine aminopeptidase
pkMetAP	Porcine kidney MetAP
PMSR	Peptide methionine sulfoxide reductase
pNP	para-Nitrophenol
ProAP	Prolyl Iminopeptidase
PP_i	Pyrophosphate

PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
R.F.U.	Relative fluorescence unit
RMSD	Root mean square deviation
scMetAP	*Saccharomyces cerevisiae* methionine aminopeptidase
siRNA	Small interfering ribonucleic acid
TFM	Trifluoromethionine
TFM-AMC	Trifluoromethionine-7-amido-4-methyl coumarin
THF	Tetrahydrofolate
TLC	Thin layer chromatography
tRNA	Transfer ribonucleic acid
uM	Micromolar
uV	Ultraviolet visible
VdW	Van Der Waal's
vHTS	Virtual high-throughput screening
yMetAP	Yeast methionine amino peptidase

All amino acids in this document are referred to by their standard one (1) or three (3) letter code unless otherwise stated

LIST OF SCHEMES

Chapter 1: Introduction

The landmark discovery of the antiobiotic penicillin has classically been regarded as one of the high points of the modern scientific era. However, in recent times, this discovery's clinical importance has become threatened with the ever increasing frequency of clinically observed drug resistance (1 - 5). Therefore, researchers are extensively laboring to develop new antibiotics, and with this, are looking for novel targets that will surmount the means of resistance developed by Gram-negative and Gram-positive bacteria against current anti-bacterial agents (6). One such line of inquiry has focused on enzymes comprising the biochemical machinery responsible for cellular protein biosynthesis. Specifically, over the last 20 years, the enzymes responsible for the incorporation and removal of the N-terminal methionine in protein biosynthesis have become increasingly important targets.

Methionine incorporation and removal utilizes the following pathway [Figure 1] (6):

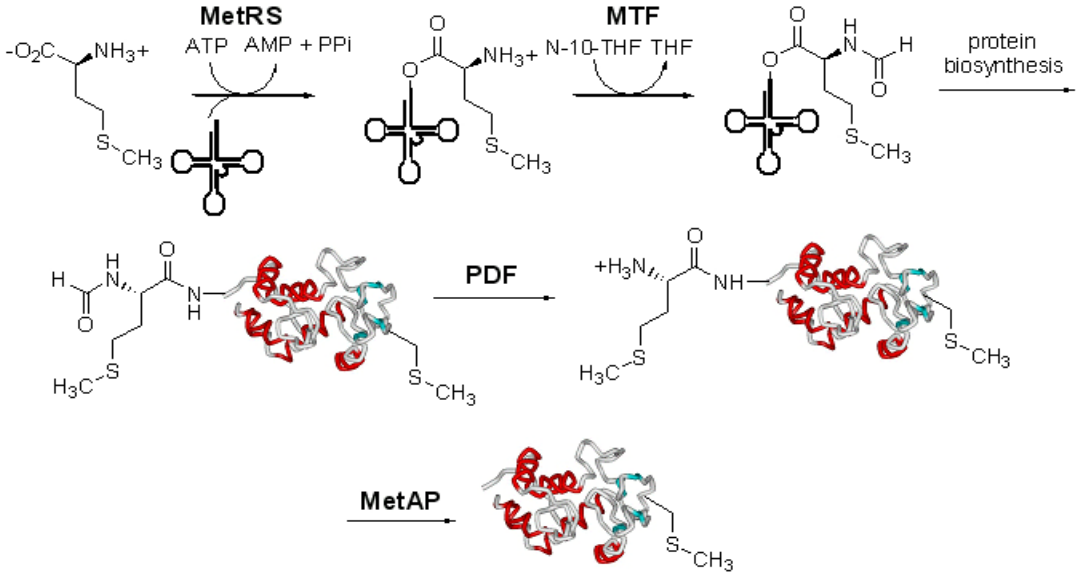


Figure 1. The pathway involving different steps and enzymes in methionine integration and excision into and from proteins. See text for abbreviations.
<Adapted from reference (6)>

The following is a brief description of the steps of the pathway:

(1) Methionyl-tRNA Synthetase (MetRS) catalyzes two reactions; first, the activation of methionine by condensing it with adenosine triphosphate (ATP) to form an aminoacyladenylate Met-AMP with concomitant release of pyrophosphate (PPi); then, the activated Met is transferred to either $tRNA_f^{Met}$ (initiator) or $tRNA_m^{Met}$ (elongator) through the creation of an ester bond to the 2'-hydroxyl group of the terminal adenosine of $tRNA^{Met}$ [Scheme 1] (6, 7);

$$E + Met + ATP \rightleftharpoons E\text{-}Met\text{-}AMP + PPi$$
$$E\text{-}Met\text{-}AMP + tRNA^{Met} \rightleftharpoons E\text{-}Met\text{-}AMP + tRNA^{Met}$$

Scheme 1

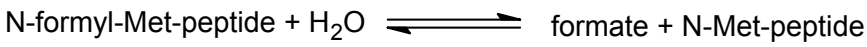
(2) In bacteria and in mitochondria, Methionyl-tRNA transformylase (MTF) then catalyzes the formylation reaction of the methionine that is esterified to the initiator methionyl-tRNA using *N*-10-formyltetrahydrofolate (fTHF), producing *N*-formylmethionyl-tRNA ($fMet\text{-}tRNA_f^{Met}$) [Scheme 2] (6).

$$Met\text{-}tRNA_f^{Met} + fTHF \rightleftharpoons fMet\text{-}tRNA_f^{Met} + THF$$

Scheme 2

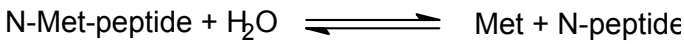
(3) At this point, protein biosynthesis occurs, utilizing $fMettRNA_f^{Met}$ (in bacteria) or $Met\text{-}tRNA_i^{Met}$ and $Met\text{-}tRNA_e^{Met}$ (in eukaryotes) (6).

(4) Peptide deformylase (PDF) removes the N-terminal formyl group from various nascent proteins that contain N-formylmethionine [Scheme 3] (8).



Scheme 3

(5) At this point, Methionine Aminopeptidase (MetAP) becomes involved and, assuming PDF has acted upon the protein in the previous step, removes the now non-formylated N-terminal methionine in some proteins [Scheme 4] (6, 9, 10).



Scheme 4

MetAP is of particular interest because its activity as an intracellular enzyme, co-translationally cleaving off the N-terminal methionine from nascent peptide chains, is integral to various biological processes (11, 12, 13). Its primary function in the cell, referred to in the literature as N-terminal methionine excision (NME), is thought to be the primary means of generating N-terminal amino acid diversity, as a consequence of revealing the various residues in the next position after methionine, with 80% of proteins undergoing this type of modification in *E. coli* (12, 13). Accordingly, MetAP is found wherever protein biosynthesis occurs, namely in the cytoplasm, mitochondria and plastids (11). MetAP is part of a larger family of enzymes called the “Pita Bread Fold” enzymes, a unique group of metalloproteases that contain a distinctive topology proposed to be responsible for their substrate specificity towards N-terminal peptide bonds (14). Amino-terminus removal (hence the name ‘amino – peptidase’, a type of N-terminal exopeptidase) of methionine is believed to be required for proper protein functioning, subcellular localization and future degradation (12, 14).

1.1. Structural Features

MetAP is divided into two classes: Type I, found in eubacteria, and Type II, which is found in both humans and archaea along with Type I (6, 12, 14, 15). X-ray crystallography has allowed a significant amount of structural information to be obtained on the various types of MetAP. For example, despite the fact that MetAP I and II exhibit low sequence homology, a similar protein fold is exhibited by both enzymes (14, 15). The 'pita bread fold', the namesake of MetAP's family and found in MetAP enzymes, has pseudo-twofold symmetry between the N- and C- termini, each of which contains two alpha-helices and two antiparallel beta-sheets [Figure 2] (14, 15). The active sites of both Type I and Type II are found near the core of a central beta-sheet (16). The primary difference between them is a 12 kDa peptide insertion 60 residues in length that forms a large helical domain in the C-terminal region of the Type II form (12, 14). The other major difference is the finding that eukaryotic Type I MetAP enzymes contain two distinct zinc finger motifs, which is in contrast to the Type II form which possesses a highly charged amino-terminus comprised of alternating polybasic and polyacidic residues (16, 17).

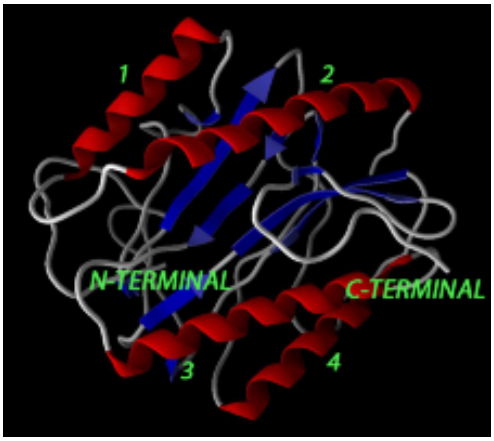


Figure 2. Three-dimensional structure of *Escherichia coli* MetAP revealing 'Pita Bread Fold' symmetry. The α-helices are numbered from N to C terminus [PDB code: 2MAT]
<Adapted from reference (12)>

1.2. Active Site

The active site of MetAP is contained within the pseudo fold (12). Both the N- and C-terminal domains contribute conserved residues to form the dinuclear metal centre here as well, where the metal ions are coordinated by either two monodentate ligands, namely His-171 and Glu-204 (*Escherichia coli* MetAP (eMetAP) numbering) and by three bidentate ligands Asp-97, Asp-108 and Glu-235 (16, 18). As well, crystallography has shown that various solvent molecules are present in this area, including a metal-bridging hydroxide ion or water molecule which is believed to act as a nucleophile (to be discussed later) (12, 19). The binding pockets where the substrate is found during catalysis are also located in and around the active site. The S_1 subsite (where the P_1 residue of an inhibitor or substrate binds) is based on residues primarily from the N-terminal domain, despite the apparent pseudo two-fold symmetry exhibited by MetAP (nomenclature as defined by Schechter and Berger) [Figure 3] (19, 20). More specifically,

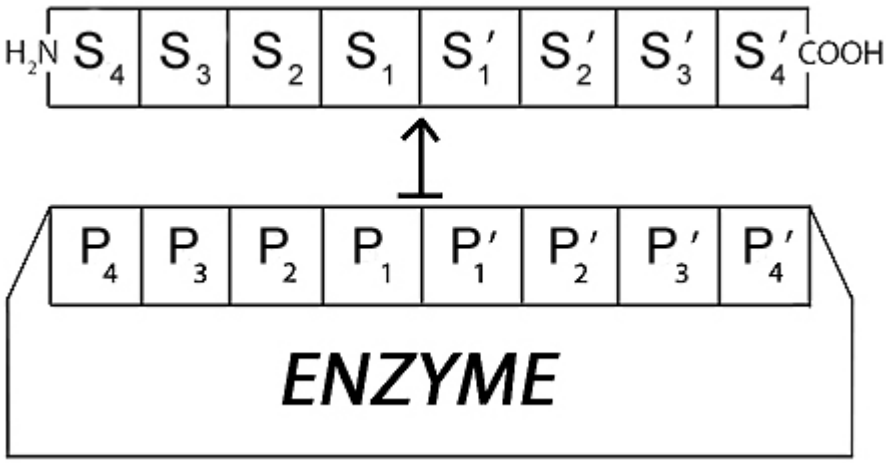


Figure 3. Schematic representation of an octapeptide in contact with an enzyme. The arrow indicates the site of cleavage. Note the numbering of the substrate and enzyme residues. <*Adapted from reference (20)*>

this region is relatively hydrophobic, and consists of amino acid residues His-79, Cys-59, Cys-70, Tyr-62, Tyr-65, Phe-177 and Trp-221 [Figure 4] (19). However, the P_1' residue of the substrate binds in the S_1' subsite, and appears to be a shallow pocket comprised of the residues Glu-204, Gln-233, Met-206 and Tyr-168 [Figure 4] (19).

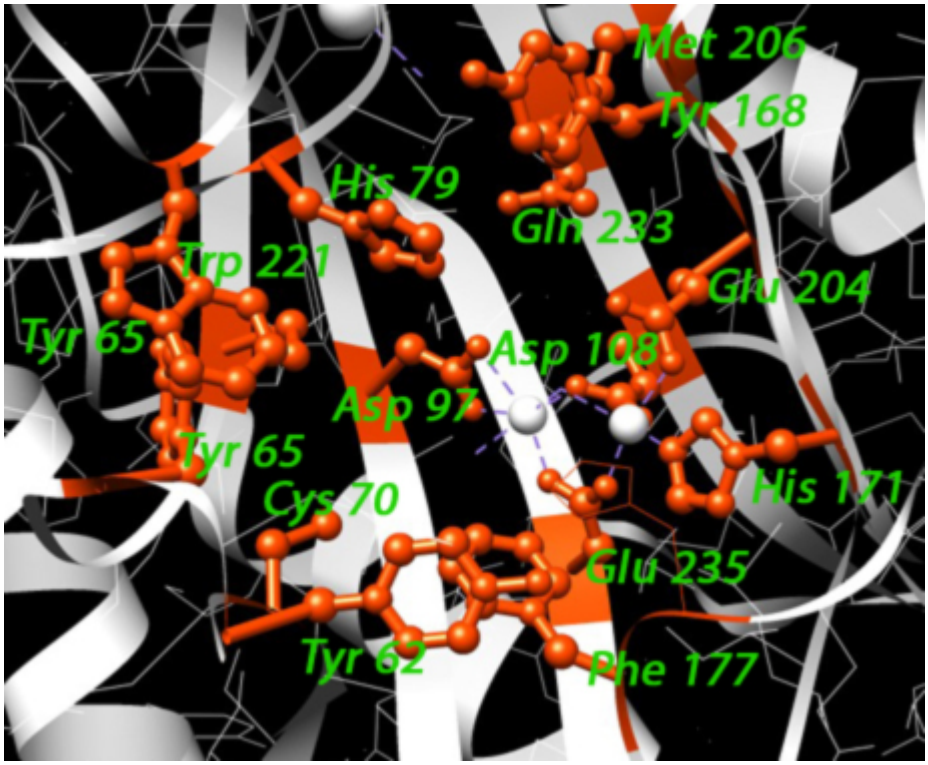


Figure 4. Active site of MetAP with important residues displayed [PDB code: 2MAT]
<Adapted from reference (20)>

Based on previous studies, His 79 is clearly an important amino acid in the active site of MetAP (6). However, its precise role has been debated, as besides its function during catalysis already mentioned, others propose that His 79 may form a hydrogen bond with the leaving scissile nitrogen atom of the peptide bond being hydrolyzed (19). Due to the funnel-like nature of the active site, such a bond may be required as the wide mouth at the opening of the active site soon becomes quite narrow deep within the cavity itself (6). It has been noted that the binding of phosphorus-based analogues of

methionine to eMetAP fosters hydrogen bond formation during catalysis at His79 (14). Further evidence from the structurally equivalent histidine residue in human MetAP (hMetAP) in complex with fumagillin, an epoxide-containing natural product, and several of its derivatives also implies these interactions (12).

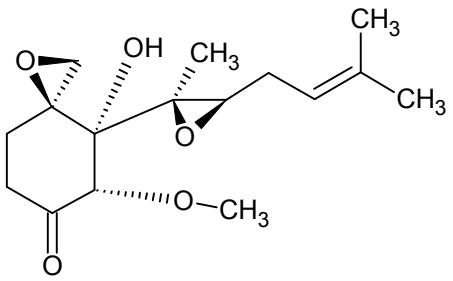


Figure 5. The anti-angiogenic compound, fumagilin.

Other residues in the active site of MetAP are important for catalysis. One such example is Asp-97 – its mutation to alanine, glutamate or asparagine effectively removes the dinuclear metal center of MetAP, based on the disruption of the metal-coordination required in the active site where Asp-97 plays a large role (21). Consequently, compared to wild-type enzyme, there is a vast reduction in enzymatic activity (21). On the other hand, altering certain residues in the active site modifies activity rather than eliminates it. This is seen in the case of His-63 which, when mutated to alanine, significantly diminishes the ability of MetAP to act upon small peptidic substrates. However, MetAP's ability to cleave the N-terminal Met of large peptides is largely unaffected, and at times, seems enhanced (22). It is believed that the loss of His-63 changes the orientation of regional residues and as a point of interest, additional hydrophobic and hydrophilic bonding interactions are created fostering a greater ability to process substrate (22).

1.3. Metal Co-Factors

The dimetal center is a distinguishing feature of MetAP enzymes [Figure 6]. The initial finding that treatment with ethylenediaminetetraacetic acis (EDTA) renders MetAP inactive uncovered the metal-dependent nature of MetAP (6). The dinuclear metal center in *E. coli* MetAP (eMetAP) is formed by the contribution of conserved residues from the N- and C- terminal domains (monodentate ligands, such as His-171 or bidentate ligands, such as Asp-97, Asp-108, Glu-235) [Figure 4]. A metal-bridging water or hydroxide ion is thought to act as the nucleophile during catalysis (19). Though Co^{2+} is historically thought to be the most common type of metal to be found in this region, Zn^{2+}, Mn^{2+}, and Ni^{2+} have also been documented to populate this region (6, 10, 12, 20, 22, 23). In addition, though Mn^{2+} and Co^{2+} maintain optimum enzymatic activity, changing the type of metal found in the dinuclear center does not eliminate activity, and in the case of yeast MetAP I (yMetAP I), Zn^{2+} actually enhances catalytic efficacy when substituted for Co^{2+} (20).

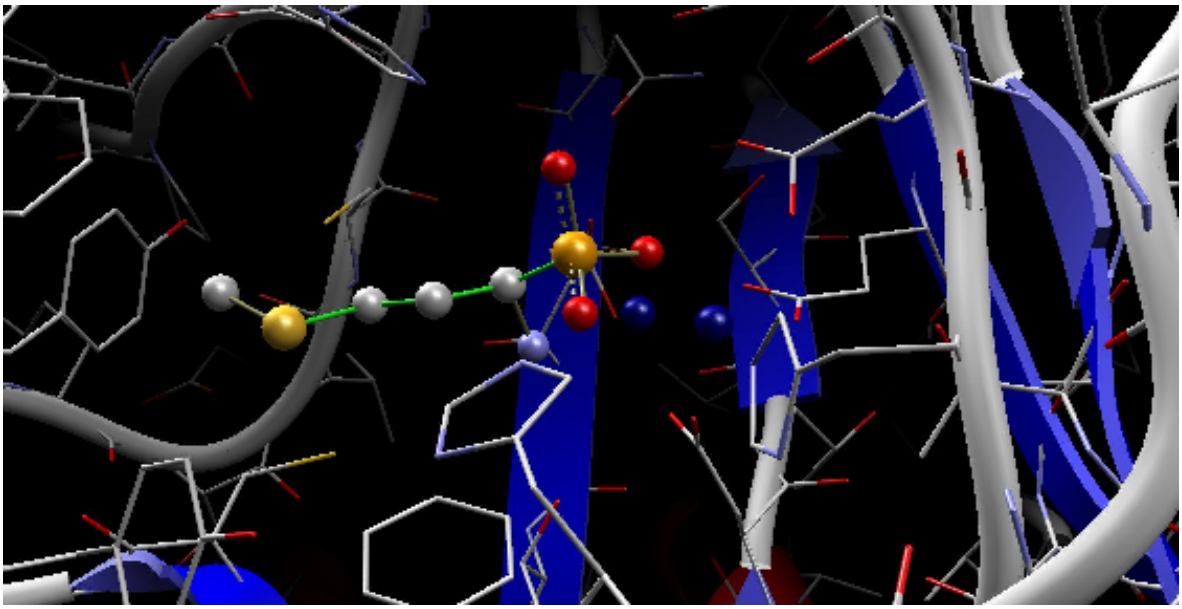


Figure 6. Three-dimensional representation of the dinuclear center of MetAP showing a phosphonate analogue of methionine bound in the active site. In dark blue are the metal ions [PDB code: ICS3] <*Adapted from reference (23)*>

This idea of metal selectivity is also quite pronounced in the search for efficient MetAP inhibitors. Several journal articles have been published that demonstrate specific types of inhibitors may be selective for different metalloforms of MetAP (24, 25, 26). For example, catechol inhibitors of MetAP containing 5- and 6-membered ring groups have been shown to inhibit MetAP much more potently in the presence of Fe(II) than other divalent metal ions commonly utilized in *in vitro* MetAP catalysis (25). This same theme permeates many different types of inhibitors (for example, Mn(II) as the metal of choice – *see Figure 7*), and its further elucidation provides insight not only into the active site of MetAP, but exactly what kind of metal selectivity, and under what conditions catalysis is optimal for this critical enzyme.

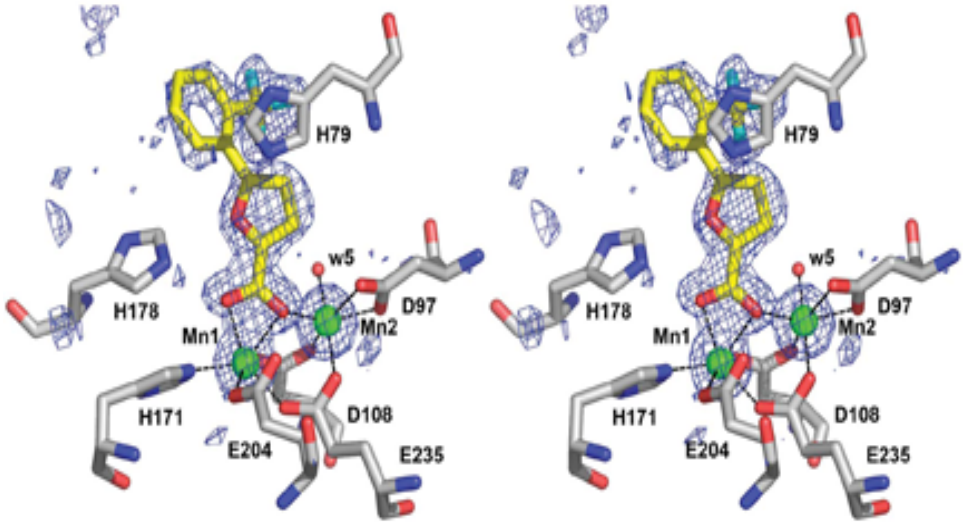


Figure 7. Stereoviews of eMetAP with a Mn(II)-selective inhibitor, demonstrating the Mn(II) coordination in the active site <*Taken from reference (26)*>

Interestingly, it has been noted that the presence of a single metal atom can allow MetAP to function fully (27). The proposed role of the second metal ion in catalysis is to bind and position the terminal amino group of the polypeptide chain. However, D'Souza and colleagues have posited that this role can in fact be carried out

by active-site amino acid groups Asp-97, Asp-108, and Glu-235, which form a negative pocket in the appropriate position to bind the N-terminal amine moiety (31). Therefore, they have suggested that the binding of the second metal ion may have no catalytic role but may play a regulatory one (31).

1.4. Physiological Role

The removal of the N-terminal methionine is a crucial stage in protein maturation (32). In fact, some proteins fail to function until this happens (12). However, this seems to occur only if the penultimate residue is small and uncharged, typically serine, threonine, valine, proline, glycine, alanine and cysteine, ultimately a function of the P_1' residue's radius of gyration [Figure 8] (6, 12, 14). It appears that these residues influence whether the initiator methionine is cleaved, controlling MetAP specificity (6). Incidentally, this substrate preference is antithetical to the 'N-end rule' for protein degradation, and MetAP will fail to cleave the N-amide bond should a 'destabilizing' residue be present (23).

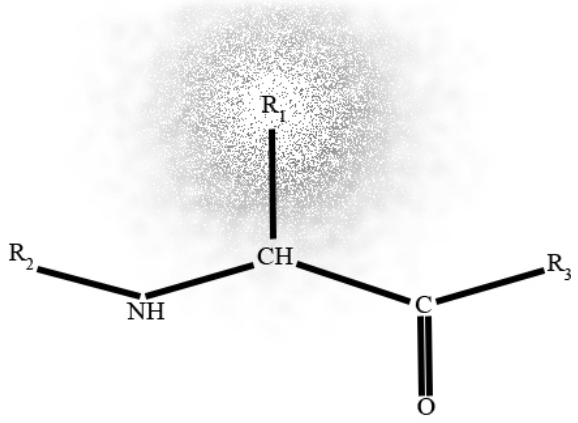
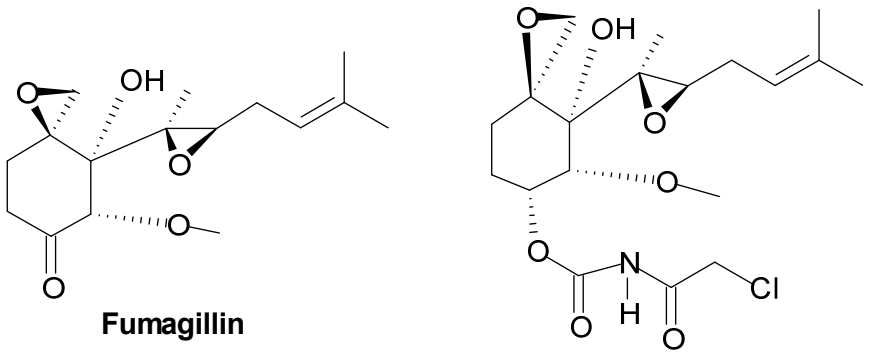


Figure 8. A representation of the radius of gyration of an amino acid side chain. The clouded region indicates the space occupied by the mobility of R_1 on account of bond rotation.

MetAP's importance biologically is underscored by the lethality of deleting or inhibiting MetAP genes or gene products in *Saccharomyces cerevisiae*, *Salmonella*

typhimurium and *E. coli* (32, 33, 34). More specifically, it has been shown that the N-terminal domain extensions of some MetAPs may be important for targeting the enzyme to the ribosome or other cellular partners via protein–protein, protein–DNA or protein–RNA interactions (14). For instance, a slow-growth phenotype has been observed for the point and deletion mutants of the zinc finger domains of yMetAP (35).

Accompanying this is the quite fascinating speculation that the N-terminal domain of the p67 protein, an hMetAP homolog, seems to promote protein synthesis by phosphorylating the α-subunit of translation initiation factor eIF2, thereby protecting it (36, 37, 38). At the level of the organism, such as Drosophillia, this results in eye and wing vein defects and an aberrant distribution of MetAP II mRNA over the course of development, suggesting a mechanism whereby MetAP exerts control at the translational level, or possibly during post-translational processing (39). MetAP is also believed to be involved in the process of angiogenesis (6, 12, 14). A very interesting study done by one of the unfortunately now deceased pioneers of the angiogenesis field, Judah Folkman, showed how knocking out MetAP2 in transgenic murine models through targeted gene disrupted leads to gastrulic defects and ultimately, necrotic defects in mouse embryos (40). As a result, MetAP is an attractive target pharmacologically because of its inactivation by anti-angiogenesis agents, fumagillin and its synthetic analogue, TNP-470, and these have thus far proven effective in clinical trials as potential anti-tumour agents [Figure 9] (6).



Fumagillin **TNP-470**

Figure 9. The natural product Fumagillin and its derivative, TNP-470.

1.5. Substrate Specificity

Though eukaryotic organisms express two types of MetAP (MetAP I and MetAP II), they both exhibit similar substrate specificities (14). However, subtle differences exist in their substrate preference, demarcated by the efficiency with which they perform N-terminal methionine excision (NME), and in their inhibition by fumagillin. For example, the yeast MetAP I (yMetAP I) is unable to cleave certain substrates that yMetAP II can process, but is selectively inhibited by fumagillin over type I (15). In addition, though both MetAP I and MetAP II exhibit discrete substrate preferences, siRNA studies have shown that depletion of either enzyme causes human cell growth impediment (41).

As mentioned already, MetAP will only remove the N-terminal Met (P_1) when the second residue (P_1') is small and uncharged (ex. Gly, Ala, Ser, Thr, Val, Pro, Cys) and the peptide is at least three residues in length (10, 42). This seems to contradict the widely accepted 'N-end rule' for protein life span and degradation, where the lifetime of

a protein is determined by the identity of its N-terminal residue (43). On the contrary, MetAP will not function if the amino acid in the penultimate position of the substrate is considered 'destabilizing' (12). However, there are exceptions. This stochastic rule emerged out of early bioinformatics analysis based on a compilation of available protein sequences of the time (which were few), and was confirmed with only about a dozen model tri- and tetrapeptides (11). The authors of these studies suggested that the process was statistical rather than stochastic and that cleavage efficiency was correlated with side-chain length or gyration radius solely at P_1' (11, 44). However, recent evidence has emerged that suggests residues in positions P_1', P_2', P_3' and even P_4' also contribute to cleavage efficiency and affect cleavage probability, and it is not simply the nature of the penultimate position that solely determines selectivity [Figure 10] (11).

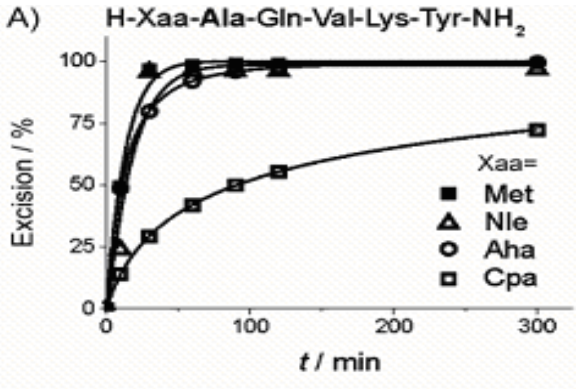


Figure 10. The speed of excision by MetAP depending on the penultimate residue with various substitutions in that position <*Taken from reference (45)*>

Interestingly, despite NME being a rapid and efficient process, the N-terminal methionine residue may erroneously be retained on some proteins that are recombinantly overproduced in *E. coli* (45). This has been especially true when strong promoters like T7 or T5 are used (45). This phenomenon is most probably due to the

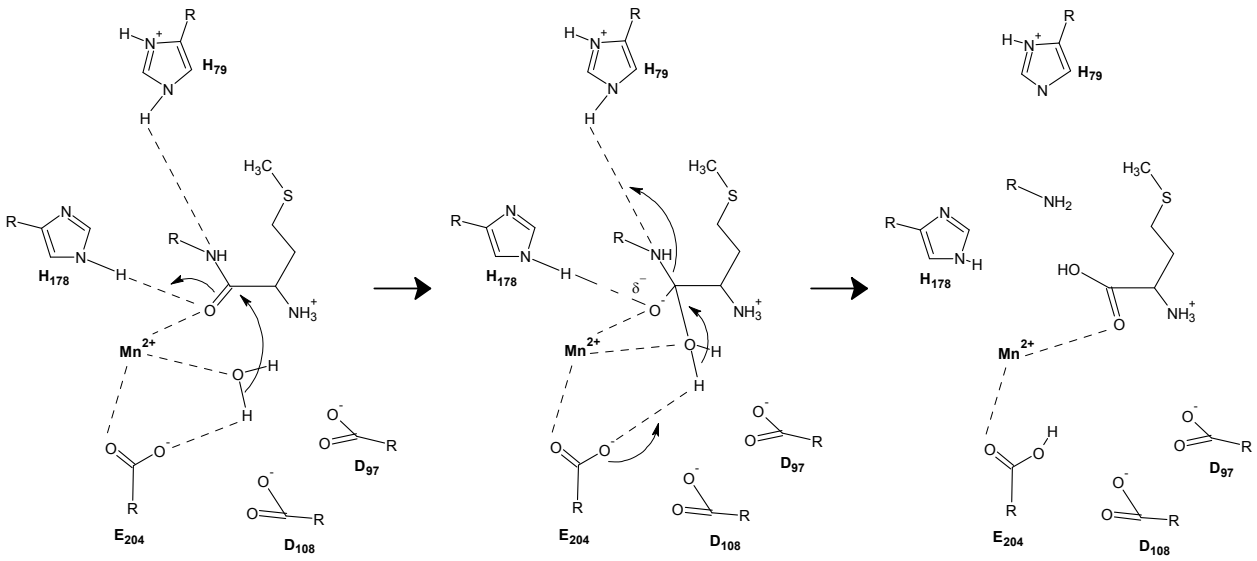
saturation of the peptide deformylase and MetAP capacities or targeting of the nascent protein, for example, into inclusion bodies (45).

1.6. Catalytic Mechanism

In spite of the dimetal centre that has been detected by X-ray crystallography within the active sites of various MetAP, and also bearing in mind that previous mechanisms proposed for MetAP make use of two metal ions, it is believed that MetAP requires only one metal ion to function, and therefore is based on a mono-metalated mechanism.

Mechanism [Scheme 5]:

Initially, there is nucleophilic attack by one of the active site water molecules on the N-terminus of the substrate. This is followed by a proton transfer from the μ-hydroxide to the carbonyl group of the substrate, and a subsequent proton abstraction from His-73. As well, Glutamate-204 abstracts the remaining proton of the carbonyl group of the substrate, all leading to a tetrahedral intermediate of the substrate. His-178 may facilitate the collapse of this intermediate, since it could act as a proton donor to the departing amino nitrogen. His-79 functions to orient the substrate for catalysis and the metal ion, in this case, Mn(II), acts in the role of coordination (19).



Scheme 5. The proposed catalytic mechanism of eMetAP. One should note that the water molecule ligated to the metal center may be present as a hydroxide. <*Adapted from reference (19)*>

1.7. Inhibitors of MetAP

To date, many synthetic compounds have been proposed to act as inhibitors of MetAP and demonstrate possible therapeutic value. These compounds include those containing keto heterocycles, aminoketones, alpha-hydroxy-beta-amino acids, triazoles, hydroxyamides, benzamides and sulfonamides, and each class has various benefits and drawbacks (46 - 51). For instance, Sheppard and colleagues designed and tested a very promising set of anthranilic acid sulfonamides, which *in vitro* inhibited MetAP II fairly strongly with an IC_{50} of 2.4 μM (50). However, the addition of human serum albumin (HSA) significantly diminished inhibition, presumably because of extensive binding of the inhibitor to HSA (99.7-99.8 % protein bound in rat plasma) (50).

Both Type I and Type II MetAP enzymes have been shown to be possible targets of several anti-fungal, anti-microbial and anti-angiogenic agents (6). However,

excitement surrounding MetAP inhibition truly began when it was found to be the specific target of the epoxide-containing anti-tumor natural product fumagillin (49). Its mode of action through irreversible covalent modification of MetAP resulted in a decrease of endothelial cell growth (52). These findings were soon followed by experimental results that implied these compounds impeded angiogenesis (53). Another compound dubbed 'TNP-470', a synthetic analogue of fumagillin, shows even higher efficacy in the therapy of specific types of cancer and is currently in clinical trials (54, 55). However, these compounds are still relatively weak inhibitors of MetAP and elicit deleterious side effects. Fumagillin, for example, has been observed to lead to severe weight loss with long-term administration (6). Therefore, the search for other inhibitors has continued. With this, many other types of compounds have been designed and tested against MetAP in the hopes of having strong inhibitory effects on MetAP.

Angiogenesis is the process by which blood vessels form and grow. During tumour metastasis, nutrients and oxygen are needed to sustain newly formed tumours and to enable their growth at new sites distal from the original site of tumor formation (6). Agents which specifically combat the formation of new blood vessels also decrease the formation of new metastatic tumors (6). As mentioned previously, MetAP is specifically inhibited by fumagillin and its synthetic analogues, such as TNP-470 and CKD-731, and results in decreased angiogenesis and endothelial cell growth upon their administration (6). As a point of interest, CKD-731 is seen to act against endothelial cell growth 1000 times more potently than fumagillin and it is thought this same affinity may also hold true in studies aimed at MetAP inhibition (56).

1.8. In Silico Medicinal Chemistry and Molecular Modeling

Medicinal chemistry has been around since the time of ancient Egypt (57). Their primitive but often innovative methods by which medicines were discovered has changed quite considerably since that era. In the 1960's, it became much more clear that a unique relationship existed between the structure of a chemical entity to be used medicinally, and the biological activity it induced (58). Since then, various strategies have been developed in order to discover therapeutic agents. Yet, few have seen as many of the technological advances in the past 20 years as computer-aided drug design and molecular modeling. In fact, by today's standards, it is almost inconceivable for the drug discovery process to even take place without the use of computational methodologies at some point during lines of inquiry in to therapeutics (59). Combining the knowledge and technology of computational chemistry, bioinformatics, conventional medicinal chemistry, biochemistry, and structural biology and chemistry, the successful application of computer-aided drug design and molecular modeling to medicinal chemistry has been a key approach to the design of medicinal agents. Moreover, the improved structural capabilities (highly efficient protein generation and purification techniques, high throughput crystallography, structure-activity relationship studies through advanced NMR methods, etc.), enhanced computational algorithms simultaneously with increasingly powerful and affordable computing power, along with the well developed high-throughput chemistry available (e.g. parallel synthesis, specific library design, etc.) have changed the face of contemporary medicinal chemistry and has undoubtedly contributed to the many successes involved thus far in the field (59).

In the last 20 years, the frequency with which molecular modeling studies have been incorporated into drug discovery-related publications has increased spectacularly [Figure 11]. For example, in 1999, 82 publications were available that looked at drug design using molecular modeling. A comparison of that figure to 10 years before, in 1989, shows that only 14 drug-design papers made use of molecular modeling. However, at the time of writing, half way through 2009, 1402 papers were published that specifically integrated molecular modeling into their drug design research. This most dramatic shift toward the use of computational methods in drug design is undeniable, and molecular modeling has clearly become an essential tool in the drug discovery process.

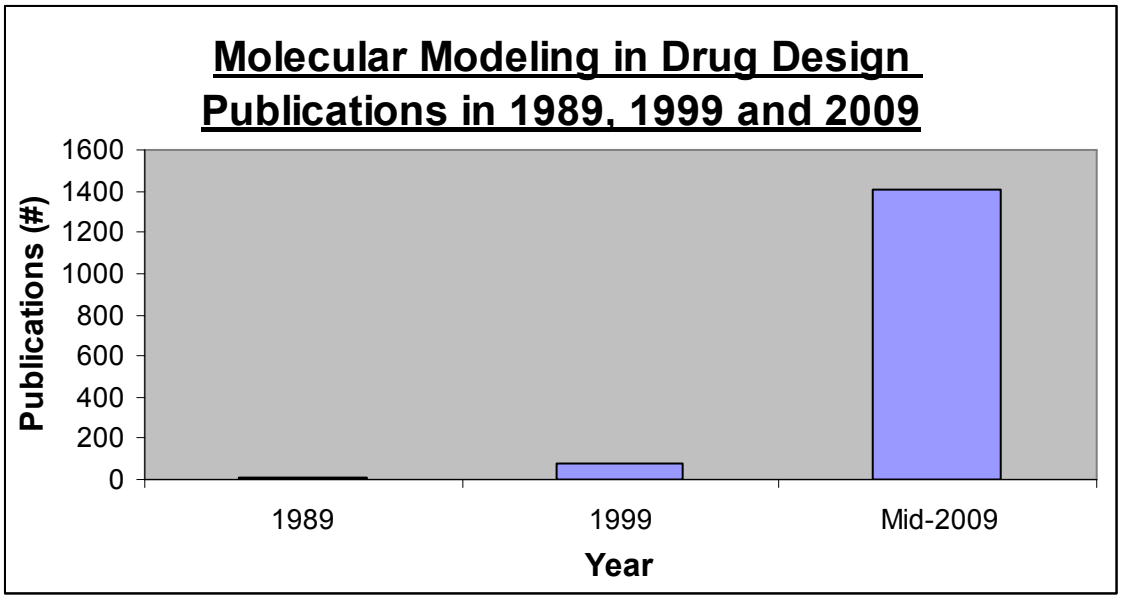


Figure 11. The number of publications in drug-design that use molecular modeling in 1989, 1999, and Mid-2009 searched on PubMed using the search terms "inhibitor" and "molecular modeling" as of August, 2009.

There has been very little work to date that studies methionine aminopeptidase using molecular modeling. That which does exist employs either a molecular mechanics view of MetAP (2 papers – *discussed later*), or uses molecular modeling in

the inhibitor design process (3 papers – *discussed later*). Considering the importance of MetAP as a potent anti-cancer target, and the number of publications that mention this while also focusing on MetAP inhibitor design, as well as the success seen using *in silico* methods in other areas of inhibitor design research like NNRTIs against HIV, it is quite surprising that more reports have not been published with MetAP [Figure 12]. Perhaps this will change if methionine aminopeptidase's importance in tumorgenesis becomes more widely known and accepted.

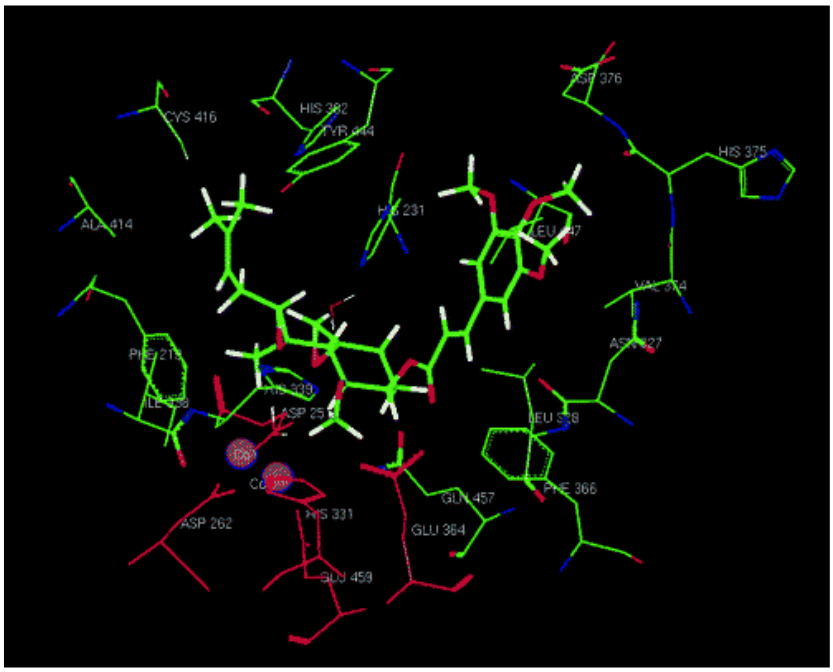


Figure 12. CKD-731, a potent analogue of fumagilin, docked *in silico* with eMet-AP2
<Taken from reference (60)>

1.9. Statement of Goal

There is great therapeutic value in understanding MetAP catalysis. Our aims with this research centre on probing the *Escherichia coli* MetAP through various means in order to get a better understanding of the enzyme's catalytic mechanism, active site,

substrate specificity, metal preferences, and inhibition. By using a fluorescence-based biochemical assay that will monitor enzyme activity, we hope to be able to assess the interactions that may exist between MetAP and various compounds possessing unique functional groups. We also intend to perform various types of molecular modeling in an effort to provide further insight into how to focus the drug design process in MetAP inhibitory studies, while also giving an idea of how the compounds we use to probe MetAP perform in this context comparatively to the assays we will perform. Taken together, pursuing these lines of inquiry will allow us to acquire a firmer grasp on MetAP and contribute to the knowledge that may one day lead to effective preventative therapeutics in MetAP-involved pathologies.

Chapter 2: Probing Methionine Aminopeptidase

2.1. Introduction

As touched upon in the introduction, MetAP has become a very important therapeutic target (6). Its involvement in the removal of nascent peptide N-terminal methionine residues is a vital component in the biosynthetic pathway, lethality resulting in its absence and functionally relevant in various pathologies (6, 32, 40). Many different kinds of compounds have been proposed to act as inhibitors of MetAP, with various functional components meant to enhance inhibitory potency utilized (46 - 51). We are interested in studying those that resemble natural substrates of MetAP, as well as compounds resembling enzyme reaction intermediates, as their assessment may also impart knowledge on the catalytic mechanism and active site of MetAP in addition to possibly diminishing enzymatic activity.

2.1.1. Transition-State Analogues

The transition state of a chemical reaction is a particular configuration along a reaction coordinate which corresponds to the highest energy of a reaction intermelong the reaction pathway [Figure 13]. Conceptually, it is a chemical reaction or process that involves the rearrangement of matter as proceeding through a continuous change or "transition state" in the relative positions and potential energies of the constituent atoms and molecules (61). Designing enzyme inhibitors as structural analogues of reaction intermediates, and as apparent analogues of transition states, is well established. This paradigm is based on replacing the main substrate functional group so as to entirely alter the reacting system, for example, substitution of a scissile bond by a non-scissile one. This allows the integration of features into a stable molecule that mimic primary

transition–state interactions of the natural substrate and the enzyme active site (62). Therefore, this approach aims to exploit the capability of an enzyme active site to stabilize a transition state structure so that the stable transition state analogue is bound with a far higher affinity than the normal substrate (63).

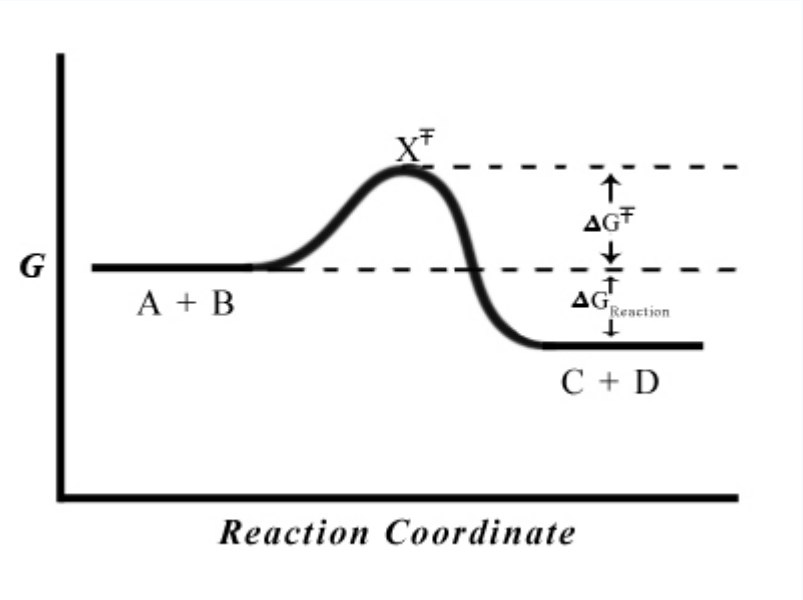


Figure 13. The transition state energy diagram; G = Gibb's free energy, A + B = reactants, P + Q = products, $X^{\ddagger}$ = transition state. Note the difference in reaction energy between the $\Delta G_{reaction}$ and $\Delta G^{\ddagger}$. <*Adapted from reference (61)*>

Methionine phosphinate and phosphonates have been shown to bind to the active site of eMetAP (46). Phosphorus-containing compounds form a third hydrogen bond between an oxygen atom of the phosphonate group and His 79, which is either manifest of movement or charge neutralization and is indicative of His 79 directly interacting with the amide nitrogen of the scissile peptide bond. The significance of such a finding is the similarity between the phosphonates / phosphinates and the natural MetAP substrate transition state (46). To a similar extent, it is noteworthy that mutations in the aforementioned active site residue His 79 inactivate the enzyme (46, 54).

Inquiry into phosphorous analogues of alpha-amino acids like methionine began in the 1940's, but underwent a renaissance during the late 1970's and early 1980's because of their observed anti-bacterial activity, potently inhibiting bacterial cell wall biosynthesis (64). More recently, their anti-angiogenic capabilities have been investigated and have resulted in further design and testing of compounds with this aim (64). The tetrahedral intermediate for the hydrolysis of the N-terminal methionine of a peptide by MetAP might be well mimicked by a phosphonate [Figure 14].

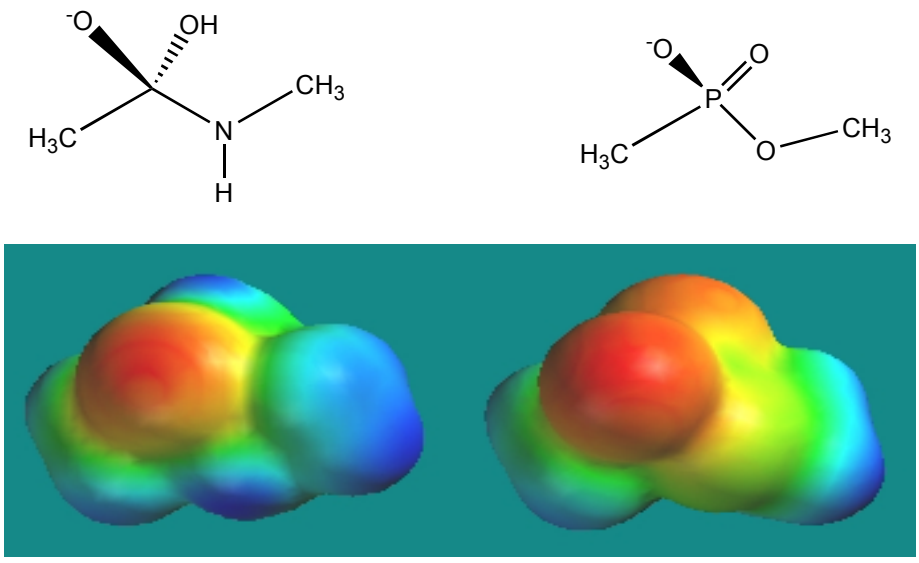


Figure 14. Top panel: The tetrahedral intermediate formed when OH attacks CH3(O)NHCH3 as a nucleophile (left), phosphonic acid, which resembles a tetrahedral intermediate (right). Bottom panel: The electrostatic potential of the tetrahedral intermediate (left) and phosphonic (right) acid mapped on to the electron density at the RHF/6-31+G* level.

It would therefore be of interest to evaluate the racemic form of amino acid phosphonate analogues with *E. coli* MAP to ascertain precisely what level of inhibition can be obtained with these simpler analogues. Although chiral analogues can also be prepared, the racemic mixture would allow for a more crude, preliminary analysis of these types of compounds, unmodified at this stage by conjugation with additional

groups on the phosphorus. In collaboration with our laboratory, Professor B. Matthews and his group (University of Oregon) have determined the molecular structure of MetAP bound with the phosphonate analogues of methionine (46). Therefore, it is of interest to build upon this previous work and continue to look into amino acid phosphonate analogues and their inhibitory potential.

2.1.2. Substrate-Mimic Inhibitors

Designing inhibitors that resemble an enzyme's natural substrate is a strategy that has been long employed in drug design with great success. One such example in the case of MetAP has been to model novel drugs based on the structure of the natural product bestatin [Figure 15] (46). By replacing the N-terminal side chain of bestatin with that of norleucine and by putting an alanine at the P_1' position, a non-hydrolyzable substrate analogue (AHHpA-Ala-Leu-Val-Phe-OMe) was produced which inhibited eMetAP. An inhibitory concentration (50%, IC_{50}) was determined to be 5 uM (19). X-Ray crystallography has shown the binding mode of the modified inhibitor in complex with the active site of eMetAP [Figure 16] (19). A number of solvent atoms are seen to be displaced relative to the metal center, with the coordination number of the first cobalt (Co_1) changing from five to six. The inhibitor interacts with the enzyme in three ways: ligation to Co1 via the keto oxygen of the pseudo-peptide linkage of the AHHpA moiety, N-terminal ligation to the second cobalt atom (Co_2), and a bridging coordination between Co_1 and Co_2 (19). Additionally, two hydrogen bonds are formed, specifically with residues His 79 and His 178 (19).

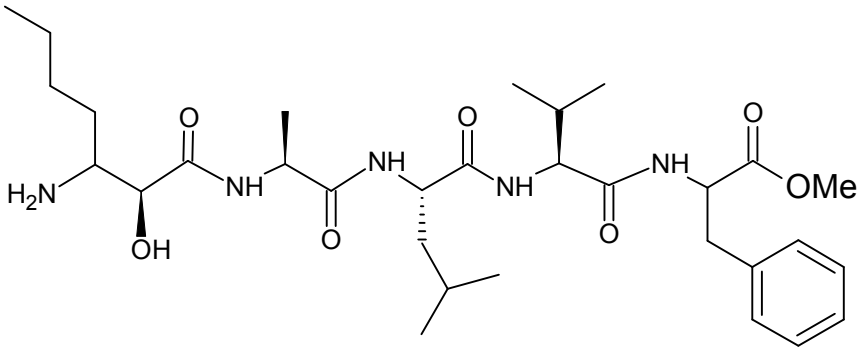


Figure 15. Bestatin, a natural product inhibitor of MetAP.

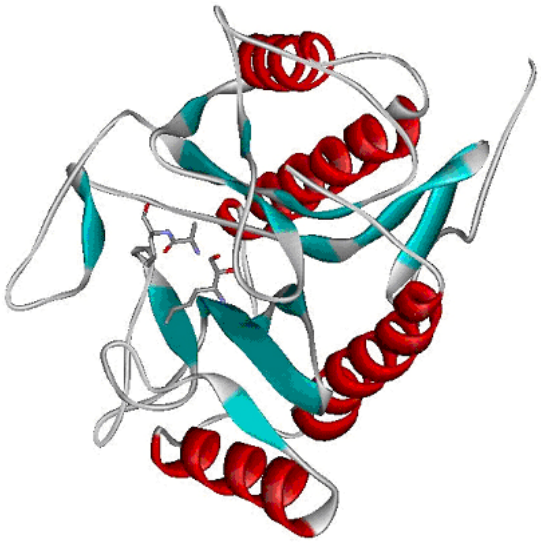


Figure 16. The bestatin-based inhibitor MetAP inhibitor containing (3R)-amino-(2S)-hydroxyheptanoic acid (AHHpA), in complex with eMetAP's active site. [PDB code: 3MAT] <*Taken from reference (19)*>

Lowther *et al.*'s findings are intriguing, and encourage exploring novel non-hydrolyzable analogues to further understand the nature of MetAP activity. Previous work has shown that thiopeptides can resist the action of a protease (65). For example, the work of Belleau and coworkers has shown that thiopeptides were effective in blocking metalloendopeptidase-mediated release of lysozymes (65). Even more compelling is the fact that Fosamprenavir and Tipranavir, two commercially popular protease inhibitors (PIs) contain sulfonate groups, while Nelfanavir, another well-known PI

contains a thioester (66). As well, with our interest in compounds that have similar structures to the natural substrates of MetAP, the sulfur atom contained within methionine is a point of curiosity, as we would like to answer the question of what might occur with MetAP if the thiomethyl functionality of methionine was replaced by a fluorinated version which altered the steric and electronics of this moiety. Doing so would give further insight into the catalytic mechanism and active site dynamics of MetAP.

2.2. MetAP's Metal Identity Crisis

MetAP, as a metalloprotease, requires divalent metal ion co-factors in its active site to function. Without metal, apo-MetAP is unable to hydrolyze the Met-AMC substrate. However, as touched upon earlier, the precise identity of the metal ion required by the *E. coli* MetAP is disputed (6, 10, 12, 20, 22, 23). Classically, Co^{2+} has been the metal most researchers believe facilitates MetAP activity (6). However, more recent studies have shown that in the case of yMetAP, Zn^{2+} is the preferred ion (20). In other reports, Fe^{2+} is believed to play this crucial role in eMetAP, with several other research groups exhibiting different metal content in the active site (25, 67, 70).

As we are interested in MetAP's role in the protein biosynthetic pathway, we decided to investigate what metal serves as a co-factor for the enzyme using the fluorogenic assay with Met-AMC using various divalent metal ions to confirm previous reports before we undertook a study of various compounds as potential inhibitors.

2.2.1. Met-AMC Fluorogenic Assay

Much of the work done in our laboratory on MetAP catalysis has been done using a high performance liquid chromatography-based assay to assess Met cleavage (67), as well as a coupled assay involving MetAP and prolyl iminopeptidase to measure enzyme kinetics (68). The HPLC assay assesses eMetAP's cleavage of a standard peptide previously utilized in our laboratory, Met-Gly-Met-Met, and whether adding inhibitors increases or decreases the accumulation of free N-terminal Met-truncated peptide (67) [Scheme 6].

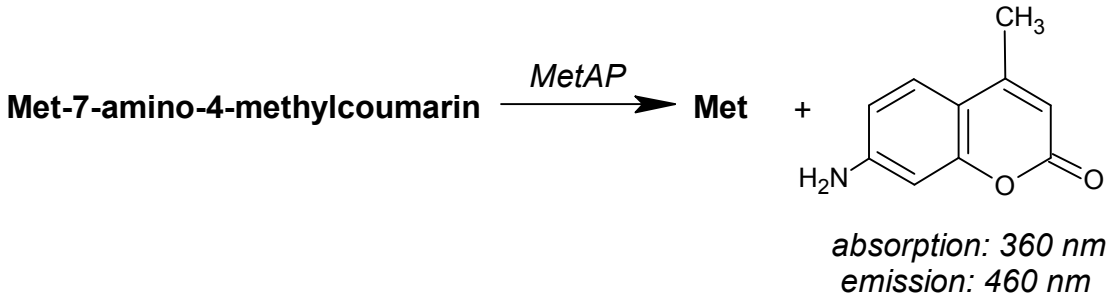
Scheme 6

The coupled MetAP reaction couples MetAP catalysis to the enzyme prolyl iminopeptidase (ProAP). A standard dipeptide substrate, methionylprolyl-*p*-nitroanilide (Met-Pro-*p*NA), has been used for optimization of the coupled assay as developed by Zhou et al. (68). ProAP is best suited for coupling to the MetAP reaction because it has a different substrate specificity from MetAP while also performing catalysis at a rate higher than the native MetAP reaction. Specifically, ProAP from *Bacillus coagulans* is employed because it exhibits a high specificity for N-terminal prolyl residues, which MetAP does not act upon, and also shows no detectable activity towards a Met-Pro-*p*NA dipeptide (68), which MetAP acts upon.

However, after several attempts to apply this assay approach in our laboratory, these methods were abandoned for several reasons. For one, our early attempts at expressing ProAP proved infeasible. The plasmid we obtained from Dr. Zhou (Ohio

State University, Columbus, Ohio) , the person who developed the assay, for ProAP did not have the correct sequence. In addition it was not possible to amplify with supplied DNA by PCR and it was not possible to express the correct protein. Another reason for changing our approach to assaying MetAP was that the HPLC assay is very lengthy and laborious, and the results are often difficult to interpret based on co-purification of the products. However, our primary reason for abandoning the previous assay protocols was because of our intent to adopt a new, simplified but powerful and fast fluorogenic assay, previously made use in a metal-selective inhibitor study on MetAP by Li and colleagues at the Chinese Academy of Sciences in Shanghai, China (69).

In Li *et al*.'s publication, the authors describe their use of the commercially available substrate Met-7-amino-4-methylcoumarin (MetAMC) from Bacham Bioscience (King of Prussia, PA), which is assayed along with MetAP on a microplate reader capable of measuring fluorescence [Scheme 7]. To this, they were able to add inhibitors from a compound library to their assay mixture, and determine whether the inhibitor's presence caused a relative decrease in fluorescence as a result of it's preferred metal selectivity (69). This inspired us to attempt to incorporate the same assay in our studies, and develop it, tailoring it for our purposes.



Scheme 7

Under appropriate conditions using a fluorescence microplate reader, one is capable of monitoring an assay in real-time, following the reaction as it proceeds, and obtaining kinetic data simultaneously. Accordingly, MetAP's cleavage of Met-AMC should cause an increase in the presence of fluorescence. The greater the cleavage, the greater the concurrent increase in fluorescence. However, with compounds present in the assay that would be thought to inhibit MetAP, fluorescence readings would be less than simply the enzyme alone with the substrate. This same train of thought would also apply to metal selectivity – if the metal ion present in the assay is conducive to MetAP activity, fluorescence readings should indicate this as a relative increase in enzymatic cleavage of the substrate, and likewise for a metal which does not activate the enzyme, no fluorescence should be observed. Appropriate background hydrolysis of Met-AMC would also be measured.

2.2.2. Experimental

2.2.2.1. Materials

Commercially obtained recombinant *E. coli* Methionine Aminopeptidase (20 ug in 20 µL) was obtained from R&D Systems, Inc. (Minneapolis, MN) and stored at -80 ^{o}C. Methionine-7-amino-4-methyl coumarin (Met-AMC) fluorogenic substrate was obtained Bacham Bioscience (King of Prussia, PA). Black 96-well opaque microplates were obtained from Corning Inc. (Corning, NY)

2.2.2.2. Met-AMC Assay

When needed for assays, reMetAP (20 ug in 20 µL) was reconstituted in a 100 µL buffer solution at pH 7.5 with 50 mM HEPES and 0.1 M NaCl. Kinetic assessment of enzyme activity was adapted from the procedure of Li *et. al.* (69) with the following conditions. Assays were performed on a SpectraMax 340 microplate reader set on fluorescence detection, from Molecular Devices (Sunnyvale, CA) and controlled by Softmax software. Assays with Met-AMC (Bacham Bioscience, King of Prussia, PA) were performed in 96-well clear black, opaque microplates. During each assay, the formation of 7-amido-4-methylcoumarin (AMC) was monitored continuously for 15 to 30 min at room temperature (25 ^{o}C) by fluorescence (λ_{ex}=360 nm, λ_{em}=460 nm). In each 200 µL well, an 80 µL reaction mixture was added that contained 50 mM HEPES, pH 7.0, 1mM NaCl, and and 0.1 uM apo-eMetAP (which translated to 0.75 µL being added to a well). The concentrations of Met-AMC substrate was varied to 1mM, which is different from the single concentration used by Li *et. al.*(69). However, unlike the aforementioned group, we did not vary metal concentrations, and kept them static at 0.1 mM. In addition to Co^{2+}, the metal typically regarded as the preferred metal in the active site of MetAP, several other metal-containing buffer solutions were prepared, at 0.1 mM concentration. In particular, the additional metals used were: Zn^{2+} / Mn^{2+} / Mg^{2+} / Ca^{2+} / Ni^{2+} / Fe^{2+}. Kinetic constants K_m and k_{cat} were obtained using the GraFit6^{TM} software through non-linear regression fitting of the rate v.s. substrate concentration curve (also referred to as the Michaelis-Menten curve).

2.2.2.3. Metal-Free Assay Conditions

It was assumed originally that our commercially purchased recombinant eMetAP was completely devoid of metal content. This was further presumed by the suppliers themselves, who confirmed as much when contacted directly. However, our initial experiments with the enzyme quickly demonstrated that such was not the case, and in fact, background metal ions were present in the original enzyme stock as the levels of Met-AMC hydrolysis with no metal present in the buffer solution were significantly higher than expected (at times, almost as high as assays run with metal present!). Therefore, a protocol was developed to insure that our assay was completely free of metal so that the results obtained with our assays would not be confounded with exogenous metal. This also served as a means to permit us to experiment with different, specific divalent metal ions in our assays, being able to assess the differential level of activation by the enzyme with various divalent metals in the active site as the particular metal ion involved in metalloprotease catalysis is hotly debated (6, 10, 12, 20, 22, 23).

To put it simply, in order to remove all possible traces of exogenous metal from the assays, everything involved in the assay had to be treated as a possible source, and dealt with accordingly. Thus, all apparatus' and materials involved, from the plastic weighing trays to the micropipette tips, were soaked in 20% nitric acid solution overnight prior to running assays. The enzyme itself (1 uM, 17 µL) was taken from the original stock solution and placed for 24 to 72 hours in 983 uL of Chelex-treated buffer (total volume: 1 mL) and incubated with 2,6-pyridinedicarboxylic acid (DPA) (0.2 mM, 0.033

g), which had been successfully utilized in our lab and in the literature to remove metals from other metalloenzymes [Figure 17] (70). Before performing assays, the dialyzed enzyme was passed through an Ultrafree-MC centrifugal filter (30 KDa MW cut-off) and reconstituted in 17 µL of original buffer solution [1 uM]. All nitric acid soaked assay relevant supplies were rinsed in Chelex-treated MQW water. Once these procedures

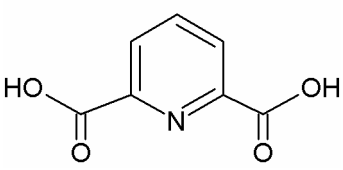


Figure 17. 2,6-Pyridinedicarboxylic acid

were carried out, externally-derived metal contamination had been reduced by approximately 80%, based on the relative fluorescence readings of assays run with the enzyme and substrate the presence of a non-metal containing buffer solution.

2.2.2.4. Fe^{2+} in Assays

E. coli MetAP is known to function optimally with Fe^{2+} during catalytic activity (71). As mentioned previously, it has even been investigated whether certain inhibitors function more potently, or are "Fe^{2+}-selective" (25). However, the precise identity of the metal which MetAP, and metalloproteases utilize, is hotly contested, with much literature on the subject having been published in order to establish a definitive role for one divalent metal over another. We decided we would like to investigate this as well, and utilize the fluorogenic assay to this end.

Unfortunately, utilizing Fe^{2+} during biochemical assays is not as simple as dissolving the ferrous sulfate ($FeSO_4$) metal crystals into the buffer solution. Fe^{2+}, once

dissolved in solution, very readily and quickly oxidizes to Fe^{3+}. In fact, if one does not expect this to occur, then they are quite shocked as almost as fast as the iron is added, the buffer solution turns purple, and then black depending on the concentration [Figure 18]. So, based on the necessity for Fe in its +2 oxidation state form to populate the

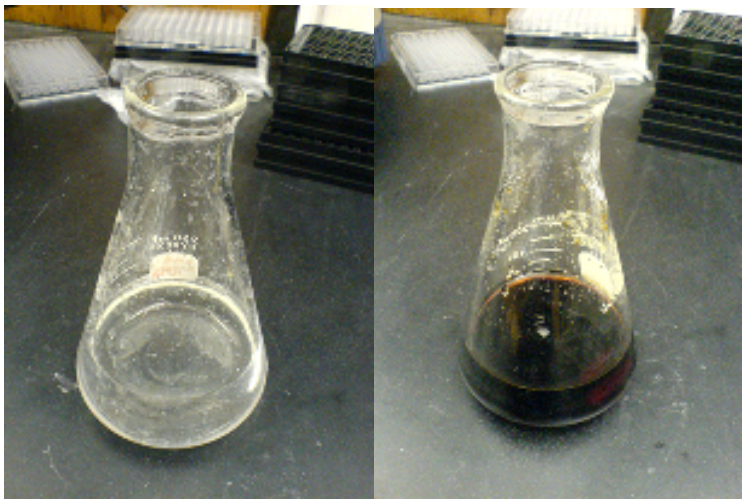


Figure 18. In the first panel, a buffer solution is shown before $FeSO_4$ is added. In the second panel, $FeSO_4$ has been added and moments later, rapid oxidation from Fe^{2+} to Fe^{3+} has occurred.

catalytic region of MetAP, appropriate steps had to be taken to ensure that rapid oxidation to the ferric species of Fe did not occur.

Wang and colleagues (2008), in their report on inhibitors selective for the Fe(II)-form of MetAP, document their usage of ascorbic acid when performing their assays as an anti-oxidant, preventing oxidation to Fe(III), with which MetAP cannot function (25). The anti-oxidative properties of ascorbic acid are well established (72). Therefore, in order to prevent ferrous oxidation from occurring in our own assays, we developed a modified version of their buffer preparation procedure. Before adding the metal, argon gas was bubbled in for approximately 30 minutes, with all oxygen being exited from the system to insure the entire environment was oxygen-free. After 15 minutes, 0.1 mM of

ascorbic acid was added to the solution while argon gas was still being bubbled into the buffer. We waited another 15 minutes while the solution continued to mix, and then carefully added 0.1 mM $FeSO_4$, returning it to the argon atmosphere for five minutes. Assays were then quickly performed, as even in spite of our cautious preparation of the buffer and containment of it during experiments, oxidation slowly occurred. This is demonstrated in Figure 19, which shows the level of relative activity of MetAP in relative fluorescence units (R.F.U.) over time using the prepared Fe^{2+} buffer.

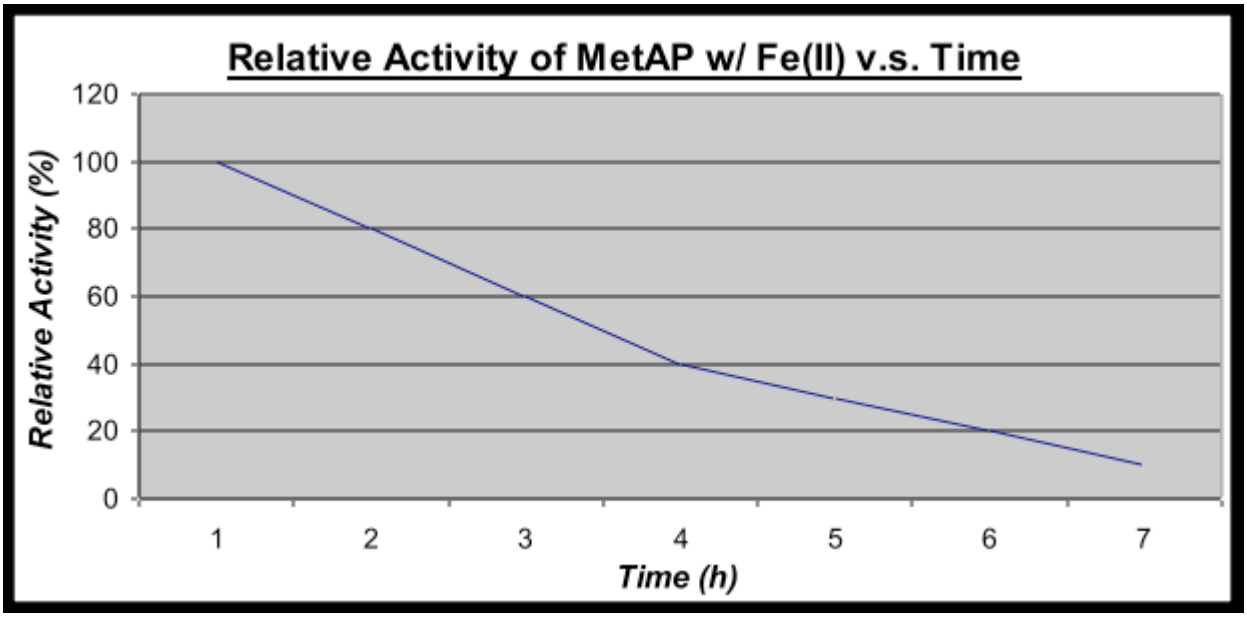


Figure 19. MetAP activity as a function of time, reported in R.F.U., using an Fe(II) buffer solution. As time passes, MetAP activity decreases, presumably due to Fe2+ oxidation to Fe3+.

As time elapses during assay experiments, there is a relative decrease in activity of eMetAP using the Fe(II) buffer prepared. When optimizing the assay originally, at times, addition of $FeSO_4$ would automatically result in a blackened solution, indicating rapid oxidation from Fe^{2+} to Fe^{3+}. When attempting to perform assays using this darkened buffer, with fresh enzyme or enzyme used before, there was virtually no activity present. This then motivated us to check if in fact there was a correlation

between the level of oxidation having occurred in the buffer (perceptible by colour), and how much relative activity there was in the assays performed with this perpetually colour-changing solution (from clear to purple to black) [Figure 18]. So, the original enzyme solution made use of in the assay was assessed for activity using the assay procedure specified in Section 2.2.2.2., with new substrate, each hour after buffer preparation with this slowly oxidizing buffer. To compare, a new batch of enzyme was used as a control, as we were not sure whether the enzyme was simply denatured or not. The results were identical whether the original enzyme batch was used, or the enzyme freshly removed from -80 ^{o}C. However, upon preparing a fresh buffer solution, activity was restored to optimal levels. Interestingly, during preliminary Fe^{2+} assays, if the buffer was not disposed of at the end of running experiments, and was left overnight, the next day, the solution was completely black and opaque, similar to the appearance of the solution in the second panel of Figure 18.

In addition to these concerns, there was also the issue that Fe^{2+}, in complex with the ascorbic acid, might function itself as an entity capable of hydrolyzing Met-AMC. Clearly, this was problematic as the results from the Fe^{2+} experiments would have to be interpreted with the proper controls. However, at the same time, despite our apprehensions, it was of interest to assess whether this indeed occurred, as it would have yielded yet another function of the almost inexhaustible roles ascorbic acid plays biologically, and a further demonstration of the role ascorbic acid and Fe play catalytically in biological systems (73, 74). Assays run with Fe^{2+} and ascorbic acid did in fact show some Met-AMC hydrolysis, ranging from between 5 to 10% of what we

used as the control experiment, being the Fe^{2+} + MetAP assay. This is quite minute however, so it was simply factored in as background signal.

2.2.3. Results and Discussion

2.2.3.1. Metal-Activation and Kinetics of MetAP with Various Divalent Metal Ions

As can be seen from Table 1, there is quite a range of catalytic activity. The most eye-catching result is probably that Fe^{2+} exhibits the highest catalytic activity rate

Metal Used In Buffer Preparation	K_m (M)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
Ni^{2+}	0.0052 +/- 0.00035	0.00342 +/- 0.00048	0.657692
Fe^{2+}	0.0038 +/- 0.00059	0.01845 +/- 0.00237	4.855263
Mn^{2+}	0.0039 +/- 0.00044	0.01435 +/- 0.00201	3.679487
Mg^{2+}	0.0015 +/- 0.00017	0.00013 +/- 0.00008	0.866667
Ca^{2+}	0.0019 +/- 0.00012	0.00335 +/- 0.00062	1.763158
Co^{2+}	0.0049 +/- 0.00036	0.01903 +/- 0.00599	3.883673
Zn^{2+}	0.0037 +/- 0.00062	0.01715 +/- 0.00403	4.635135

Table 1. The K_m, k_{cat}, and k_{cat}/K_m values for the assays performed using various metals.

of all seven metals tested, giving the highest k_{cat}/K_m value. To those familiar with MetAP, this might seem surprising, considering that in the past, Co^{2+} has been understood to be the native eMetAP metal co-factor (6). However, in more recent times, it has been shown that instead, Fe^{2+} is the preferred metal for eMetAP, and our results are in line with this finding (71). In our assays, Zn^{2+} also exhibited quite high catalytic activity, as did Mn^{2+}. As stated previously in this document, Zn^{2+} is considered the native co-factor in yeast, but it is quite intriguing that it is demonstrably preferred over Co^{2+} as well in eMetAP. Mn^{2+}, the other metal with relatively elevated observed

activity levels, is also shown in the literature to be capable of activating MetAP efficiently, so its high level of activation is not surprising.

The Co^{2+}-MetAP is the most prevalently used metalated form in *in vitro* studies (6). However, as others have shown, activity is maintained (and in some cases enhanced or inhibited) using other divalent metals, including Zn^{2+}, Mn^{2+}, Ni^{2+}, Mg^{2+}, Ca^{2+}, Cu^{2+} and Fe^{2+}(6, 10, 12, 20, 22, 23, 76). In the case of yMetAP, Zn^{2+} is the enzyme providing the optimal level of activation (23). In eMetAP, Fe^{2+} has been found to be the metal most favored by the enzyme (71). Further, the hMetAP type II is considered to prefer Mn^{2+} over all other divalent metal ions (67). Therefore, our studies serve not only as support for those previously observed results in eMetAP, but also as a comprehensive assessment of eMetAP with all divalent metals (except Cu^{2+}) used previously in the literature. Our findings of Fe^{2+} being the most potent activator of MetAP is very much in agreement with those of Chai et al. (2008) that show Fe^{2+} is probably the native co-factor of MetAP (71). Moreover, our observation that Ca^{2+} does not facilitate significant enhancement of MetAP activity is also what Chai and colleagues noted, as well as other groups (71). In the case of our Mn^{2+} assays, the high levels of activation and activity, as defined by k_{cat} / K_m were expected. Mn^{2+} has been shown to bear preference to various MetAPs, including yMetAP, hMetAP-II, eMetAP, and *Pyrococcus furiosus* MetAP Type IIb (pfMetAP-IIb) (71, 76) , and has been shown to bind preferentially in the active site by Li and Chang (1996) and D'Souza et al. (2002) (28, 76). The fact that Ni^{2+} activates eMetAP most poorly is not that surprising. It has been shown that although it is capable of activating eMetAP, the results are minimal

(27). What's more is that if the enzyme is depleted of its metal, and then re-metalated by Ni^{2+}, activity is almost negligible (27). In other types of MetAP, like bstMetAP-Ia, Ni^{2+} has even been found to inhibit activity (77).

We were slightly surprised by the result of Mg^{2+}-containing assays having such low amounts of MetAP activity, only marginally greater than Ni^{2+} which has been observed to be a poor activator of MetAP (77). In porcine kidney MetAP (pkMetAP), Mg^{2+} has been found to facilitate activity to a great extent, as is the case with bovine lens MetAP (77). Though it may not have been expected to be a particularly strong outlier, the fact that Mg^{2+} performed so relatively poorly is contrary to what has been seen in other organisms, like *Bacillus stearothermophilus* MetAP *(*bstMetAP) (77). However, with hMetAP-IIb, Yang *et al.* (2001) found Mg^{2+} only a moderate activator of MetAP, especially in comparison to Co^{2+} and Zn^{2+}, which in our assays, proved quite efficacious (78). However, it was more unexpected that our Zn^{2+} assay demonstrated increased MetAP catalysis over Co^{2+}, as other studies have shown that Zn^{2+} confers significantly weaker activity in eMetAP versus Co^{2+}, one of the reasons Co^{2+} becoming the classically used metal of choice with MetAP (78). However, as mentioned previously, Zn^{2+} is the preferred metal in yMetAP, and so it's fostering of elevated catalytic activity may be more widespread amongst other organisms and warrants further study.

If one takes these results as a whole, what can be seen in a broad sense is that all divalent metals tested facilitate catalytic functioning by MetAP. Though some metals,

like Ni^{2+} did so quite minimally, there still existed catalysis. Clearly, the 'one metal, one enzyme' belief that researchers previously held for metalloenzymes may require some revision. If this were true, then there would have been much different results in our assays, especially based on our care to insure that no metal contamination was present during our assays. One, perhaps two of the divalent metals tested, would have fostered MetAP activity quite appreciably, while the rest would have more or less remained at the same low catalytic levels, indicating a more definitive metal for our enzyme. It may simply be the case that MetAP, and possibly other metalloproteases, are capable of functioning within a variety of metalated-environments, effectively adapting to the surroundings of the organism.

2.3. MetAP: A Moonlighting Protein?

It is becoming more widely accepted that the one gene – one protein – one function rule is less applicable to complex biological systems of higher organisms (79). It has been reported in a number of journal articles that many proteins can be involved in cells in processes or structures that are not related to their established or normally "main" or originally described function [Figure 20]. This does not even include the many instances of where proteins have undergone gene fusions, homologous but non-identical proteins, splice variants, proteins whose post-translational modifications can vary and proteins that have a single function but can operate in different locations or utilize different substrates (79). These proteins are called "moonlighting" proteins since other functions for them have been detected in addition to their first described function. Thus, it is always exciting to find an alternative role for a protein under study.

Enzyme	Secondary Biological Activity
PutA proline dehydrogenase	Transcriptional repressor
Phosphoglucose isomerase	Autocrine motility factor
Thymidine phosphorylase	Platelet-derived endothelial cell growth factor
Uracil-DNA glycosylase	Glyceraldehyde-3-phosphate dehydrogenase
EMS2 mismatch-repair enzyme	Hypermutation of antibody variable regions
Ribosomal proteins	Development regulation
Thrombin protease	Ligand for cell surface receptors
Mitochondrial LON protease	Chaperone
Band-3 anion exchanger	Regulator of glycolysis
Bacterial FtsH chaperone	Metalloprotease

Figure 20. Examples of "moonlighting" proteins and their secondary biological activity. ***<Based on data from reference (79)>***

Based on the interest in MetAP within our laboratory, we are always curious to know as much about an enzyme as possible. One area that has not been investigated is MetAP's activity in other biochemical domains. Its role as a metalloprotease involved in NME has been almost definitively stated in the literature, however, one known alternative function of MetAP does exist. In rats, rat initiation factor-2 associated protein (p^{67}) plays a vital in regulating translation through preventing the phosphorylation of the α subunit of eukaryotic initiation factor-2 (eIF-2). However, the human homologue, was found to exhibit activity MetAP-like activity, with similar kinetics, substrate specificity and even divalent metal-dependance (80). Human p^{67} (Hp^{67}) has even been found to bind the natural product potent MetAP inhibitor, fumagilin, boosting Hp^{67} obstruction of eIF-2 α subunit phosphorylation (81). With this unique function in mind, we hypothesized whether MetAP has the capability of hydrolyzing substrates not previously thought suitable for it to act upon. Since we were already investigating the

functional implications of phosphonates as possible transition-state inhibitors of MetAP, we wondered whether phosphate-containing substrates could be processed by MetAP at all. Therefore, we evaluated the compounds 4-nitrophenyl-O-phosphate (4NP) and bis(4-nitrophenyl)phosphate (BNPP), since both had been shown to be acted upon by a variety of enzymes (82, 83, 84), we speculated whether MetAP could also treat those compounds as substrates and yield free p-nitrophenol, seen as a typical yellow colour in somewhat basic solution.

2.3.1. Experimental

2.3.1.1. Materials

The compounds 4-nitrophenyl-O-phosphate (> 97% purity) and bis(4-nitrophenyl)phosphate (> 95% purity) were purchased from Sigma Aldrich (St. Louis, MO). Recombinant *E. coli* MetAP was obtained from R&D Systems, Inc. (Minneapolis, MN).

2.3.1.2. MetAP Phosphatase Assays

The assay procedure specified in Section 2.2.2.2. was modified to utilize the SpectraMax 340 microplate reader's capacity to read absorbance rather than fluorescence as done previously in these studies. This was done with the appropriate UV absorbance of both 4NP and BNPP [Figure 21] (found to be optimal at 405 nm) on a

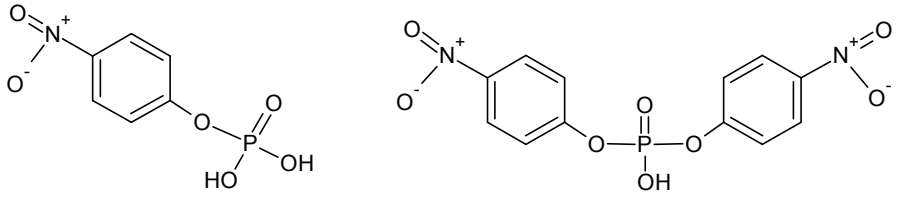


Figure 21. The compounds 4NP and BNPP.

clear 96-well microplate rather than a black opaque one. A buffer solution was then prepared similarly as in Section 2.2.2.2. However, to the 96-well microplate, 1.0 mM of 4NP or BNPP was added rather than Met-AMC as the substrate. The assay was then begun by reading the UV absorbance over time (72 h).

2.3.2. Results and Discussion

2.3.2.1. MetAP Phosphatase Activity

Initially, it was quite frustrating because there was no change in absorbance after the usual assay run times (30 minutes) with either 4NP or BNPP. So therefore, the assay times were increased from 30 minutes to an hour, then from an hour to two hours. After two hours, there was an ever-so-slight change in absorbance, but it had to be magnified within the Softmax software in order to see it with any worthwhile resolution. As we were frustrated not having seen any significant catalytic activity of MetAP during the 4NP / BNPP assays, we moved on to perform other experiments in another area of the lab. However, the 96-well microplate containing the 4NP / BNPP assay mixtures were forgotten in the microplate reader. As it happened, a group from another lab had to use the microplate reader, and inadvertently took the microplate from the reader and placed it in the 37 ^{o}C oven without inquiring, with the thought to save the contents from decaying at room temperature as their own assays required physiological conditions and temperatures. About three hours elapsed and upon our return to this area of the lab, the other group was still using the reader, and alerted us to what had transpired. Recovering the microplate, we saw that the wells containing the reaction

mixture had turned slightly yellow. After an endpoint reading of these wells and comparing those numbers to newly prepared 4NP / BNPP-containing wells that had not had any enzyme added yet, it was clear that pNP had been liberated from our substrates. With this, our approach to the 4NP and BNPP assays had unquestionably changed.

Rather than continuously following MetAP catalysis of 4NP and BNPP, we decided to take endpoint readings every 15 minutes as it seemed that the accrual of free p-nitrophenol was not sufficiently rapid to monitor with the kinetic setting of the microplate reader. As well, our assays were run at both room temperature and 37 ^{o}C, to see if there was indeed a difference when run at physiological temperature based on what had seemed to happen in the oven [Figures 22 and 23].

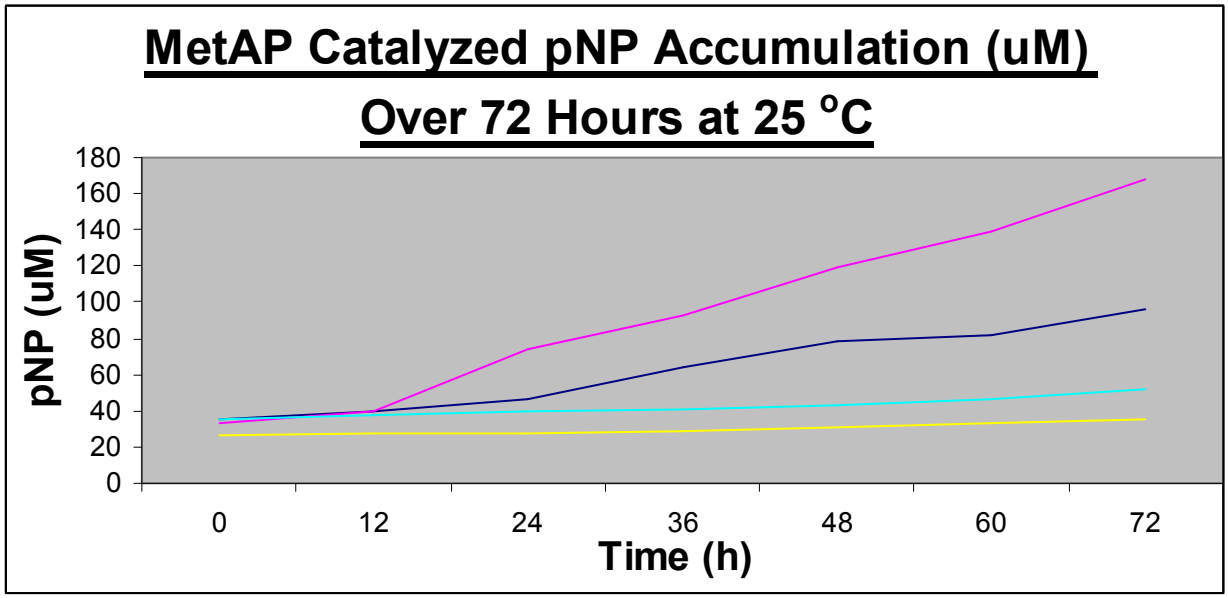


Figure 22. The accumulation of pNP through MetAP activity on 4NP and BNPP (Blue: 4NP [1.0 mM]; Magenta: BNPP [1.0 mM]; Yellow: 4NP [1.0 mM] (No Enzyme); Light Blue: BNPP [1.0 mM] (No Enzyme)) at 25 ^{o}C over 72 hours.

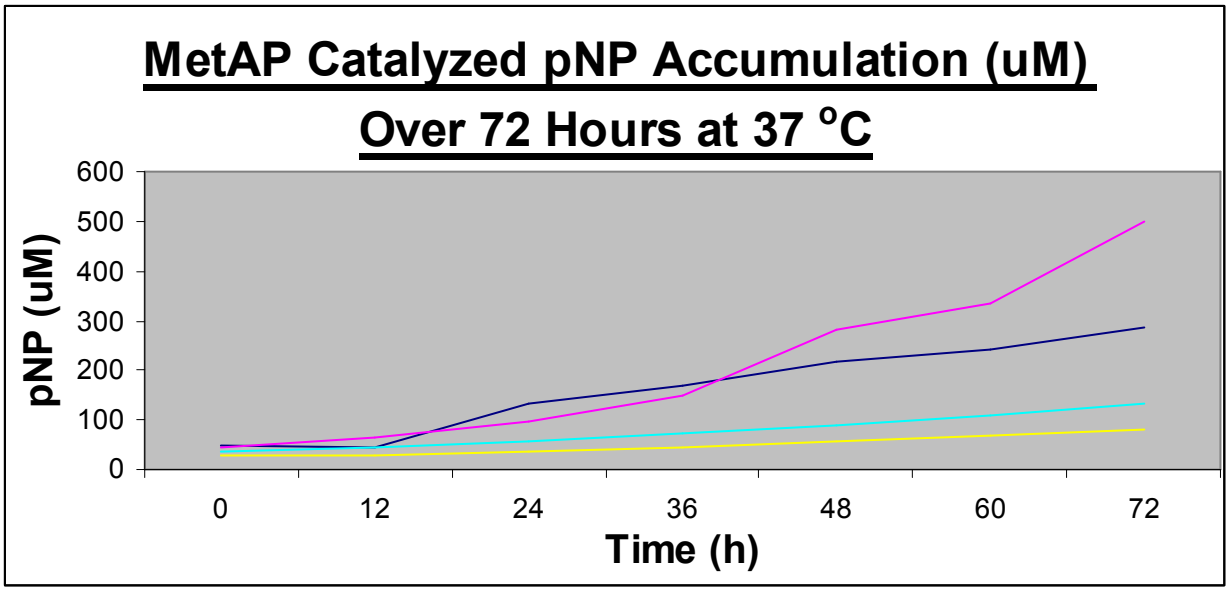


Figure 23. The accumulation of pNP through MetAP activity on 4NP and BNPP (Blue: 4NP [1.0 mM]; Magenta: BNPP [1.0 mM]; Yellow: 4NP [1.0 mM] (No Enzyme); Light Blue: BNPP [1.0 mM] (No Enzyme)) at 37 °C over 72 hours.

As can be seen from viewing tables Figures 22 and 23, the results indicate that MetAP had catalyzed the hydrolysis of the 4NP and BNPP molecules. This is further supported with Figure 23, a photograph taken of the clear 96-well microplate used in the experiments housing the assay mixtures after 72 hours – the characteristic yellow colour indicates pNP formation [Figure 24]. Moreover, this hydrolysis is greater at higher temperature, which is expected in most enzymatic reactions.

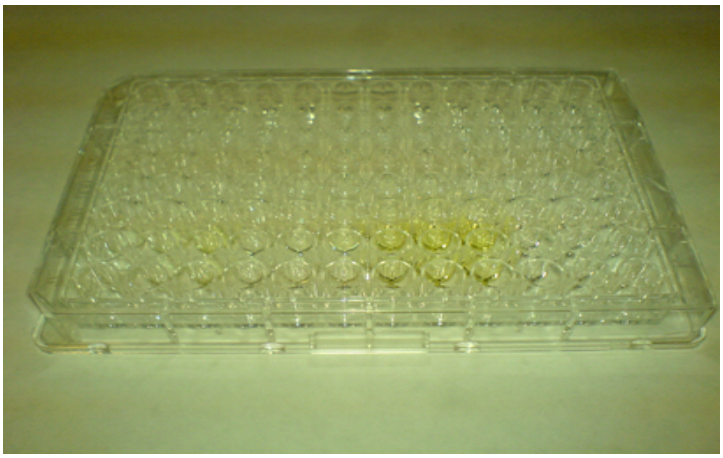


Figure 24. The yellow colour in the bottom center of the clear 96-well microplate is indicative of reactions with 4NP / BNPP having formed free pNP.

BNPP appears to be better processed than 4NP, probably based on its size – the two nitrophenyl groups create much more hydrophobic interactions, while the extra nitro group may also foster greater polar interactions in the active site of MetAP than the single one present on 4NP. This appears to be in agreement with the heightened inhibitory potency of BNPP against 4NP observed in Section 2.4.2. (*to be discussed later*). Conversely, one might have assumed prior to performing the experiments, if MetAP would indeed act upon either of these substrates, the size of BNPP versus 4NP may have prohibited it from being processed to the same extent as 4NP based on its size being almost twofold larger, but this could be explained simply based on MetAP's active site being more than adequately wide enough to have room for BNPP's entry. The one unusual result, in spite of its accordance with what was previously observed in terms of continuous kinetic monitoring, is that it takes approximately 12 hours to notice a significant change in pNP concentration in both substrates we examined. Up until that point, MetAP does not seem quite as apt at dealing with these substrates, in spite of their rapid blocking of activity when treated as inhibitors. This might be due perhaps to what was mentioned concerning the interaction of moonlighting proteins with different types of molecules, as oppose to their native substrate, which may cause the enzyme in question to exhibit alternate biochemical roles (79).

These observations represent a novel catalytic reaction of MetAP, an enzyme previously believed to be only able to process substrates containing N-terminal Met residues. Further investigation is definitely merited especially to determine if this

phosphodiesterase reaction is a physiologically important reaction or just a laboratory curiosity.

2.4. Phosphorus- and Sulfur-Based Probes of MetAP

Based on our interest in molecular probes of MetAP, we decided to synthesize and purchase a series of compounds that contained phosphorus- and sulfur-based functional groups to assay with MetAP and to examine what happens to the enzyme in their presence, specifically investigating whether they were capable of inhibiting MetAP catalytic activity or not. Our rationale was based on our previous work on the use of phosphonate compounds previously demonstrated to bind to the active site of MetAP, as well as studies that demonstrated sulfur-containing molecules also successfully decreased metalloprotease enzyme activity (46, 65). The compounds used are all fairly simple in chemical structure. Our intent was to understand the structure-activity of their interaction with the enzyme.

2.4.1. Experimental

2.4.1.1. Materials

The phosphorus- and sulfur-based probes purchased were: Aminomethane sulfonic acid (97%) (Sigma Aldrich, St. Louis, MO), Aminomethylphosphonic acid (99%) (Sigma Aldrich, St. Louis, MO), Methylphosphonic acid (98%) (Sigma Aldrich, St. Louis, MO), Pentane sulfonic acid (97%) (Sigma Aldrich, St. Louis, MO), Methylene diphosphonic acid (98%) (Sigma Aldrich, St. Louis, MO), Iminobis(methylphosphonic)

acid (97%) (Sigma Aldrich, St. Louis, MO), p-nitrobenzenesulfonic acid (98%) (Sigma Aldrich, St. Louis, MO), para-toluene sulfonic acid (99%) (J. T. Baker, Phillipsburg, NJ), 4-Nitrophenyl-O-phosphate (97%) (Sigma Aldrich, St. Louis, MO), bis(4-nitrophenyl)phosphate (>95%) (Sigma Aldrich, St. Louis, MO) and Triethyl phosphonoacetate (>97%) (Sigma Aldrich, St. Louis, MO).

Diethyl phosphite (98%), valeraldehyde (97%), (diphenylmethyl)amine (97%), dichloro-5,6-dicyanobenzoquinone (DDQ) (98%), and anhydrous toluene (99.8 %) were obtained from Sigma Aldrich (St. Louis, MO). 4 Å crushed molecular sieves were purchased from EMD Chemicals Inc. (San Diego, CA). Anhydrous N,N-Dimethyl formamide (DMF) (>99%) was purchased from Caledon Laboratories Ltd. (Georgetown, ON).

Commercially obtained recombinant *E. coli* Methionine Aminopeptidase was obtained from R&D Systems, Inc. (Minneapolis, MN). The fluorogenic substrate 7-amido-4-methyl coumarin was acquired from Bacham Bioscience (99%) (King of Prussia, PA). 96-well black opaque microplates were obtained from Corning Inc. (Corning, NY).

All solvents were evaporated *in vacuo* on a rotary evaporator unless otherwise specified.

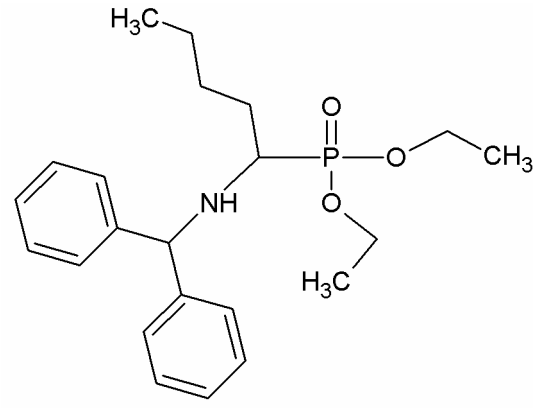
2.4.1.2. Synthesis of an Aminophosphonate

Besides commercially purchasing our compounds, we decided to try and make one as well. An earlier report by Yadav *et al.* at the Indian Institute of Chemical Technology, Hyderabad, India, indicated the facile use of microwave synthesis in a one-pot synthesis of alpha-amino phosphonates, and we decided to synthesize an analogue using their method as well (85). We focused on the synthesis of the norleucine phosphonate, an analogue that did not contain the thiomethyl functionality. This analogue had been previously synthesized and crystallized with MetAP, as reported by Lowther *et al.*, however, the inhibitory potency of this compound was not assessed at that time (46).

2.4.1.3. Synthesis of (Diphenylmethyl)norleucine aminophosphonate diethyl ester

KSF clay powder (1.5 g) (Sigma Aldrich, St. Louis, MO) was added to a microwave reaction vessel with a magnetic stir bar. To this was added 644 ul (0.005 mol) diethyl phosphite (Sigma Aldrich, St. Louis, MO), 532 ul (0.005 mol) valeraldehyde (Sigma Aldrich, St. Louis, MO), and 861 ul (0.005 mol) (diphenylmethyl)amine (DPM – a protecting group) (Sigma Aldrich, St. Louis, MO). Anhydrous dimethylformamide (DMF) (Caledon Laboratories, Georgetown, ON) was added until the total reaction mixture was 5 mL. At this point, the microwave pyrex tube was sealed with a wax-bottom top specifically made for use with microwave tubing, and set in the microwave to undergo irradiation at 150 ^{o}C for 10 minutes. After completion and cooling of the solution, the reaction mixture was evaporated, dried and re-dissolved in chloroform, then eluted on a

silica gel column with a mixture of ethyl acetate – hexane (3:7) to afford the protected norleucine aminophosphonate.

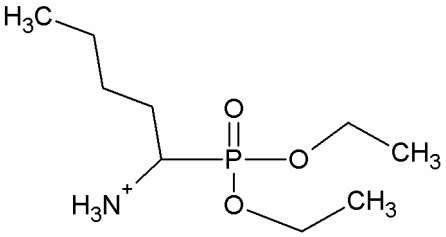


(Diphenylmethyl)norleucine aminophosphonate diethyl ester yield was 751 mg (85%). R_f (EtOAc / Hexanes 3:7) 0.11; ^{1}H NMR (500 MHz, $CDCl_3$) δ (relative to TMS, 0.00 ppm) 0.85 (m, 3H, C*H*$_3$ from alkane chain), 1.2-1.33 (m, 6H, CH_2 alkane, C_β, C_γ, C_δ), 1.24 (m, 6H, OCH_2CH_3), 2.75 (m, 1H, C_α), 4.08 (m, 4H, $POCH_2$), 5.25 (s, 1H, Ph_2CH-), 7.20-7.35 (m, 10H, arene), LRMS (ESI) calculated for $C_{22}H_{32}NO_3P$: 389.2120, observed at 390.22 ($M+H^+$).

2.4.1.4. Norleucine aminophosphonate diethyl ester

The following procedure was adopted from a previous graduate student in our laboratory, Dr. Peter Sampson (87). A solution of anhydrous toluene (3 mL) (Sigma Aldrich, St. Louis, MO) and 2 g of 4 A crushed molecular sieves (EMD Chemicals Inc., San Diego, CA) (stored in a 100 ^{o}C oven overnight) was added to a double-neck round bottom flask and placed in an argon atmosphere, with a thermometer placed in one neck and the other kept sealed. After 15 minutes, 0.038 g (5 mmol) of the DPM-

protected compound was added, followed by the addition of 0.34 g (1 mmol) of DDQ (Sigma Aldrich, St. Louis, MO) after which the complete reaction mixture was heated to 60 ^{o}C for 2 hours. The flask itself was covered with aluminum foil and the fume hood that this experiment took place in was kept in darkness as the conditions called for no light. Initially, the solution was deep red, however, after the 2 hours elapsed, the reaction mixture took on a more orange-ish appearance, and little flecks of purple precipitate could be seen. The mixture was then cooled to room temperature, and filtered on a column with Brockmann 1 basic alumina, washed with toluene to elute the newly deprotected aminophosphonate. After removing the toluene *in vacuo*, a yellow oil remained, which was dissolved in 10 ml diethyl either and 10 ml 0.1 N HCl. The mixture was stirred vigorously for 6 hours, followed by separating the phases using a separatory funnel. Each phase was washed twice with the other solvent being used, so in the case of diethyl either, this phase was washed with the HCl, and vica versa (2 x 5 ml). Combining the aqueous layers and then evaporating the HCl gave a hydrochloride salt of norleucine aminophosphate as a yellow oil (236 mg).



Norleucine aminophosphonate diethyl ester yield was 23.6 mg (31%). R_f (EtOAc / Hexanes 3:7) 0.09; ^{1}H NMR (500 MHz, $CDCl_3$) δ (relative to TMS, 0.00 ppm) 0.85 (m, 3H, alkane C*H*$_3$), 1.15-1.35 (m, 12H, methylene and OCH$_2$C*H*$_3$), 3.60 (m, 1H, tentative NHC*H*PO), 4.10 (m, 4H, POC*H*$_2$), 7.30 (broad s, N*H*); LRMS (ESI) calculated for $C_9H_{22}NO_3P$: 223.1337, observed at 224.15 ($M+H^+$).

2.4.1.5. Inhibitor Assays

All assays run with the molecules purchased and synthesized were performed as specified in Section 2.2.2.2., with the exception of Co^{2+} being the metal used. All compounds were dissolved in 3% - 10% DMSO solution, and added in diminishing concentrations starting from 50 uM to the buffer solution, sequentially decreased in 20% decrements down to approximately 0.5 uM. DMSO controls were used to verify that levels of DMSO up to 10% showed no effect on normal activity of MetAP. IC_{50} values were determined using the GraFit6™ software for Windows.

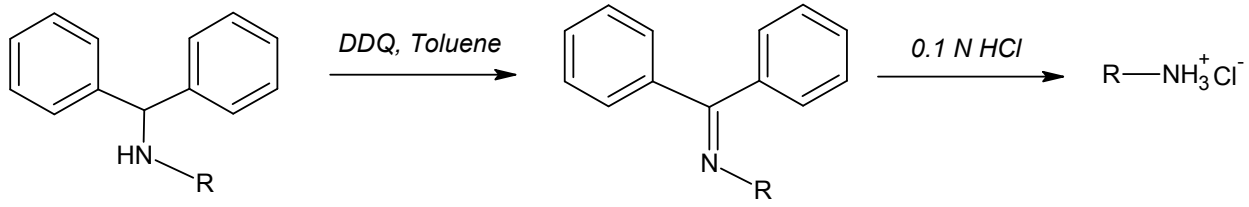
2.4.2. Results and Discussion

2.4.2.1. Synthesis of norleucine aminophosphonate diethyl ester

The protocol of Yadav *et al.* made use of montmorillonite KSF clay as a catalyst, as in more recent times, clay and zeolite catalysts are becoming more commonly used in organic chemistry because they are inexpensive, environmentally compatible, reusable, operationally simple, non-toxic, and non-corrosive while offering greater selectivity and ease of isolation (85). They are quite popular among those performing anionic reactions, Diets-Alder reactions, acid-catalyzed reactions, aromatic chlorination and nitration transformations (85, 86). So, in creating the DPM-protected norleucine phosphonate, we utilized KSF clay in our microwave irradiation synthesis.

After initially synthesizing (diphenylmethylamino)norleucine phosphonate diethyl ester using microwave irradiation, we had to remove the DPM group. This had been previously performed in our laboratory by Dr. Peter Sampson, a former graduate

student, whose results were published (87). At the time, this was done to answer the question of how to readily remove the DPM protecting group, frequently used for its stability in strongly acidic conditions, in situations where mildly acidic conditions must prevail. Though the option of exposing our DPM-protected compound to powerful acid to remove it existed, we could not carry this out because of the possible hydrolysis of the ethylene groups on the phosphonate oxygens. Therefore, we removed DPM employing an oxidative approach, whereby the secondary amine is oxidized to an imine that can then be put in mild acid and the DPM hydrolytically removed (87). This was accomplished through reacting our DPM-protected compound with DDQ, a high oxidation potential molecule that has seen much use in organic synthesis as a dehydrogenation agent [See Scheme 8].



Scheme 8

Though we successfully synthesized the aminophosphonate protected derivative of norleucine, being norleucine aminophosphonate diethyl ester, this reaction took quite a while to optimize. Initially, it was not understood just how integral maintaining an argon atmosphere was to this step. Initial reaction execution was performed by only bubbling argon in to the toluene / molecular sieve solution for 15 minutes, and then carrying out the reaction. This gave shockingly low yields (approximately 5%). Moreover, it was also quite important to consistently maintain the 60 ^{o}C temperature requirement. A heat gun was used first in order to heat the reaction, as it was important

to us not to potentially damage the phosphonate through fragmentation. However, it was not possible to sustain the temperature needed with any kind of stability using this approach. So, once we switched to heating using a water bath, in concert with the argon atmosphere, yields dramatically increased, probably fostered by the persistent argon infusion as well.

Finally, it must be noted that purifying anything involving the DPM group is very, very difficult. Not only does it tend to contaminate every fraction you purify, in spite of how fine you attempt to resolve your purification, but it even remains affixed to the Mass Spectrometer so that samples that do not include DPM still appear to have strong DPM MS signals present. Exceptionally low flow rates were eventually utilized to success, but the entire procedure took an incredibly long period of time, and is inefficient because of this.

2.4.2.2. Phosphorus- and Sulfur-Based Compounds as Inhibitors of MetAP

Viewing Table 2 along with Figure 25 gives the initial impression that the larger and more complex a molecule is, the more effectively it behaves as an inhibitor, with the largest molecule assayed, bis(4-nitrophenyl)phosphate acting as the most powerful compound acting against MetAP catalysis. If this is correct then the greater number of

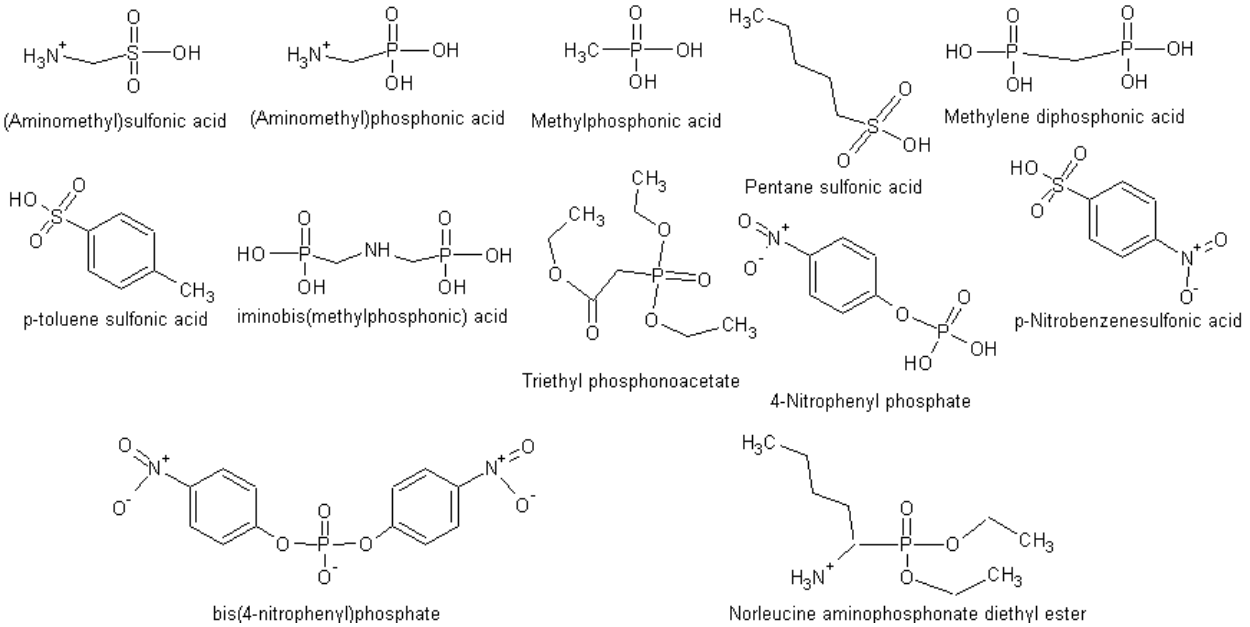


Figure 25. The compounds used as possible inhibitors of MetAP.

Compound	**IC_{50} (μM)**
Bis(4-Nitrophenyl)Phosphate	*1.9 +/- 0.30*
4-Nitrophenyl Phosphate	*3.2 +/- 0.41*
Triethyl Phosphonoacetate	*4.7 +/- 0.67*
Iminobis(methylphosphonic) Acid	*5.2 +/- 0.39*
p-nitrobenzenesulfonic acid	*8.2 +/- 0.72*
Methyl-p-toluene sulfonate	*7.1 +/- 0.45*
Methylene Diphosphonic Acid	*10.2 +/- 0.62*
Pentane Sulfonic Acid	*12.8 +/- 1.0*
Norleucine Aminophosphonate Diethyl Ester	*11.4 +/- 0.69*
(Aminomethyl)-phosphonic Acid	*10.5 +/- 0.47*
Aminomethane Sulfonic Acid	*12.5 +/- 1.1*
Methylphosphonic Acid	*12.9 +/- 0.9*

Table 2. IC_{50}'s (uM) of the phosphorus- and sulfur-based compounds examined as possible inhibitors of MetAP

interactions and bonding events that occur within the active site of an enzyme with a given molecule, the more tight the binding, and the greater inhibitory value that molecule has as an inhibitor. Based on the similar IC_{50} values determined for (aminomethyl)phosphonic acid and the aminomethanesulfonic acid it appears that both these functional groups can contribute somewhat to inhibitor binding. It is possible that

the phosphate-possessing group would be more similar to the transition state produced during a MetAP catalytic event.

It is interesting to note that the racemate of norleucine aminophosphonate diethyl ester was not highly active as an inhibitor of the enzyme using Met-AMC as an assay substrate. Once dissolved (which took a higher concentration of DMSO than the other compounds, and much more tedious mixing, in spite of it being a hydrochloride salt like some of the others commercially purchased) and assayed, it exhibited a fairly weak IC_{50} relative to several of the other compounds, although all compounds were in the micromolar range for their IC_{50} values. However, for our purposes, this compound served as a preliminary norleucine analogue that has a neutral charge on the phosphonate, yet still maintains its tetrahedral structure, which might resemble the tetrahedral intermediate in the enzymatic reaction. However since it lacks the negative charge on the oxygen, future studies should test the monoethyl ester as well as the completely deprotected norleucine phonsphonate). Overall, these results allude to the fact that phosphate / phosphonate and sulphonate-containing compounds are able to inhibit MetAP. Further studies are warranted with these compounds to enhance inhibitory potency which perhaps might be improved by the addition of moieties that resemble a di- or even a tripeptide structure.

2.5. Fluorinated Methionine Analogues

One of the other areas of research our laboratory has interest in is in fluorinated amino acid analogues and their effects on protein structure and function. The

substitution of hydrogen for fluorine confers various changes in methionine's physical properties. Not only does it become more hydrophobic, but the dihedral angle rotary energy barrier is significantly reduced as there is a shortening of the carbon-sulfur bond (F_3C-S) and an increase in the distance of the sulfur-carbon (S-CH_2-R) bond (88). As well, electron and charge density are diminished, causing native state steric interactions with other neighboring molecules to be modified (88). However, it has also been stated that fluorine is "relatively nonsterically demanding", and in certain instances, can replace hydrogen atoms of biological molecules with negligible ramifications structurally, which is also beneficial depending on the system under study (89). Taken together, these findings from altering methionine through fluorine addition allows for the creation of a very unique probe of the MetAP enzyme.

One of the most exciting ways fluorinated-amino acid analogues have been utilized by our group and others is as a means to gather structural information about protein structure and interactions by using ^{19}F Nuclear Magnetic Resonance Spectroscopy (NMR) (89, 90, 91). Our laboratory in particular has been quite involved in studying the production of and the change in structure in biochemically modified proteins, and their interactions with other proteins, with F-Met analogues using this technique. We were inspired to examine whether fluorine incorporation into substrates processed by MetAP would be processed at a rate less than non-fluorinated ones (92, 93). Therefore, we decided to further develop the Met-AMC assay by synthesizing fluorinated analogues of Met-AMC, and assessing whether there was a decrease in catalysis based on the presence of fluorine atoms in the thiomethyl functionality in

methionine, utilizing the fluorogenic Met-AMC assay. Our laboratory has observed decreased fluorinated methionine N-terminal processing for DFM-incorporated in *Pseudomonas aeruginosa* alkaline protease (AprA) (93). A similar finding was reported by Budisa *et al.* at the University of Milan with trifluoromethionine (TFM) in fluorinated green fluorescent protein (92). To date, these are the only studies looking at enzyme activity on fluorinated methionine substrates.

2.5.1. Experimental

2.5.1.1. Materials

Difluoromethionine (DFM) and Trifluoromethionine (TFM) had been previously synthesized in our laboratory by Dr. Mark Vaughan, a former graduate student in our laboratory (94) and by Ignace Moya, a current graduate student in the Honek laboratory.

Anhydrous methylene chloride (>99.8%) and 9-fluorenylmethyl *N*-succinimidyl carbonate (97%), 7-amino-4-methyl coumarin (99%) were obtained from Sigma Aldrich (St. Louis, MO). Anhydrous N,N-dimethyl formamide (DMF) (>99%) was purchased from Caledon Laboratories Ltd. (Georgetown, ON).

The resins PS-carbodiimide, PS-trisamine, and MP-TsOH were obtained from Biotage (Charlotte, NC).

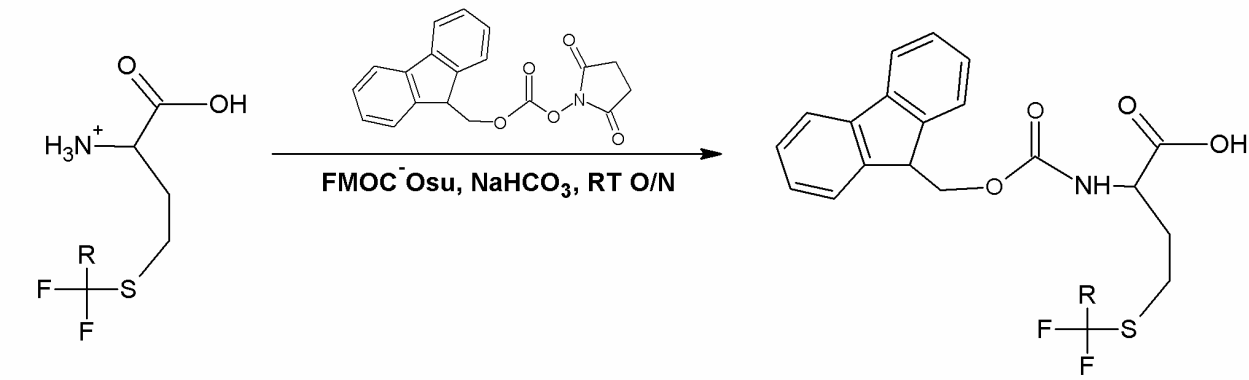
Commercially obtained recombinant *E. coli* Methionine Aminopeptidase was obtained from R&D Systems, Inc. (Minneapolis, MN). The fluorogenic substrate 7-

amido-4-methyl coumarin was acquired from Bacham Bioscience (99%) (King of Prussia, PA). 96-well black opaque microplates were obtained from Corning Inc. (Corning, NY).

All solvents were evaporated *in vacuo* on a rotary evaporator unless otherwise specified.

2.5.1.2. FMOC-Difluoromethionine and FMOC-Trifluoromethionine

FMOC-DFM and FMOC-TFM were previously synthesized in the work of Dr. Mark Vaughan, a former graduate student in our laboratory (94) [Scheme 9]. A 10 ml 1:1 solution of H_2O and acetone was placed in a 25 ml round-bottom flask with a stir



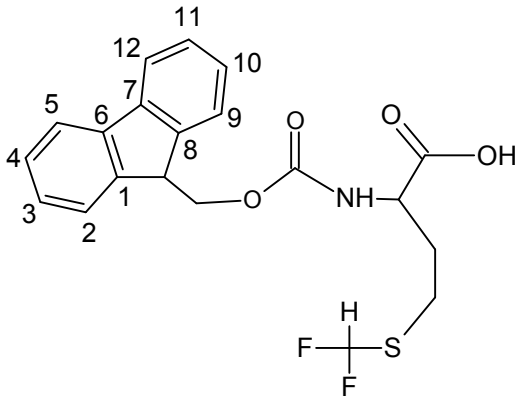
R = **H**, *Difluoromethionine*; **F**, *Trifluoromethionine*

Scheme 9.

bar. Sodium bicarbonate (22.5 mg, 0.265 mmol) was added to function as a base, followed by 50 mg (0.27 mmol) of DFM or 54.5 mg (0.27 mmol) TFM and 90.5 mg (0.27 mmol) 9-fluorenylmethyl *N*-succinimidyl carbonate (Sigma Aldrich, St. Louis, MO) being added to the solution. This reaction mixture was then allowed to sit on ice for 1 hour, and then was stirred at room temperature overnight. Afterwards, it was acidified to pH 2

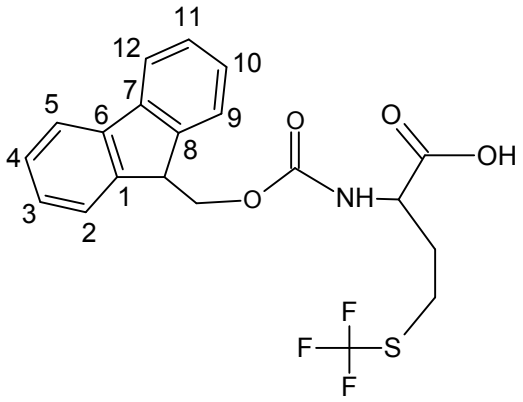
using 0.1 N HCl. Solvents were removed *in vacuo*, and the aqueous layer was extracted with chloroform (3 x 1 ml). The organic layer was then washed with H_2O until reaching a neutral pH. The solution was then dried over anhydrous sodium sulfate and evaporated to give a yellow oil. When checking purity with mass spectrometry and TLC, it appeared that unreacted FMOC and DFM / TFM was still present (Rf ($CHCl_3$/EtOAc 3:7; 0.9 and 0.05). Therefore, FMOC-DFM and FMOC-TFM were treated as >95% pure for the next step of the reaction

The FMOC-DFM and FMOC-TFM compounds created here were spectroscopically identical to the compounds produced by Dr. Mark Vaughan in our laboratory previously (94). Our characterizations of both of the currently synthesized compounds are in agreement with those findings.



FMOC-Difluoromethionine yield was 34.1 mg (85%). R_f ($CHCl_3$/EtOH 3:1) 0.52; ^1H NMR (500 MHz, $CDCl_3$) δ (relative to TMS, 0.00 ppm) 2.06 (m, 1H, $C_\beta H_B$), 2.32 (m, 1H, $C_\beta H_A$), 2.86 (m, 2H, $C_\gamma H_2$), 4.25 (t, 1H, $C_\alpha H$, J = 7.0 Hz), 4.45 (m, 2H,

COCH_2), 4.56 (m, 1H, fluorene-CH-CH_2-)), 5.45 (1H, CONH), 6.81 (t, 1H, C$H$$F_2$, J_{HF} = 56 Hz), 7.33 (2H, arene), 7.42 (2H, arene), 7.58 (2H, arene), 7.77 (2H, arene),10.65 (s, 1H, CO_2H); LRMS (ESI) calculated for $C_{20}H_{19}NO_4F_2S$ (M+) 407.1003 observed at 408.1356 ($M+H^+$) (also observed: (M + Na^+) calculated 430.0901, observed at 430.1009).

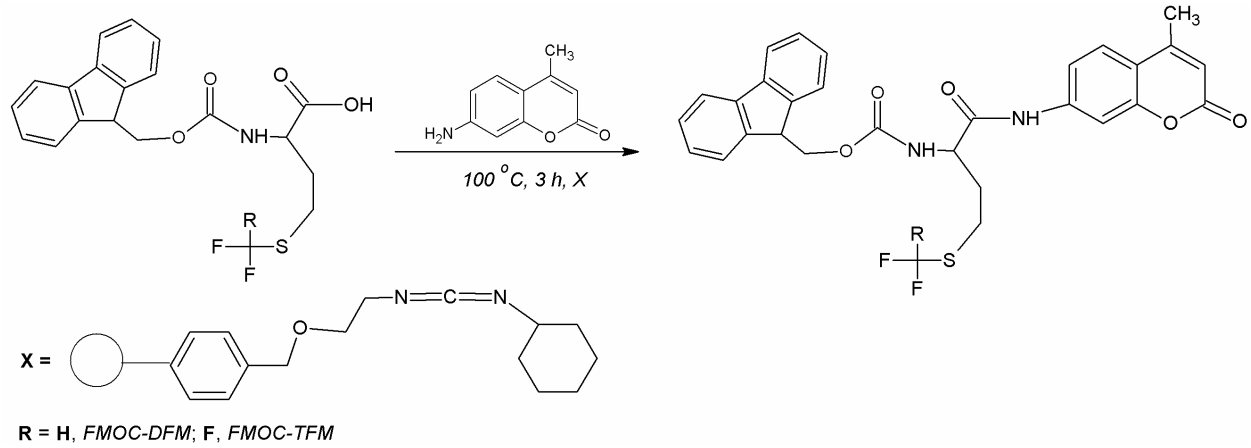


FMOC-Trifluoromethionine yield was 31.5 mg (79%). R_f ($CHCl_3$/EtOH 3:1) 0.54; ^{1}H NMR (500 MHz, $CDCl_3$) δ (relative to TMS, 0.00 ppm) 2.03 (m, 1H, $C_\beta H_B$), 2.34 (m, 1H, $C_\beta H_A$), 2.87 (m, 2H, $C_\gamma H_2$), 4.10 (t, 1H, C_αH, J = 7.0 Hz), 4.40 (m, 2H, COCH_2), 4.55 (m, 1H, fluorene methine), 5.33 (1H, CONH), 7.32 (2H, arene), 7.39 (2H, arene), 7.58 (2H, arene), 7.75 (2H, arene), 10.68 (s, 1H, CO_2H); LRMS (ESI) calculated for $C_{20}H_{18}NO_4F_3S$ (M^+) 425.0909 observed at 426.4223 ($M+H^+$) (also observed: (M + Na^+) calculated 448.0806, observed at 448.0984).

2.5.1.3. FMOC-Difluoromethionine-AMC and FMOC-Trifluoromethionine-AMC

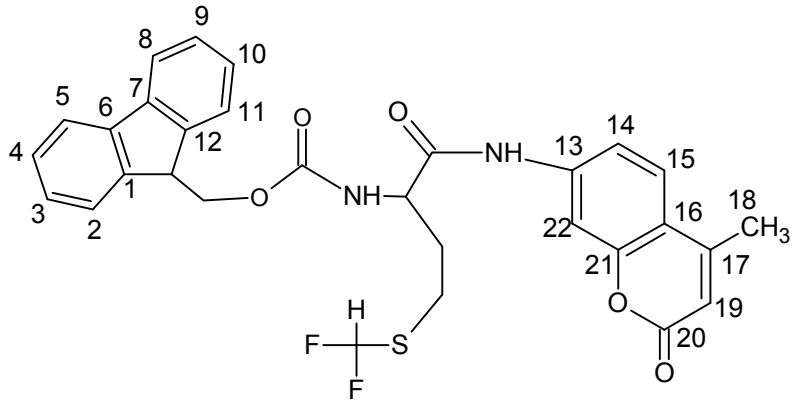
The protected fluorinated analogues, 34.1 mg FMOC-DFM (0.084 mmol) and 33.7 mg (0.079 mmol) FMOC-TFM products from the previous step were dissolved in 3 ml of anhydrous methylene chloride (Sigma Aldrich, St. Louis, MO). PS-carbodiimide

(Biotage, Charlotte, NC; 146.9 mg (2 eq. based on the company's analysis of 3.8 mmol/g of resin) was added. PS-carbodiimide, as reported by Booth and Hodge (95), consists of a coupling agent carbodiimide that is attached to a resin bead where the coupling reaction takes place [Scheme 10]. AMC (14.4 mg, 0.082 mmol) (Sigma Aldrich, St. Louis, MO) was dissolved in 0.3 ml anyhydrous DMF (Caledon Laboratories Ltd., Georgetown, ON), which was then added to the resin / TFM / DFM mixture in a microwave test tube. The complete

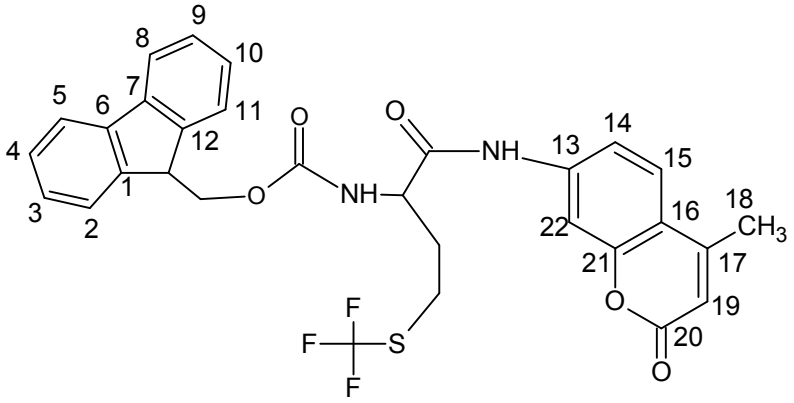


Scheme 10 (X = PS-carbodiimide resin)

reaction mixture was then exposed to microwave irradiation for 3 h at 100 °C in the Biotage Initiator Microwave (Biotage, Charlotte, NC). Previous experiments determined that this temperature and this length of time were optimal to avoid decomposition of the product as well as to take into account the reduced nucleophilicity of the AMC amine.



FMOC-Difluoromethionine-7-amino-4-methyl coumarin: R_f ($CHCl_3$ / EtOAc 3:7) 0.60; contamination with some unreacted AMC which co-purified with the compound. Semi-pure product was used in the next step, which was the removal of the FMOC group followed by purification. ^{1}H NMR (500 MHz, $CDCl_3$) δ (relative to TMS, 0.00 ppm) tentative assignments: 2.23 (m, 1H, $C_\beta H_B$), 2.30 (m, 1H, $C_\beta H_A$), 2.90 (s, 3H, coumarin-CH_3), 2.97 (m, 2H, $C_\gamma H_2$), 4.02 (t, 1H, possible fluorene methine), 4.11 (t, 1H, $C_\alpha H$, J = 7.0 Hz), 4.31 (m, 2H, fluorene-CH-CH_2), 5.95 (s, 1H, possible coumarin arene), 6.35 (broad s, 1H, possible N*H* of amide), 6.75 (t, 1H, C*H*F_2, J_{HF} = 57 Hz), complex arene region 7.23-7.7 unable to deconvolute due to complex mixture, 8.02 (broad s, 1H, possible C_αN*H*); LRMS (ESI) calculated for $C_{30}H_{26}N_2O_5F_2S$ (M^+) 564.1530 observed at 565.1791 ($M+H^+$). (Also observed: (M + Na+) calculated 587.1428, observed at 587.1448; (M + K+) calculated 603.1168, observed at 603.1386).



FMOC-Trifluoromethionine-7-amino-4-methyl coumarin: R_f ($CHCl_3$ / EtOAc 3:7) 0.63; ^{1}H NMR (500 MHz, $CDCl_3$) δ (relative to TMS, 0.00 ppm) 2.10 (m, 2H, $C_\beta H_2$), 2.28 (s, 3H, CH_3), 2.90 (m, 2H, $C_\gamma H_2$), 3.96 (d, 2H, fluorene-CH-CH_2, J = 6 Hz), 4.05 (t, 1H, $C_\alpha H$, J = 6 Hz), 4.30 (m, 1H, fluorene-methine), 5.93 (s, 1H, arene), 7.23-7.35 multiplets (5H, arene) 7.57 (2H, arene), 7.73 (2H, arene), 7.94 (2H, arene), 8.12 (s, 1H, C_αN*H*), 8.83 (s,

broad, N*H*) LRMS (ESI) calculated for $C_{30}H_{25}N_2O_5F_3S$ 582.1436, observed at 583.1713 ($M+H^+$). (Also observed: (M + Na+) calculated 605.1334, observed at 605.1464; (M + K+) calculated 621.1073, observed at 621.1348).

2.5.1.4. Removing Unreacted AMC

In order to remove unreacted AMC, the scavenging resin MP-TsOH (Biotage, Charlotte, NC) was used [Figure 26]. FMOC-DFM-AMC and TFM-FMOC-AMC were dissolved in 5 ml chloroform, to which we added the resin to fill the bottom portion of the round bottom flask, qualitatively. This was then stirred at room temperature until no

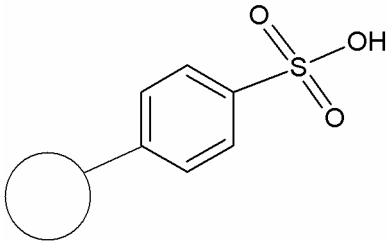


Figure 26. MP-TsOH Resin (Biotage, Charlotte, NC)

AMC was present in the solution as based on visualization using thin-layer chromatography (TLC) (Rf ($CHCl_3$/EtOAc 3:7) of AMC – 0.51, DFM-AMC, 0.06). If AMC was still present, small amounts of the resin was added again, and stirring continued. The reaction was finished after approximately 30 minutes, and the resin removed by vacuum filtration.

2.5.1.5. Removing FMOC

FMOC is a base-sensitive protecting group, so in order to remove it, the PS-trisamine resin was added to the crude FMOC-DFM-AMC and FMOC-TFM-AMC mixtures. Removal of base-sensitive FMOC was done using 0.34 g (10 eq.) of PS-

trisamine resin (Biotage, Charlotte, NC) (based on the company's analysis of 3.9 mmol / g of resin) [Figure 27] an approach suggested before by Booth and Hodge (95) and by personal communication with the technical division of Biotage, Inc. . The reaction mixture containing FMOC-DFM-AMC was dissolved in anhydrous methylene chloride (3mL) (Sigma Aldrich, St. Louis, MO) and placed in a microwave test tube. The resin was then added and the entire mixture was exposed to microwave irradiation for 3 minutes at 100 ^{o}C. TLC confirmed the absence of FMOC in the reaction mixture (Rf ($CHCl_3$/EtOAc 3:7) FMOC – 0.9, DFM-AMC – 0.05)

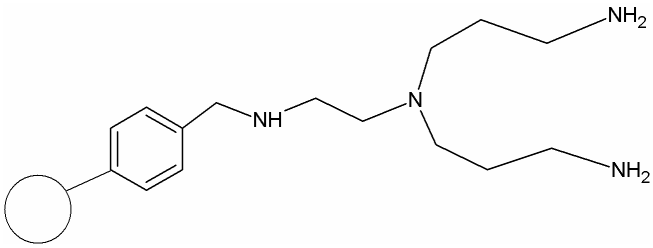
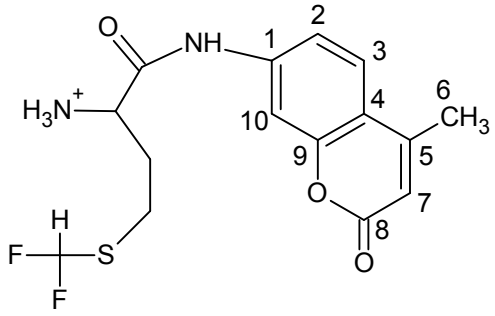


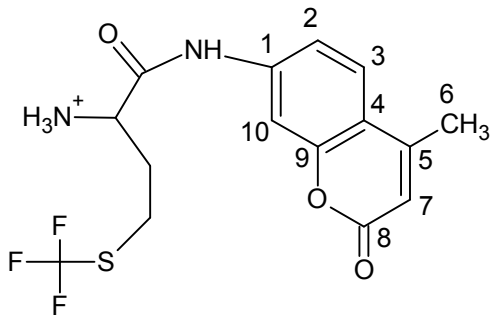
Figure 27. The PS-Trisamine resin from Biotage (Biotage, Charlotte, NC)

Once FMOC was removed, DMF-AMC and TFM-AMC samples were purified using the Biotage Flash System (Biotage, Charlotte, NC) before physical characterization using NMR and mass spectrometry. Unfortunately, this did not adequately purify the compounds, with some impurities still left in the reaction mixture. Whether it was because of the lack of reaction completion with the resins, or simply hydrolyzed AMC confounding the refinement, the compounds were not pure. Therefore, in order to obtain a finer resolution, a miniature gravity column was made using a Pasteur pipette packed with silica gel and cotton. Purification was much more lengthy (at least 8 hours for each compound), but the end result was a pure product. Even still,

the final yields for the truly pure DMF-AMC and TFM-AMC were 1.2% and 1.7% (1.1 mg and 1.7 mg) respectively.



Difluoromethionine-7-amino-4-methyl coumarin yield: 1.1 mg; R_f ($CHCl_3$ / EtOAc 3:7) 0.12; ^{1}H NMR (500 MHz, $CDCl_3$) δ (relative to TMS, 0.00 ppm) 2.76 (m, 2H, $C_\beta H_2$), 2.88 (m, 2H, $C_\gamma H_2$), 2.40 (s, 3H, CH_3), 4.30 (m, 1H, $C_\alpha H$), 6.06 (s,1H, arene), 6.84 (t, 1H, CF_2H, J_{HF} = 56 Hz), 7.55 (m, 1H, arene), 7.70 (m, 1H, arene), 7.98 (1H, arene), 8.16 (broad s, 1H, NH), 9.76 (broad s, 1H, NH); LRMS (ESI) calculated for $C_{15}H_{16}N_2O_3F_2S$ 342.0850, observed at 343.0918 ($M+H^+$).



Trifluoromethionine-7-amino-4-methyl coumarin yield: 1.7 mg; R_f ($CHCl_3$ / EtOAc 3:7) 0.1; ^{1}H NMR (500 MHz, $CDCl_3$) δ (relative to TMS, 0.00 ppm) 2.45 (s, 3H, CH_3), 2.57 (m, 2H, $C_\beta H_2$), 2.67 (m, 2H, $C_\gamma H_2$), 4.26 (m, 1H, $C_\alpha H$), 6.17 (s, 1H, arene), 7.58 (m,1H, arene), 7.66 (m,1H, arene), 7.74 (m,1H, arene), 8.03 (1H, tentative NH_3^+); LRMS (ESI) calculated for $C_{15}H_{15}N_2O_3F_3S$: 360.0755, observed at 361.0734 ($M+H^+$).

2.5.2. DFM-AMC / TFM-AMC Assay

To get a sense of whether our synthesized fluorinated methionine analogues are processed by MetAP with any level of efficiency, we decided to assay them using the Met-AMC assay.

The DFM-AMC / TFM-AMC Assay was performed as outlined in Section 2.2.2.2. in the Met-AMC assay. However, DFM-AMC and TFM-AMC were dissolved in 20% and 18% DMSO solution, and added to the buffer solutions in concentrations up to 1.5 uM. For comparison, Met-AMC was added in the same way. Kinetic constants K_m and k_{cat} were determined using GraFit6^{TM} software for Windows through non-linear regression fitting of the rate v.s. substrate concentration curve (also referred to as the Michaelis-Menten curve).

2.5.3. Results and Discussion

2.5.3.1. Synthesis of DFM-AMC and TFM-AMC

The commonly used fluorenylmethyloxycarbonyl (FMOC) amino protecting group was added to DFM and TFM in order to prevent reactions taking place at the amino terminal of the analogues when linking them to the AMC chromophore at the carboxyl group and additionally to prevent polymerization reactions during the coupling of the chromophore to the amino acid. The chromophore, 7-amino-4-methyl coumarin (AMC), was coupled to FMOC-DFM and FMOC-TFM using the PS-carbodiimide resin to create FMOC-DFM-AMC and FMOC-TFM-AMC. Then using the resins PS-Trisamine and MP-TsOH, the FMOC group of the compounds, and any unreacted AMC, were removed, yielding DFM-AMC and TFM-AMC.

Overall, the reaction was successful and fairly easy to perform, albeit, the purification was quite arduous. Flash chromatography on a large scale did not sufficiently purify DFM-AMC and TFM-AMC for our needs. Therefore, small-scale purification using a gravity column was necessary in order to obtain a finer resolution, and yielding truly pure compounds. Though our final yields DFM-AMC and TFM-AMC were quite low (1.1 mg and 1.7 mg, respectively) they were adequate for our requirements of characterization and utilization in fluorescence assays.

2.5.3.2. Kinetic Parameters for DFM-AMC and TFM-AMC

As Table 3 shows, Met-AMC is superior as a substrate to both DFM-AMC and TFM-AMC. The catalytic efficiency (k_{cat}/K_m) is superior to both analogues, and this was partially expected. It has been seen that L-TFM affects processing of N-terminal substrates adversely (92, 93), and so the fact that in our experiments, it showed only

Substrate	K_m (M)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
DFM-AMC	0.0073 +/- 0.00085	0.01395 +/- 0.00171	1.91
TFM-AMC	0.0094 +/- 0.00091	0.01014 +/- 0.00098	0.541
Met-AMC	0.0051 +/- 0.00042	0.01872 +/- 0.00266	3.67

Table 3. The K_m, k_{cat}, and k_{cat}/K_m of DFM-AMC, TFM-AMC, and MET-AMC.

15% of catalytic processing by MetAP as compared to Met-AMC is not surprising. The additional steric bulk imposed on TFM by the additional fluoride substitution is apparently sufficient enough that even in protein systems engineered like lysozyme bacteriophage lambda (LaL) to express it, it is much more difficult to integrate and is found with fairly low frequency (90). Its poor performance in these assays are consistent with what has been qualitatively observed in protein processing. The DFM-AMC substrate was processed with efficiency in between methionine-AMC and TFM-AMC. It is interesting to note that DFM incorporation into proteins is very high, almost

suggesting that the two fluorines contribute little steric hindrance to the normal biochemical processing of this analogue (90).

2.6. Summary

We have demonstrated various unique means of querying MetAP activity and biochemical character. Divalent metal activation kinetics and the metal-dependance of MetAP has been shown, with Fe^{2+} seeming to be eMetAP's native metal co-factor, as has been reported in the literature (76). With this, we have further developed the fluorogenic assay of MetAP activity with the chromophore Met-AMC. We have also potentially demonstrated an additional 'moonlighting' role of MetAP – phosphatase activity - not previously known, and undoubtedly deserving of further investigation. We have also shown MetAP to be inhibited differentially by various phosphorus- and sulfur-containing compounds, more potently inhibited by those more closely resembling a tetrahedral intermediate. Finally, we have synthesized novel fluorogenic substrates with fluorinated substitutions that have advanced our understanding of the substrate specificity of Met-AMC, and the effect of fluorination on MetAP catalysis, our enzyme showing preference for the di- versus tri-fluorinated substrate.

Chapter 3. In Silico Chemistry and Molecular Modeling

3.1.1. Introduction

One of the most important current uses of molecular modeling is the virtual docking of potential ligands on to protein targets so as to design pharmaceuticals *in silico* by identifying new chemical structures that may potentially bind to a protein active site and to also optimize lead candidate structures (96). Docking algorithms are used to find lead candidates by attempting to identify the optimal binding mode, termed a "pose", of small molecules on the active site of a macromolecular target. Thus, virtual docking software attempts to find compounds that more strongly bind to a given protein target than the natural substrate [Figure 28]. With this, the biochemical reaction catalyzed by the macromolecular complex can be inhibited or modified in the hopes of

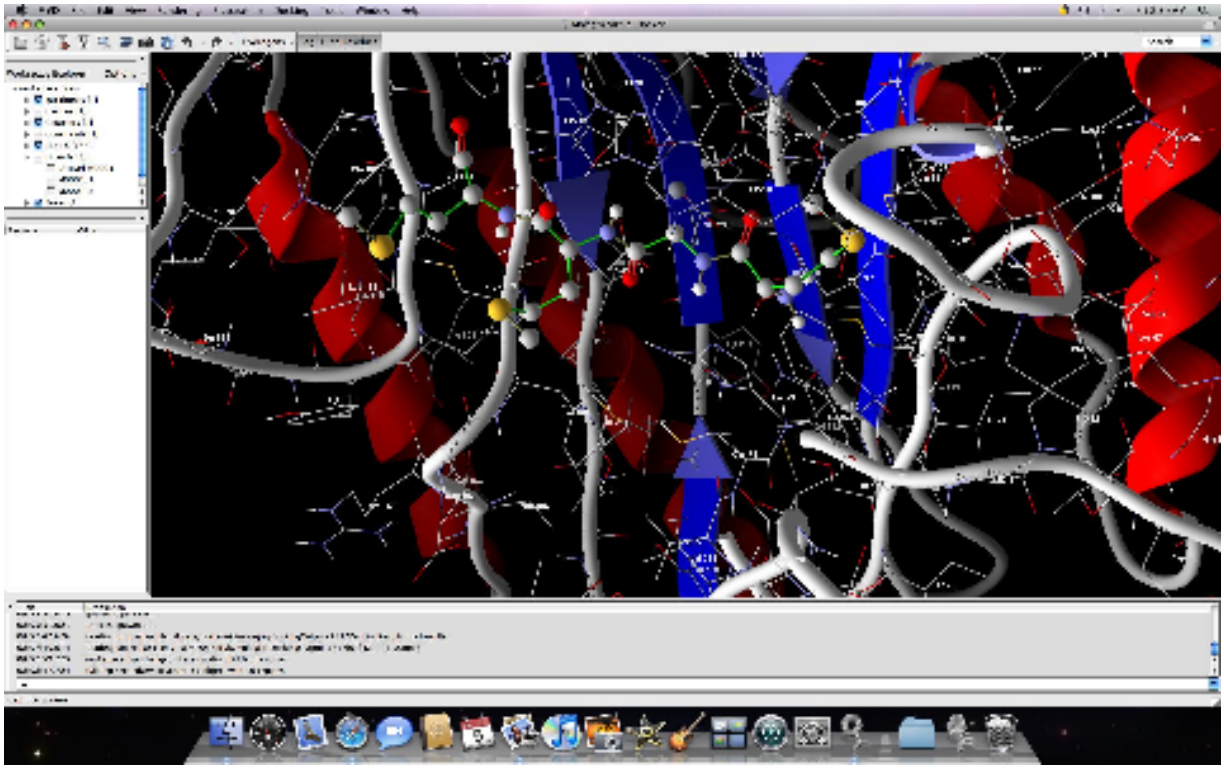


Figure 28. A desktop view of the Molegro (Molegro ApS, Aarhus, Denmark) virtual docking software suite, used in our *in silico* chemistry experiments.

eliminating the deleterious effect of continued reaction.

High throughput screening methods are useful to measure the activity of a large number of compounds against a specific biochemical target (96, 97). Unfortunately, this practice tends to be costly, and extremely laborious. However, with the availability of 3D structures, molecular docking simulations can provide a very useful and practical tool in drug-discovery (97). Using an *in silico* approach, one is permitted a faster and cheaper strategy to identify promising drug candidates by screening virtual entire compound libraries in the hopes of finding promising leads. Once this portion of the process is complete, laboratory experimentation, be it synthesis, biochemical assays, toxicological assessment, or even clinical trials, can be conducted to further inspect the drug candidates identified.

Virtual screening is based on software docking methods that use an energy-based scoring algorithm, called a “scoring function”, to identify the energetically most favorable ligand conformation when bound to the target (97). In a broad sense, lower energy scores represent better protein−ligand bindings compared to higher energy scores. Molecular docking can be viewed as an optimization tool, meant to identify the conformations of compounds that have the tightest binding.

As mentioned earlier in this work, molecular modeling and *in silico* chemistry has become more ubiquitous within the drug discovery process over the last 20 years, having produced and expanded upon many potential candidate drugs [Figure 29] (97).

For example, when viewing Figure 29, taken from a review by Shoichete and his collaborators, one sees the target tumorgenic enzyme CDK4, a kinase in cell cycle G1 phase progression, where the lead inhibitory compounds available for this enzyme were immensely improved using molecular modeling techniques (97). The best inhibitor available prior to the studies by the original research group, Honma and colleagues at Merck in Japan, had an IC_{50} of 44 uM (98). Through molecular modeling, and modifications to the lead compounds, Honma *et. al.* ultimately came up with *N*-[(9b*R*)-5-oxo-2,3,5,9b-tetrahydro-1*H*-pyrrolo[2,1-*a*]isoindol-9-yl]-*N'*-pyridin-2-ylurea, with an IC_{50} of 0.011 uM when tested against CDK4 (98). This is an absolutely amazing improvement that clearly demonstrates the merits of computer-aided drug design and *in silico* chemistry in general. Such findings have not only been fostered by the availability of greater computing power at a lower cost, but also because of the greater availability of compound libraries available both publicly and from the private sector (97). Work has typically focused on enzymes and their inhibition (97, 99). The inhibitors discovered usually are quite novel, having little similarity to the enzyme's known ligands, and most having affinities in the low-micromolar range (97). One of the most exciting exceptions, however, has been new sulfonamide inhibitors of carbonic anhydrase II – though they are much more potent (nanomolar potency), they are not completely novel (97). In any event, the high number of lead candidates discovered has undoubtedly been enhanced by the availability of many new docking methods.

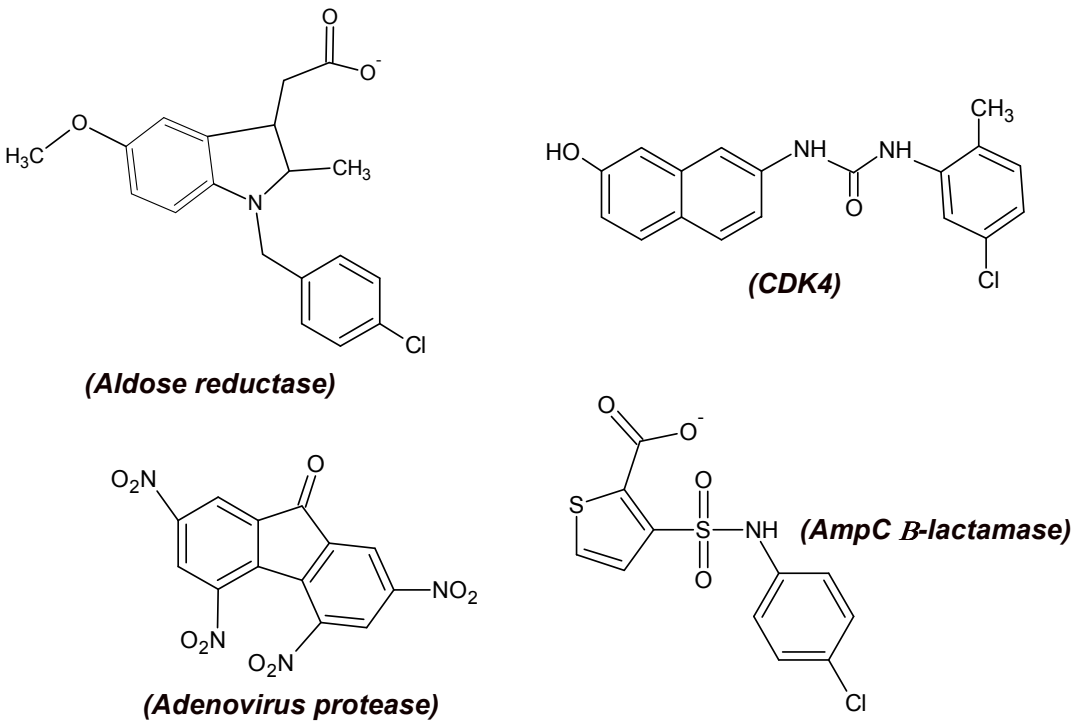


Figure 29. Examples of various lead candidates discovered using virtual screening in the last decade. Enzyme inhibited by the compound shown is in brackets. <*Adapted from reference (97)*>

It is important to note that one of the original criticisms molecular docking faced was that it did not sufficiently take into account flexibility of the residues in the active site of the protein, and therefore, the reality of a ligand-substrate binding event (99). A different complexion is placed upon the matter entirely when the software evaluates the ligand and macromolecular active site as dynamic and in flux, rather than rigid and static. Contemporary software does indeed factor flexibility into scoring, however, this creates a much higher number of potential ligand poses and protein side-chain conformations to be considered. Though computing power has increased dramatically since the beginnings of molecular modeling, docking remains a very challenging dilemma as the high number of possible docking conformations (within the field called "poses") prevents a systematic brute-force approach (100). Therefore, in order to

properly assess potential conformations when docking ligands upon proteins *in silico*, thereby properly handling flexibility, search heuristics are required. It should be noted however that the most critical weaknesses in docking programs is the inability of these programs to faithfully capture the intermolecular interaction potentials between the ligand and the protein. These concerns will be expanded upon later in this chapter.

3.1.2. Molecular Docking and Virtual Screening

Molegro (Molegro ApS, Aarhus, Denmark) was chosen for the docking and virtual screening assays for a variety of reasons. One reason was the affordability. Some molecular modeling software can run many thousands of dollars, making it more difficult for laboratories outside of industry to gain access to these programs. As well, Molegro is available in a very easy to use format for Apple and PC platforms, which means that those unfamiliar with operating systems based on UNIX are not prohibited from using it. However, most importantly, Molegro is centered around the MolDock scoring function, the newest and potentially one of the most accurate and reliable scoring function currently available (100) [Figure 30].

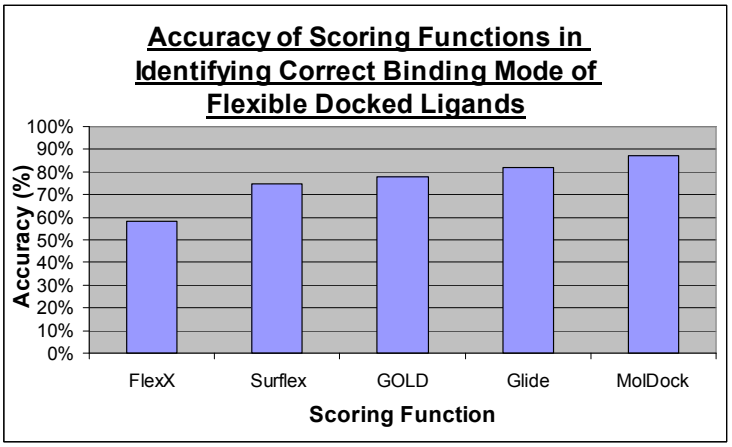


Figure 30. A comparison of docking accuracy with a selection of widely used docking algorithms. <*Adapted from reference (100)*>

Released in 2005, MolDock has been shown to identify the correct binding mode of flexible ligands 87% of the time in a docking study with a defined subset of ligands and proteins (100). This result surpasses the reliability of other, widely used scoring functions, such as GOLD, found to be reliable 78% of the time in the same study, or Glide, which had 82% accuracy in distinguishing the correcting binding mode of the same flexible ligands (100). Nevertheless it should be kept in mind that these values could vary for these programs depending upon the chosen protein-ligand sets used in the comparison.

Various mathematical algorithms are currently in use that search through thousands of different poses during the molecular docking process to find the most ideal conformation balanced with the most realistic. Based on the attempt to find the most "optimized" solution, these algorithms are referred to as "heuristic search algorithms" (100). The most popular heuristic search algorithms in molecular docking are simulated annealing, based on combining the parameters involved in the annealing of solids and statistical mechanics (100); tabu search, Glover's landmark heuristic search algorithm which is a procedure that iteratively moves from one solution to another potentially better solution within a specific search space (100); and evolutionary algorithms, which much like their name denotes, provide a set of solutions to a problem and those most fit to answer the problem "outlive" other solutions too costly for their own survival with increasing complexity applied to the solutions incrementally (100).

MolDock is based on a new search heuristic algorithm, called “guided differential evolution” (100). Guided differential evolution algorithm is a hybrid, combining the differential evolution with a cavity prediction algorithm. Differential evolution (DE), introduced by Storn and Price at Berkeley in 1995 (101), is a mathematical technique used to find a global minimum, based on computational evolutionary methods, amongst large amounts of multivariable and multifactorial data, and does so with precise probability (101). This procedure has seen previous success in its application to molecular docking, through its handling of flexible ligands and large macromolecular complexes (100). DE, combined with the use of predicted cavities within sizeable protein molecules by restricting the search space to a particular region, allows for quick and reliable identification of potential binding modes, or “poses” of lowest energy (100).

MolDock is an extended version of a previous evolutionary-based scoring function called GEMDOCK, introduced by Yang and colleagues in 2004, with the addition of a piecewise linear potential (PLP) from Gehlhaar (100). The docking scoring function is given by Equation 1:

$$\text{Eq. 1. } E_{Score} = E_{Inter} + E_{Intra}$$

where E_{Inter} is the ligand-protein interaction energy (100) (Equation 2):

$$\text{Eq.2 . } E_{Inter} = \sum_{i \in \text{ ligand}} \sum_{j \in \text{ protein}} \left[E_{PLP}(r_{ij}) + \frac{332.0\ q_i q_j}{4r^2_{ij}} \right]$$

This equation is applied to all heavy atoms in the ligand and all heavy atoms in the protein, including any cofactor atoms and water molecule atoms that might be present (100). The E_{PLP} term is the piecewise linear potential. It uses two different sets of

parameters: one set for approximating the steric (van der Waals) term between atoms and the other stronger potential for hydrogen bonds (100). The PLP is given by Equation 3:

$$\text{Eq. 3. } E_{Total} = E_{H\text{-}Bond} + E_{Inter} + E_{Repulsion} + E_{Contact}$$

$E_{H\text{-}Bond}$ is the contribution to the total energy by hydrogen bonding, E_{Inter} is the contribution to energy by the total intermolecular forces, $E_{Repulsion}$ is the internal atomic repulsion and $E_{contact}$ is the atomic contribution of contact potentials (100). This powerful equation however, is extended with a further expansion of the definition of hydrogen bonding energy, which with MolDock takes into account hydrogen bond directionality (previously ignored with earlier scoring functions that account for hydrogen bonding energy), given by the H-Factor (Equation 4):

$$\text{Eq. 4. } H_{Factor} = \Phi\,(\Delta_{D\text{-}H\text{-}A};\ 90^{o};\ 100^{o}) \cdot \Phi\,(\Delta_{H\text{-}A\text{-}AA};\ 90^{o};\ 100^{o}) \cdot \Phi\,(\Delta_{D\text{-}A\text{-}AA};\ 90^{o};\ 150^{o})$$

In the above equation, a bond is considered a hydrogen bond if one of the atoms can donate a hydrogen atom and the other atom can accept it (100). The E_{PLP} in Equation 2 only depends on the distance between atoms. Therefore, as in Equation 4, to take into account the directionality of hydrogen bonding, the geometry of the hydrogen bond is examined, and the following factor, H_{Factor}, is multiplied by the PLP hydrogen bond strength (100). AA (the Aacceptor Antecedent) denotes a heavy atom connected to the acceptor (A), while D denotes the donor, and H is the specific donated hydrogen atom (100). The ramp function Φ is defined as $\Phi\,(A;\ A_{min};\ A_{max}) = 0$ for $A <= A_{min}$ and $\Phi\,(A;\ A_{min};\ A_{max}) = 1$ for $A >= A_{max}$ and is linearly interpolated between these values for $A_{min} < A < A_{max}$ (100). In the case that it is impossible to calculate one of these factors, it is left out (100). This pertains to cases like hydroxyl groups, where the

particular location of the hydrogen is not investigated during docking, and these two factors cannot be calculated (100).

In equation 2, the second term describes the electrostatic interactions between charged atoms (100). It is a Coulomb potential with a distance-dependent dielectric constant given by D(r) 4r (100). The numerical value of 332.0 constrains the units of the electrostatic energy to kcal/mol (100). For dealing with distances that are less than 2.0 Å, the electrostatic energy is cut off at the level corresponding to a distance of 2.0 Å to ensure that no energy contribution can be higher than the clash penalty. It is important to note that in spite of the electrostatic energy contribution having the theoretically predicted magnitude, all other energy terms are experimentally based, and the actual binding affinity is not necessarily compatible with total energy (100).

E_{Intra} is given by the following (Equation 5):

$$\text{Eq. 5. } E_{Intra} = \sum_{i \in \text{ligand}} \sum_{j \in \text{protein}} E_{PLP}(r_{ij}) + \sum_{\text{flexible bonds}} A[1 - \cos(m \cdot \Theta - \Theta_o)] + E_{clash}$$

The double summation is between all atom pairs in the ligand, excluding atom pairs that are connected by two bonds or less (100). The second term is a torsional energy term, parameterized according to the hybridization types of the atoms involved in bonding and the term responsible for taking into account bond rotability both in ligand structure and side chain residue conformational variability. Θ, the torsional angle of the bond, is not necessarily determined uniquely in each instance, but at times, under appropriate circumstances, is generally defined based on known bond geometry (100). Molegro treats two types of bonds as rotatable, namely sp^2-sp^3 bonds and sp^3-sp^3 bonds.

According, for sp2-sp3 bonds, the Θ_o in Equation 5 is 0.0, m is 6 and A is 1.5, while for sp3-sp3 bonds, Θ_o is π, m is 3 and A is 6. The last term, E_{clash}, assigns a penalty if the distance between two atoms (more than two bonds apart) is less than 2.0 Å. Therefore, the E_{clash} term is meant to penalize ligand conformations that are not possible (100). This issue is important especially with regards to side chain flexibility, a feature of Molegro that will be discussed in greater detail later.

3.1.3. General Experimental

3.1.3.1. Materials

Throughout our series of molecular modeling experiments, the following software was used: MacSpartan 8.0 (Wavefunction Inc., Irvine, CA) for structure based ligand building and MAC Molegro v3.2 (Molegro ApS, Aarhus, Denmark) for docking and virtual screening of compounds against MetAP. All experiments were performed on an Apple Mac Pro with two quad-core Intel Xeon "Nehalem" processors (4 x 2 processors = 8 processors total) with 16GB RAM.

3.1.3.2. Docking With Molegro

Docking in our experiments using Molegro was performed as follows: the target protein was searched for on the Protein Data Bank (www.rcsb.org), and the 3D structure, in the form of a PDB file, was downloaded. This PDB file was then imported into the Molegro work space, usually excluding waters as is the standard procedure in most virtual screening and molecular docking studies (96). Ligands built in Spartan and minimized at the B3LYP / 6-31+G* level were then imported in to the same work space.

A built-in feature of Molegro is to “prepare” the molecules for docking - correct any errors in the PDB file (ie. looking for missing bonds) and the Spartan-built ligand, while adding hydrogens to the protein structure, so this was performed next. Afterwards, the ‘cavity’, or area of the 3D structure most likely to bind a ligand, was detected automatically using Molegro (however, it is worth noting that one also has the option of manually selecting the area of the protein you would like to dock with ligands). Following this, docking was performed in Molegro. The only setting typically changed from default when using this feature is the number of iterations (changed from ‘15’ to ‘100’ – the maximum permitted), and the default search space, or the area the software uses on the protein to dock ligands – this was changed from the default centre of the protein to the specific cavity detected earlier. After running the docking simulation, the results were imported, and evaluated. Docking scores are reported by MolDock in kcal/mol form.

The default reporting of MolDock scores do not factor in all the various elements that Molegro uses in arriving at the lowest energy pose of a ligand. For example, in spite of its factoring in forces like van der Waal’s forces and hydrophobic interactions in its algorithm, it does not rank the poses generated initially using some of these intuitively important scoring factors. This is probably because the developers of Molegro wanted to provide the user with the option of ranking poses with whatever factors they wanted to choose, which would probably be different depending on the aims of the experiments and the types of compounds and target proteins used in docking. Therefore, in addition to a comparison with the “ReRank score” Molegro

provides (however, once more, this does not provide a score complete in terms of all potential factors involved in the scoring process), all relevant scoring factors, such as short- and long-term interactions, and as to be seen later, in specific cases water molecule interactions, were selected, and the poses re-ranked to obtain the most accurate pose with the most accurate lowest energy possible. The pose with the lowest energy will be referred to henceforth as the lowest energy pose (LEP).

3.1.3.3. Standard Error

Before proceeding any further, it was important to find out whether the MolDock algorithm gave the same score every time if the ligands and 3D structures, as well as settings, remained equal. Intuitively, one might believe that this is the case, as many kinds of software often function like a calculator – the same numbers put in and performed with the same arithmetic always yields the exact same answer. However, during preliminary phases of using Molegro, it became apparent that this was not the case, and in fact each time a value for the energy of an LEP was given after a docking simulation, it was different from the previous one, even if the pose topographically appeared to be the same. To test this definitively, the PDB file 3MAT, a structure deposited by one of our previous collaborators Matthews and Lowther at the University of Oregon, was imported into Molegro, and its bestatin-based inhibitor, (3R)-amino-(2S)-hydroxyheptanoyl-L-Ala-L-Leu-L-Val-L-Phe-OMe [Fig] was docked with the di-metalated (Co(II)) eMetAP structure according to the procedure in Section 3.3.1.2. (*discussed later*) , with the LEP saved into the workspace, and the docking process taking place again with the original ligand. This process was repeated five times.

The energetic score for each pose of the highest energy obtained after the simulations run were: -130.102, -129.323, -126.407, -131.799, and -135.196. Therefore, when docking the same ligand against the same substrate in replicate, MolDock actually gives a slightly different score after every run. Calculating the standard error gives a scoring variation of approximately 3%. This seemed strange initially, as it appeared visually that the LEPs in each run were in the identical position. However, upon further examination, it was the case that there are slight deviations in position within the workspace only discernable through precise calculations using inborn functions of the Molegro workspace to measure discrete distances and angles. The random nature of the iteration process, a necessity built-in to the algorithm in order to prevent bias, is something that probably also contributes to this result (100). In any case, for all molecular docking simulations, to be accurate, all values reported are the average of runs in triplicate, and a standard error is assigned where appropriate. It should be added that this variation in energetics in a docking algorithm is likely also produced in all docking software although little is mentioned in the scientific literature concerning the magnitude of these variations.

3.2. Assessing Molegro Accuracy

In spite of the report that MolDock is the most advanced and reliable docking algorithm currently available (96, 100, 102), based on previous experience in our lab, and the generally acknowledged consensus that molecular modeling, and *in silico* chemistry in general, is still not as reliable as the 'lab bench'; it was important to identify the confidence level of our Molegro docking results. Therefore, a three-tier approach

was developed in order to examine the accuracy of Molegro in predicting previously published results. In this way, it was believed that if Molegro was faithful in reproducing the results, or obtaining very similar results, to published articles, then our own experiments with the software could be taken more seriously, and our docking predictions might have increased predictive power.

3.2.1. Experiment #1: Comparing Previously Published Inhibitor Potency With Predicted Molegro Scoring

In this experiment, it was our desire to assess whether in fact Molegro and the MolDock scoring function could predict the order of inhibitor potency of previously published small molecule inhibitors. If it could, then we would expect that the order of the energetic scores provided by Molegro would correspond to the order of potency, given as an IC_{50}, of the inhibitors.

3.2.1.1. Experimental

A publication was selected based on fulfilling several criteria: 1) The target of inhibition had to be MetAP; 2) The inhibitors used had to be metal dependant or take into account the coordinating role that metals play in the active site; 3) There was a range of inhibition between the weakest and strongest compounds relative to IC_{50}. Therefore, the paper by Hu and colleagues (2004), "*Peptidyl Hydroxamic Acids as methionine aminopeptidase inhibitors*" was chosen as it satisfied the requirements we put forth in designing this docking experiment (103).

In this paper, five peptide-based compounds were synthesized, 1a – 1e [Figure 31]. These compounds were drawn and energetically minimized in Spartan at the B3LYP / 6-31+G* level, and then imported into Molegro where they were docked within the active site of a di-metalated (Co(II)) MetAP according to the procedure in Section 3.1.3.2. The PDB file 1CS3 was used for this experiment as the assumption by the authors was that the two metal ions were required for inhibitory action based on the chelating nature of the chemical agents designed. Therefore, it was our feeling that for true faithfulness to the lab bench and experimental reality, it was important to keep the authors assumptions with the *in silico* model being designed in Molegro. However, as a control, we also tested these same inhibitors on a mono-metalated species of MetAP with Mn(II) in the active site [PDB Code: 2GTX], to assess whether these virtual inhibitors could differentiate between the virtual active sites with different metal ions presented in MetAP's catalytic centre. The incorporation of metal ions in to the active site calculations adds another level of complexity to docking calculations since electrostatic charge and metal ion sterics need be taken into consideration. Molegro takes into account these aspects of molecular interactions in three ways. The van der Waals radii (VdW) is used, as well as the covalent radii between atoms of the protein substrate, ligands, and the metals themselves. The specific radii for the metals we docked, Mn(II) had a VdW radii of 1.7 Å, while the covalent radii is 1.35 Å. Molegro also assigns a charge to metal ions, +2 in the case of divalent metals that would be found in MetAP's active site. These factors, along with the fact that two Co(II) metals were present in 1CS3 while only one Mn(II) ion was present in the 2GTX PDB file, were expected to affect the docking results of this particular *in silico* experiment.

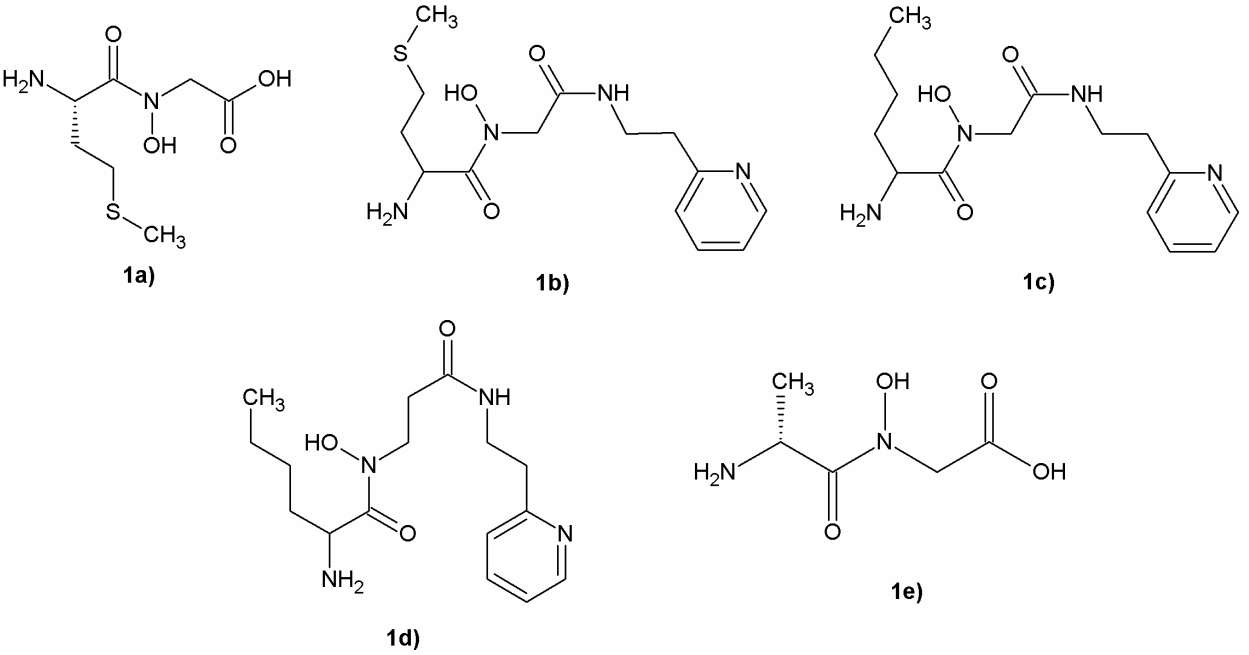


Figure 31. Peptidyl hydroxamic acid inhibitors used in Hu and co-workers' paper adapted for use in our molecular docking studies
<Adapted from reference (103)>

3.2.1.2. Results and Discussion

	Co(II)-metalated eMetAP	***Molegro Score (kcal/mol)***	***Molegro Score (kcal/mol)***
Compound	**IC_{50} (uM)**	**Mn (II)**	**Co (II)**
1a	170 +/- 16	-72.437	-64.954
1b	2.5 +/- 1.0	-137.382	-123.461
1c	59 +/- 6	-114.342	-119.549
1d	48 +/- 10	-112.129	-113.475
1e	8.8 +/- 2.7	-119.193	-120.057

Table 4. Comparison of Hu *et al.*'s compounds' IC_{50} values and their Molegro scores with different metal ions.

Two patterns become clear after viewing Table 4.:*1) In this instance, Molegro appears reasonably predictive in ordering of the compounds that have IC_{50} values far*

apart; *2) Molegro is not totally reliable when comparing compounds that have IC_{50} values fairly close to one another*.

In analyzing the calculations for compounds 1a-c, it appears that if they are ranked in order of IC_{50}, the order is 1b < 1c < 1a. This is also the result if the compounds are arranged by either Molegro score provided in Table 4. Therefore, with the previous assumption in mind, that a lower Molegro score (more negative) is correlated to a tighter binding, one can see that with IC_{50}'s reasonably far apart, ie. 170 uM versus 2.5 uM, it is possible to have confidence in the suggestion that Molegro can be capable of predicting lead candidate compounds.

In the case of IC_{50}s that are quantitatively similar, as in the case of compounds 1c and 1d, despite 1d having a lower IC_{50} value, its Molegro scores were greater in both Mn(II)- and Co(II)-metalated enzymes, implying that the inhibitory potential would be lesser, when in experimentally the opposite is the case. This was expected. It has been our experience within our laboratory that compounds at the upper and lower echelons of inhibitory potential are quickly discerned with virtual screening methods, whereas compounds that lie in the middle range (in this case, between 8.8 and 59 uM) are not separated with the same ease.

There was not much of a difference in Molegro scoring for utilizing different 3D protein structures with different metals in the active site. In spite of subtle differences in actual energetic scores, most probably due to the changes of orientation in the active

site between the di-metalated Co(II) and mono-metalated Mn(II) forms (27), the differences did not impact the ranking as previously described, both in the cases of IC_{50} values and molecular docking rankings.

Ultimately, we were satisfactorily assured that after performing this comparison of Molegro's results with that of peer-reviewed literature, our experience would not be a computerized simulation that did not at all represent what actually happens in the process of 'wet science'. Rather, we obtained initial evidence that as a tool for screening strong candidates for therapeutic usage out of a larger group, Molegro, and its MolDock scoring function, could be of use as a preliminary evaluatory technique given the caveats just mentioned.

3.2.2. Experiment #2: Comparing the Spatial Orientation of Ligands Bound Within the Active Site of a 3D Protein Structure With Docked Ligands in Molegro

Next, it was important to evaluate whether the docking performed by Molegro would yield docked structures similar to the 3D structures deposited in the Protein Databank. Previous literature findings (104), along with experience in our laboratory, has shown that on occasion, the docked structure is almost completely opposite in orientation to the original ligand-bound 3D structure, something that will be elaborated upon later on this in this Chapter. However, at that time, it was absolutely necessary to be sure that the docked structures obtained and the resultant energies acquired through MolDock scoring were legitimate and realistic, especially when compared to the published structures and their own energetics. Otherwise, the docking scores could

potentially be irrelevant if the docking conformation does not relate to what actually happens during crystallography and inhibition in solution.

3.2.2.1. Experimental

A structure obtained in our lab in conjunction with Matthews and Lowther at the University of Oregon was used (PDB code: 1CS3), namely of eMetAP bound to difluoromethionine (DFM). This was then compared to a Spartan-built duplicate of the difluoromethionine docked to the same PDB file, albeit without the original probe present. Docking was performed as in Section 3.1.3.2., first with the original DFM ligand that is included in the original PDB file, and then with the Spartan-made version.

3.2.2.2. Results and Discussion

Initially, no structure obtained through docking matched the ligand-eMetAP conformation in the original PDB file. Therefore, a modified version of the docking procedure had to be employed. Molegro has a feature that allows the user to add in constraints, user-specified parameters that influence the behaviour of the ligand docking within the active site of 3D structure, to the work space. In this case, the constraint to be considered was distance, specifically between particular atoms within the work space. Based on the mechanism of MetAP, it was essential to account for the metal ion within the active site, as its coordination to the substrate during catalysis is paramount in the chemical mechanism of the enzyme. In the case of the PDB file ICS3 structure, Co(II) is present as the metal ion. Although two Co^{2+} ions are found in the ICS3

structure, neither appeared proximal to the carboxyl oxygen atoms of the inhibitor in these docked poses, a structure that would be expected based on the mechanism of the enzyme. Therefore, a distance constraint was added that dictated the carboxyl of DFM could be no greater than 2.8 Å away from one of the metal ions (the precise details to be discussed in a later section), which is within the limits seen in the 3D structures of most 3D MetAP structures deposited thus far. Once this was done, docking was performed as before, and a more realistic portrayal of the *in silico* built DFM ligand was seen docked with MetAP, as compared to the original 3D structure with DFM as the ligand [Figure 32]. As can be seen from the panels in Figure 32, the location of DFM, both the original and that prepared with Spartan, are almost identical. The actual

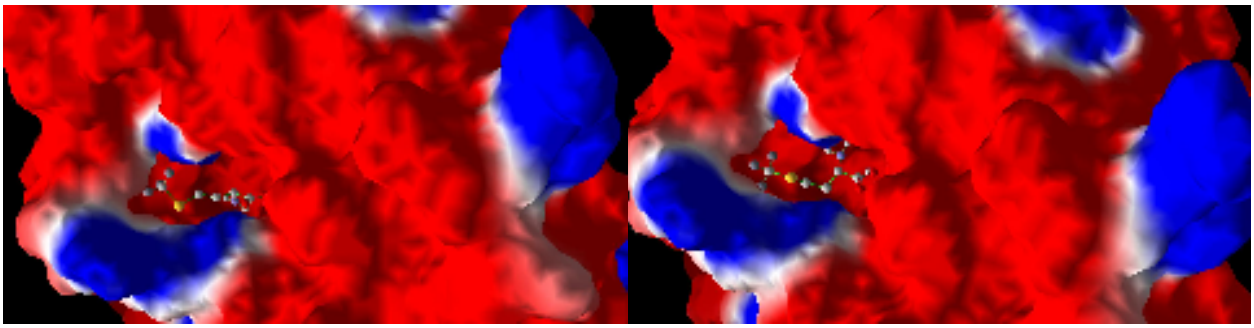


Figure 32. The active site of MetAP containing DFM in A) Original PDB File with DFM ligand; B) Spartan-made DFM docked with MetAP (PDB code: 1CS3)

difference in distance between the two is approximately 0.15 Å in rms value based on Molegro's RMSD matrix feature, which calculates differences in distance between ligands bound to a substrate. As well, it can be seen when viewing Figure 32, that in spite of ligand flexibility, both the original DFM and the DFM created with Spartan exhibit almost identical conformations. This is further corroborated by the scores obtained with the MolDock scoring function. The original DFM ligand was scored at -90.15 +/- 2.19 kcal/mol for the LEP, while the Spartan-made version's LEP was scored at -88.93 +/- 1.96 kcal/mol. Therefore, with the addition of constraints accounted for,

Molegro is reliable in terms of docking ligands in the same position and conformation to a protein of which a 3D structure is available.

3.2.3. Experiment #3: Predicting Biochemical Literature Findings with Molegro

The final test of Molegro came in the form of examining whether Molegro was capable of corroborating the findings of previously published work with respect to specific biochemical features of enzyme structure. In theory, modifying the structure of the active site would have effects upon the ability of molecules to bind. Though Molegro is only capable of docking ligands based on the information the user provides, we felt that the results should also be able to sufficiently support biochemical findings rooted in active site structure, as MolDock scoring is partially based on the electrostatic interactions and concomitant energetics of the macromolecular protein involved. As well, Molegro possesses a unique feature allowing the user to mutate residues anywhere in the protein, and determining whether those mutations alter the energy of a ligand-protein interaction. Therefore, it was of interest to investigate whether mutating key residues in the active site known to be integral to MetAP catalysis had any effect on the binding energy of a ligand within the active site of MetAP.

3.2.3.1. Experimental

The substrate 4-nitrophenyl phosphate (4NP), a classically used substrate with alkaline phosphatases since the 1940's and underscored by the work of Attias in the the 1970's (105, 106), was selected not only to investigate whether it, having been studied with so many enzymes before, would in fact be capable of binding to MetAP *in silico*,

but also whether it could act as a substrate or inhibitor of MAP (see Chapter 2). Therefore, 4NP was drawn and minimzed at the B3LYP / 6-31+G* level in Spartan and the resulting structure was imported in to Molegro [Figure 33]. Next, a mutation was made within Molegro at residue His 79 of MetAP, believed to be an important residue in the MetAP's active site during NME (6, 19). Specifically, the mutation changed the histidine residue to a glycine. Docking was then performed as described previously in Section 3.1.3.2., with the only modification to the standard procedure being the constant for distance was added based on the observation that it was critical in order to obtain a realistic 'pose' (the conformation of a lead candidate) in the active site.

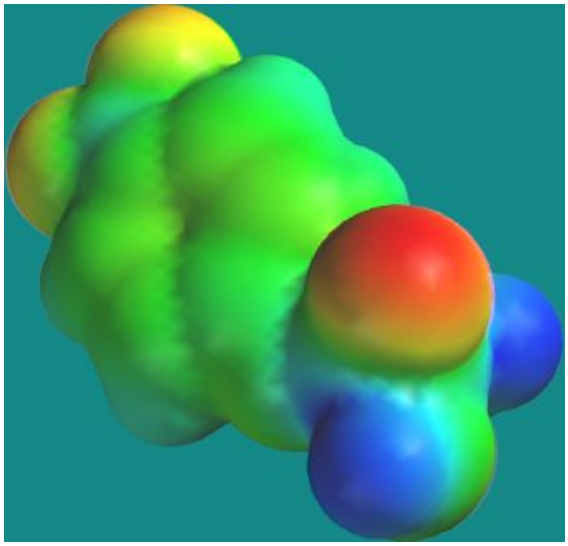


Figure 33. The electrostatic potential of 4-nitrophenyl phosphate (4NP) mapped on to the electron density at the B3LYP / 6-31+G* level in MacSpartan 8.0.

3.2.3.2. Results and Discussion

Upon examing the results, it was found that the WT MetAP bound to 4NP *in silico* exhibited an energy of -115.7 +/- 4.218 kcal/mol for the LEP (again, presumably the tightest binding). In the MetAPH79G mutant, the energy was -61.950 +/- kcal/mol. Clearly, such a drastic change in His79 rendered the 4NP ligand almost 50% less able

to bind to MetAP's active site in this computer simulation. This is most likely due to a loss of important electrostatic interactions in the active site of MetAP with the ligand. These interactions, as mentioned earlier, are believed to be important for orienting the substrate properly during catalysis, and though this evidence does not prove the importance of His79, it nicely mirrors the findings in the literature, and further demonstrates the reliability of Molegro as a virtual screening tool.

Evidently, based on these lines of inquiry, our examination of the Molegro docking software demonstrated some of its important capabilities (and limitations!) to explore ligand-protein binding interactions and provided a fairly accurate qualitative assessment of compound binding within the active site of MetAP. Therefore, we felt it was possible to proceed further with a certain level of confidence to continue using Molegro to dock compounds used in actual biochemical assays and assess their ability to bind to the active site of MetAP.

3.3. Sulfur- and Phosphorus-Based Probes of MetAP Revisited In Silico

As seen in Chapter 2, various phosphorus- and sulfur- containing molecules were used to assess their inhibitory activity against MetAP using the fluorogenic Met-AMC assay, as had been previously described (69). It was of interest to us to assess whether in fact these probes behaved similarly *in silico* as they did *in vitro*, as some of the previous experiments in this chapter have attempted to ascertain in a broad sense. Similiarly as before, we were particularly interested in the relationship sulfur-containing chemical moieties possess with the active site of MetAP, like the natural methionine

substrates, as well as the phosphonate behaviour as a tetrahedral intermediate during catalysis (46), hence the selection of these compounds. Not only do the results give insight into the active site of MetAP during catalytic functioning, but they also provide some possible information into the mechanism of MetAP, and what drives enzymatic activity during NME. Finally, because of the use of two differentially metalated active sites in the docking simulations, it was interesting to inspect whether the type of metal, and the number of ions present, had any effect on the ability of the compounds to bind to the active site of eMetAP. It is quite apparent in the literature that compounds may exhibit specificity to a particular metalated form of the enzyme, such as the Fe^{2+}-metalated form of MAP (25). This was especially appealing to us because of our belief that contrary to previous reports, MetAP is based on a mono-metalated mechanism, not a di-metaled one, in spite of these metalloproteases' catalytic centres being often referred to as the "di-metal center" (6, 19).

3.3.1. Experimental

All ligands were built in MacSpartan, minimized using a B3YLP / 6-31+G* level set of calculations and then the three dimensional structures imported into Molegro [Figure 34 and Figure 35]. In addition, the PDB files 2GTX (containing a single Mn(II)) or 1CS3 (containing two Co(II)) were also brought into Molegro to serve as the protein used in the docking experiments. Docking was performed as per the procedure specified in Section 3.1.3.2. but the number of maximum iterations was input to be 100, rather than 25. Docking calculations were then begun, and took approximately four hours per run.

3.3.2. Results and Discussion

It is interesting to see when viewing Tables 5 and 6 that the more complex and large a molecule is, the lower the energy is and thus, a presumably stronger binding affinity is predicted to exist between the compound and active site of MetAP [Figures 33 and 34]. This is probably due to the fact that the larger the compound, the greater number of interactions are present between the protein and the compound, in the form of hydrogen bonding, van der Waals interactions, hydrophobic interactions and ionic bonding. The phosphate compounds serve as a case in point, where both the Co(II) form and the Mn(II) form of the enzyme exhibit the greatest binding affinity with BNPP and 4NP. This is well correlated with the findings of Chapter 2 in Section 2.3.2.2., where the extremely low kinetic speed of catalysis of MetAP upon both these compounds was seen. When taken with the findings of these docking experiments, perhaps the reason for their unique behaviour was that their high binding affinity contributes to the low turnover of the enzyme and therefore, lower k_{cat} and V_{max} values. In addition, the enzyme will contribute transition state stabilization when acting upon the normal substrate, and MetAP may not contribute optimal transition state stabilization to these unusual substrates.

Compounds	LEP (kcal/mol)
Bis(4-Nitrophenyl)Phosphate	-115.7350
4-Nitrophenyl Phosphate	-89.4953
Triethyl Phosphonoacetate	-85.9436
Iminobis(methylphosphonic) Acid	-85.9206
p-Nitrobenzenesulfonic Acid	-80.3912
Methyl-p-toluene sulfonate	-73.6429
Methylene Diphosphonic Acid	-73.5820
Pentane Sulfonic Acid	-72.2908
Norleucine Aminophophonate Diethyl Ester	-68.5079
(Aminomethyl)-phosphonic Acid	-56.5613
Aminomethane Sulfonic Acid	-55.3685
Methylphosphonic Acid	-45.7636

Table 5. LEP of phosphorus- and sulfur-based compounds docked with Molegro using PDB file 2GTX containing a single Mn(II) ion. The phosphonate and sulfonate groups were protonated during docking.

Compounds	LEP (kcal/mol)
Bis(4-Nitrophenyl)Phosphate	-120.903
4-Nitrophenyl Phosphate	-98.084
Triethyl Phosphonoacetate	-92.397
Iminobis(methylphosphonic) Acid	-91.588
p-nitrobenzenesulfonic acid	-79.326
Methyl-p-toluene sulfonate	-68.959
Methylene Diphosphonic Acid	-66.166
Pentane Sulfonic Acid	-64.546
Norleucine Aminophosphonate Diethyl Ester	-63.423
(Aminomethyl)-phosphonic Acid	-52.998
Aminomethane Sulfonic Acid	-51.774
Methylphosphonic Acid	-39.901

Table 6. LEP of phosphorus- and sulfur-based compounds docked with Molegro using PDB file 1CS3 containing two Co(II) ions. The phosphonate and sulfonate groups were protonated during docking.

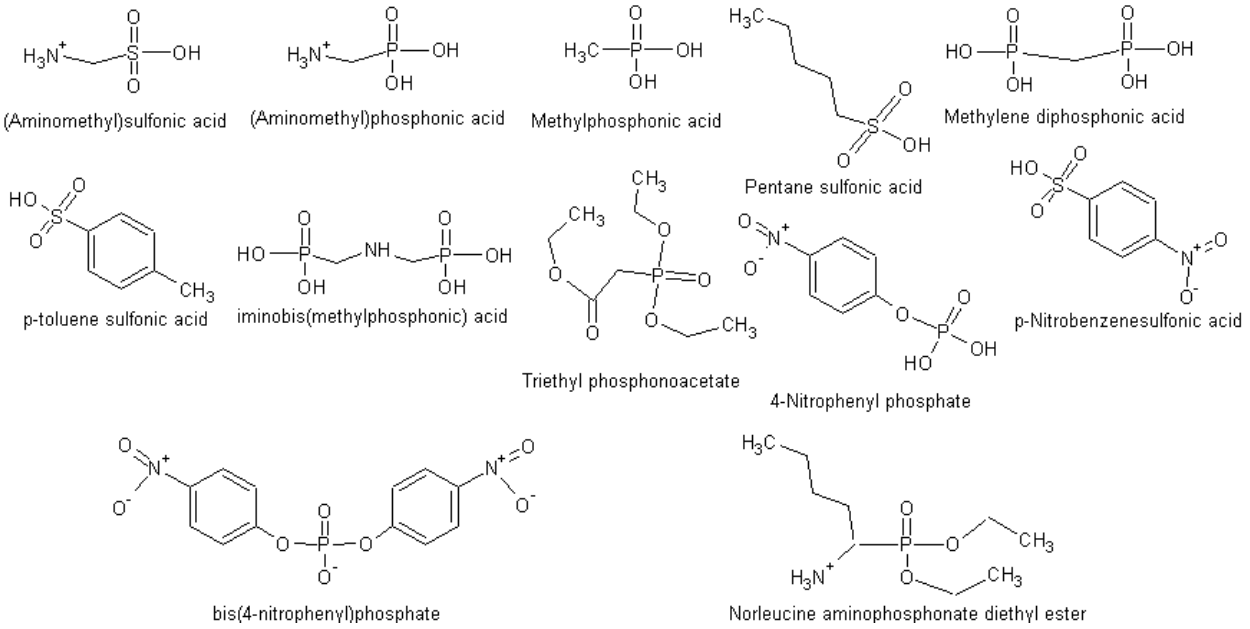


Figure 34. Line drawings of compounds from Chapter 2 used for the molecular docking studies using Molegro.

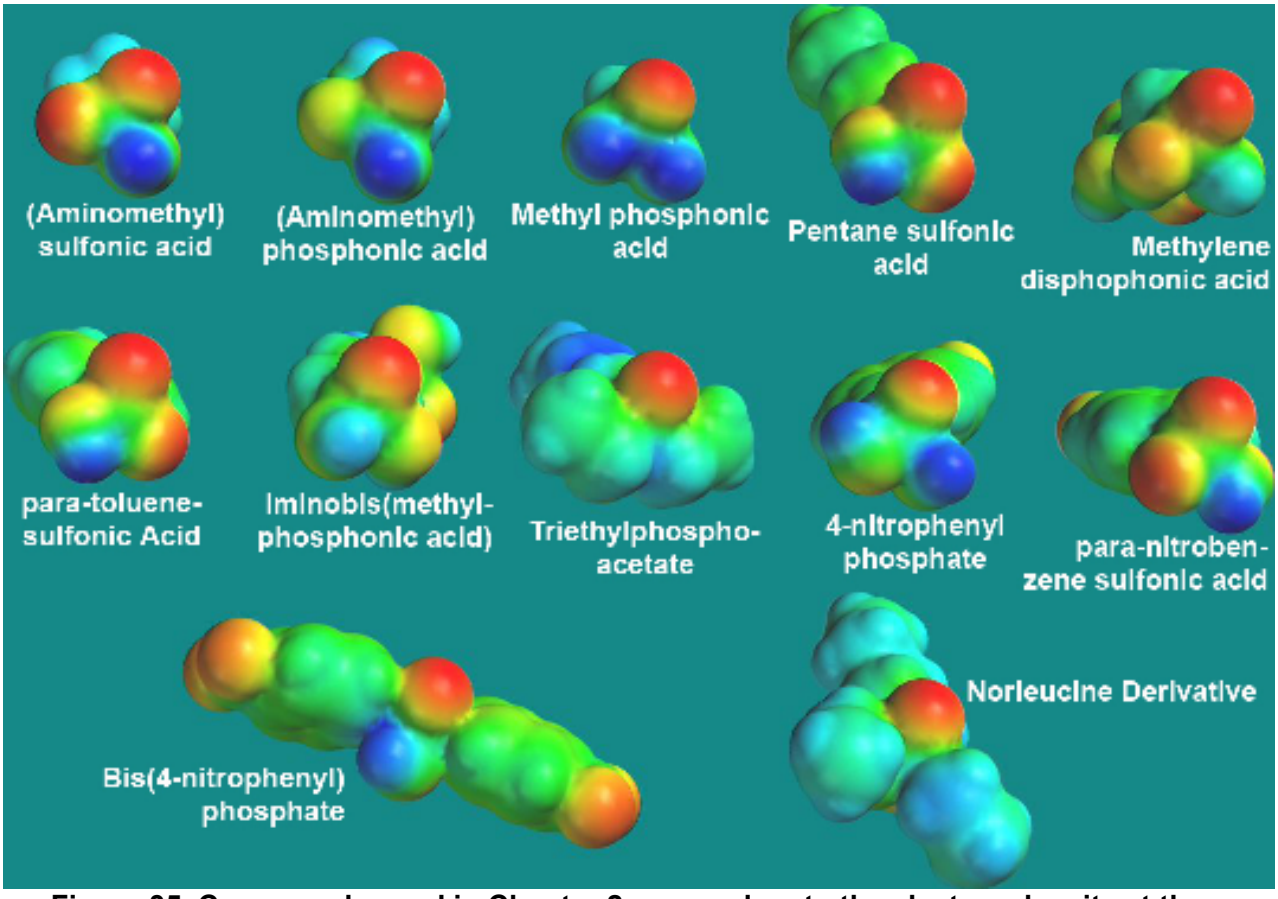
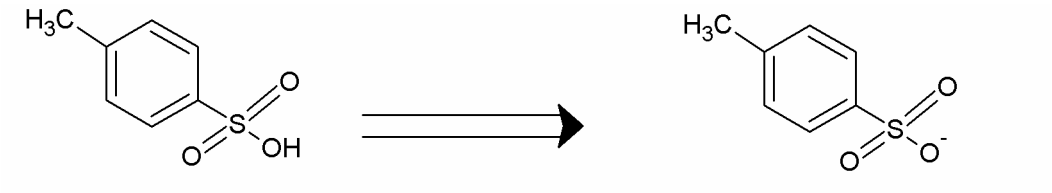


Figure 35. Compounds used in Chapter 2, mapped on to the electron density at the B3LYP / 6-31+G* level in MacSpartan 8.0.

Both the sulfonates and phosphonates, as well as the phosphates, demonstrate binding affinities and it depends on other parts of the chemical structure to improve upon the interactions of these chemical moieties with the enzyme. It also seems to be evident that as compounds become larger, LEP scores become lower. For example, when one docks aminomethane sulfonic acid versus p-nitrobenzene sulfonic acid, lower energetic scores are recorded (again, potentially corresponding to what may be tighter binding). This is interesting to deliberate because in general, MetAP k_{cat} values increase as peptide size increases, inferring greater enzymatic turnover of substrate (78). The tighter binding observed with the large substrate would typically correlate with a decreased ability to rapidly process substrate. In the case of 4NP and BNPP, this may correlate with the fact that they, being phosphate diesters (RO-P(O_2)-OR), might resemble the tetrahedral intermediate moreso and that is the reason for their highest ranking in terms of binding, both *in silico*, and as seen in Chapter 2, *in vitro* as well.

3.3.3. Ionization States

The next step was to look into whether the ionization states of the phosphorus- and sulfur-based compounds affected binding [see Figure 36 for an example]. *A priori*, it was assumed that changing the ionization state would lower energetic scoring and therefore, result in greater binding affinity based on the enhancement of electrostatic interactions fostered by such a modification. With this, it is important to take into account the pKa for the compounds we used. The pKa of sulfonic acid is -2.6, the pKa of a phosphonate is 2.75 and the pKa of phosphates is 1.92 (the second pKa being 6.5), all indicating that deprotonation at physiological pH would be quite likely to occur.



p-toluene sulfonic acid p-toluene sulfonic acid (anion)

Figure 36. The negative ionization state of p-toluene sulfonic acid, which was docked in Molegro

3.3.3.1. Experimental

Ionization state changes were made either in Spartan, through removal of the hydrogen atom and applying a negative (or doubly negative) charge to the compound built, or in Molegro where one can manually change the charge and number of hydrogens present on an atom. Docking was described in Section 3.1.3.2.

3.3.3.2. Results and Discussion

As can be seen in Tables 7 and 8, the differences in energy of the lowest pose between neutral molecules and those conferred with anionic or dianionic character is small. To use the example given in Figure 36, with p-nitrobenzenesulfonic acid, the neutral species' MolDock score is -73.642, while the negatively charged ion's score is -76.496, a 3-4% difference. This is different than what was expected, where we had believed that differentially charged species would be exceptionally lower in energy (and tighter binding) than neutral molecules docked in the active site. However, when bearing in mind the standard error between docking simulations is roughly 3% alone, it is apparently the case that ionization states alone do not play much of a role in determining the molecular docking binding affinities, and would not dramatically alter the choice of lead candidates in a virtual screen of a large library. We also believe this to

Compounds	Negatively-Charged Species Overall (-1) (LEP)	Negatively Charged Species Overall (-2) (LEP)
Bis(4-Nitrophenyl)Phosphate	-118.342	-120.327
4-Nitrophenyl Phosphate	-100.993	-102.103
Triethyl Phosphonoacetate	-88.249	*N/A*
Iminobis(methylphosphonic) Acid	-87.104	-89.312
p-nitrobenzenesulfonic acid	-83.15	*N/A*
Methyl-p-toluene sulfonate	-77.643	*N/A*
Methylene Diphosphonic Acid	-75.328	-77.613
Pentane Sulfonic Acid	-74.117	*N/A*
Norleucine Aminophosphonate Diethyl Ester	-71.441	-74.001
(Aminomethyl)-phosphonic Acid	-59.322	-60.949
Aminomethane Sulfonic Acid	-57.316	-58.107
Methylphosphonic Acid	-47.978	*N/A*

Table 7. LEP of phosphorus- and sulfur-based compounds docked with Molegro using PDB file 2GTX containing a single Mn(II) ion having different ionization states. (N/A = not present in dianion form).

Compounds	Negatively-Charged Species Overall (-1) (LEP	Negatively Charged Species Overall (-2) (LEP
Bis(4-Nitrophenyl)Phosphate	-122.144	-124.582
4-Nitrophenyl Phosphate	-92.315	-95.316
Triethyl Phosphonoacetate	-95.018	*N/A*
Iminobis(methylphosphonic) Acid	-94.235	-95.573
p-nitrobenzenesulfonic acid	-81.447	*N/A*
Methyl-p-toluene sulfonate	-71.288	*N/A*
Methylene Diphosphonic Acid	-68.519	-70.226
Pentane Sulfonic Acid	-66.384	*N/A*
Norleucine Aminophosphonate Diethyl Ester	-66.496	-67.941
(Aminomethyl)-phosphonic Acid	-55.782	-57.387
Aminomethane Sulfonic Acid	-54.697	*N/A*
Methylphosphonic Acid	-42.066	-43.899

Table 8. LEP of phosphorus- and sulfur-based compounds docked with Molegro using PDB file 1CC3 containing two Co(II) ions having different ionization states. (N/A = not present in dianion form).

be in part supported by Molegro's handling of charge interactions in general. The MolDock score in Molegro typically treats charge interactions as 'soft', meaning the energetic penalty (or reward) is minor for electrostatic clashes and interactions. This would easily explain the minute changes between the ionization states when ranking docked ligands.

At this point, it is also important to note we decided that it was no longer important to continue to dock ligands in the active site of MetAP that would contain Co(II) ions, specifically in PDB file 1CS3. This conclusion was based on the fact that the general trends observed to this point, like phosphate superiority in binding overall, remained the same when docking with either PDB file. As well, there was not much of a change in numerical values when comparing the results between the Mn(II) active site docking and Co(II) active site docking. These factors, combined with our ultimate belief that MetAP catalysis functions in a mono-metalated scheme, made us discontinue docking with the di-metalated Co(II) enzyme. Therefore, from this point forward, docking was performed exclusively with the PDB file 2GTX.

3.4. Addition of Constraints: Making In Silico Assays More Realistic

Based on the findings of the previous section which incorporated the native ionization states of the compounds docked with MetAP into the virtual screening process, it became further relevant to see whether adjusting other parameters, in an effort to further integrate more realism into molecular docking, would alter docking results. As is the case in molecular docking and *in silico* chemistry in general, the

results provided by software like Molegro have to be interpreted with the knowledge that in spite of whatever one does in terms of modeling molecular phenomenon, nothing replaces wet lab work. Therefore, in spite of the interesting results that have been obtained through the experiments performed thus far, in order to make the modeling as faithful to reality as possible, as well as to obtain an improved understanding on the real strength of Molegro's modeling abilities, various parameters were modified to foster greater fidelity to the binding assays performed in Chapter 2. Within the molecular work space, these parameters are introduced as "constraints", which basically infers that one is creating limitations to confine specific elements of the search space MolDock works within when arriving at LEP of a docked ligand.

In addition, it was visibly noted that at times, similar to what was mentioned in Section 3.2.2., the LEP would be completely unrepresentative of what was actually occurring in the active site based on the protein X-ray structure. In fact, for several of the phosphorus- and sulfur- containing molecules used in this work, the position of the molecule would almost be completely opposite in conformation to the ligand deposited with the protein in the 3D structure obtained from the Protein Data Bank, even when the ligands were quite similar (104). Although this phenomena has been observed in other studies (104), it did not fit well with our intent in reproducing reality as best as possible within the confines of molecular modeling. Therefore, after a range of modifications to the Molegro work space in the hopes of discovering what would eliminate this problem, the constraints described below were found to best correct this dilemma.

The constraints used in our experiments with Molegro are also of interest as previous literature using Molegro has not utilized them in their virtual screening. Nine papers, including the publication that introduced the MolDock scoring function by Thomsen and Christensen in 2006 (100) have so far used Molegro in their virtual docking studies to different ends. However, none of the published articles went beyond the default settings when performing their *in silico* experiments. It is was our belief prior to adding the constraints that though the "initial screen" (the term we dubbed to refer to the basic docking performed without the addition of constraints and modifications to the Molegro work space) provided useful and valuable information regarding compounds binding within the active site of MetAP, more information that was potentially more realistic, based on adding previously observed experimental data through constraints, would be functionally of greater value. For example, relying on the 'initial screen' means that the user is also implying the active site of the enzyme is completely static, and no residue movement or larger, general structural motif mobility is observed. This situation could be interpreted as implausible, as even the most amateur enzymologist knows of examples of conformational changes observed during most substrate binding events during enzyme catalysis. Therefore, in order to evaluate whether our assumptions were true, it was necessary to further understand how constraints affected the Molegro docking progress, and how they function within the context of molecular modeling in general.

Each constraint was added, followed by the next constraint in the sequence described below such that once a constraint was added, the results were kept and the

subsequent constraint was put in to the Molegro work space. This means that following the first instance below of constraint implementation, subsequent results are a sum of total constraints in use within the Molegro work space during virtual docking experiments.

3.4.1. Distance Constraints

As mentioned earlier, when docking certain ligands into MetAP's active site, the LEP, which would seemingly represent the conformation most tightly bound to the active site, visibly appeared to be almost diametrically opposite in conformation to that observed in the deposited structure. An example of this can be seen in Figure 37, where in contrast to Panel A, Panel B shows the Spartan-built aminomethyl phosphonic

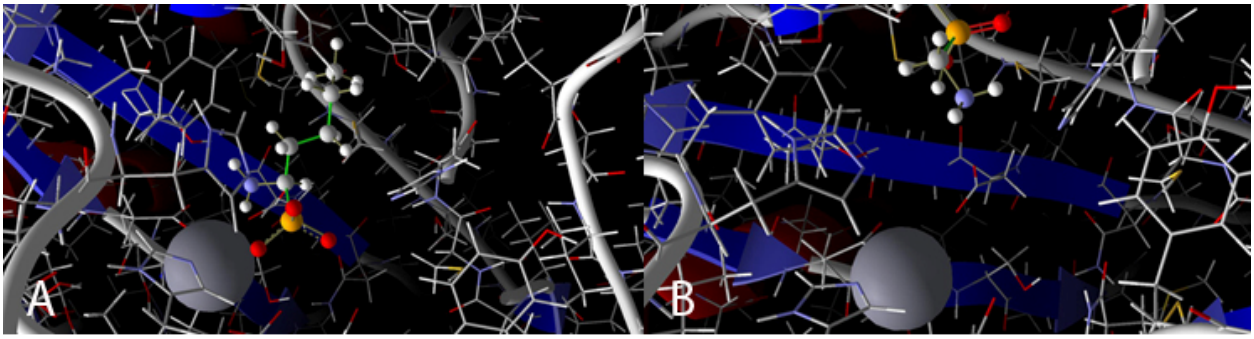


Figure 37. The active site of MetAP [PDB code: 2GTX] with Mn(II) ion (Grey, in Space-fill for emphasis) with A) Included Norleucine ligand; b) Spartan-made Aminomethylphosphonic acid ligand docked using Molegro (initial scan).

acid oriented a way that does not conform to what should be observed. The phosphonate group, for one, where a negative charge would be present and therefore attracted to the positive charge of Mn^{2+}, appears to be directed away from this area and instead, the methyl group is proximal to the Mn(II), which would not be the case based on the general rules of inorganic chemistry. One pose in each of the multiple runs was found by MolDock that did more closely bear resemblance to the norleucine ligand's

docked position. However, they were different in terms of docked location by an average rms value of approximately 10 Å (found using the RMSD Matrix feature in Molegro), and their LEPs were between 15 to 20 percent higher, indicating weaker binding. Therefore, the software has found a pose of best fit topologically, without taking into account reality, where in spite of MolDock being considered the best scoring function currently available (100), it apparently has mistaken various interactions to occur when in fact they would not.

One of the constraints that can be added to the Molegro workspace is the "distance constraint". As alluded to earlier, the distance constraint functions to limit the ability of a specific atom to deviate a specific distance away from another atom the user prescribes. For example, if it was previously determined that a specific group on a ligand remains close to a certain residue of an active site, then one can choose that group on the ligand, and specify the maximum distance allowed it to be oriented away from a specific atom on the residue in question. In this case, it would be relevant with regard to His79 or His178, both known to be proximal to the Met group of newly synthesized peptides for their roles in nucleophilic attack and coordination, respectively (19).

3.4.1.1. Experimental

In our experiments, a distance constraint was created such that sulfate and phosphonate, or phosphate groups in the case of 4NP and BNPP, would not be allowed to be more than 2.8 Å away from the Mn(II) metal ions present in the active site. This limitation was based on literature precedents (26, 107), as well as the actual distance

measured in the original PDB files between similar groups on the ligands deposited and the metal ions present. We were also interested in whether there would be a significant change in energy scores of the ligands as a result of implementing this constraint. Docking was then performed as in the procedure set forth in Section 3.1.3.2.

3.4.1.2. Results and Discussion

From Table 9, it can be seen that except for the phosphate compounds 4NP and BNPP, the distance constraint did not change very much energetically for the compounds used in this work, except BNPP which was drastically reduced. This outlying result is probably based on the massive size of BNPP and its ability to no longer be docked in any position possible with the introduction of this constraint. In terms of binding strength order, the only compounds that changed positions were p-nitrobenzenesulfonic acid and methyl-p-toluene sulfonate, albeit, the difference between them in LEP was marginal. However, the actual appearance of the docked ligand (in

Compounds	LEP (kcal/mol)
Bis(4-Nitrophenyl)Phosphate	-99.532
4-Nitrophenyl Phosphate	-76.966
Triethyl Phosphonoacetate	-73.911
Iminobis(methylphosphonic) Acid	-73.892
Methyl-p-toluene sulfonate	-71.156
p-nitrobenzenesulfonic acid	-67.959
Methylene Diphosphonic Acid	-66.169
Pentane Sulfonic Acid	-63.785
Norleucine Aminophosphonate Diethyl Ester	-60.287
(Aminomethyl)-phosphonic Acid	-53.733
Aminomethane Sulfonic Acid	-52.601
Methylphosphonic Acid	-43.475

Table 9. LEP of phosphorus- and sulfur-based compounds docked with Molegro using PDB file 2GTX containing a single Mn(II) ion after application of a distance constraint centered around the Mn(II) metal ion.

the form of the lowest energy pose) became much more realistic, and closely resembled the original ligand provided in the deposited 3D structures. This was observed in all compounds docked, where functional groups such as the phosphonate or sulfonate group were appropriately positioned relative to the metal ion, and the alkyl groups (if present) arranged expectedly from sources of electrostatic charge [Figure 38]. For one, it would be an unfortunate situation if the energies within the newly constrained simulations had dramatically changed, as it would then be hard to believe that Molegro

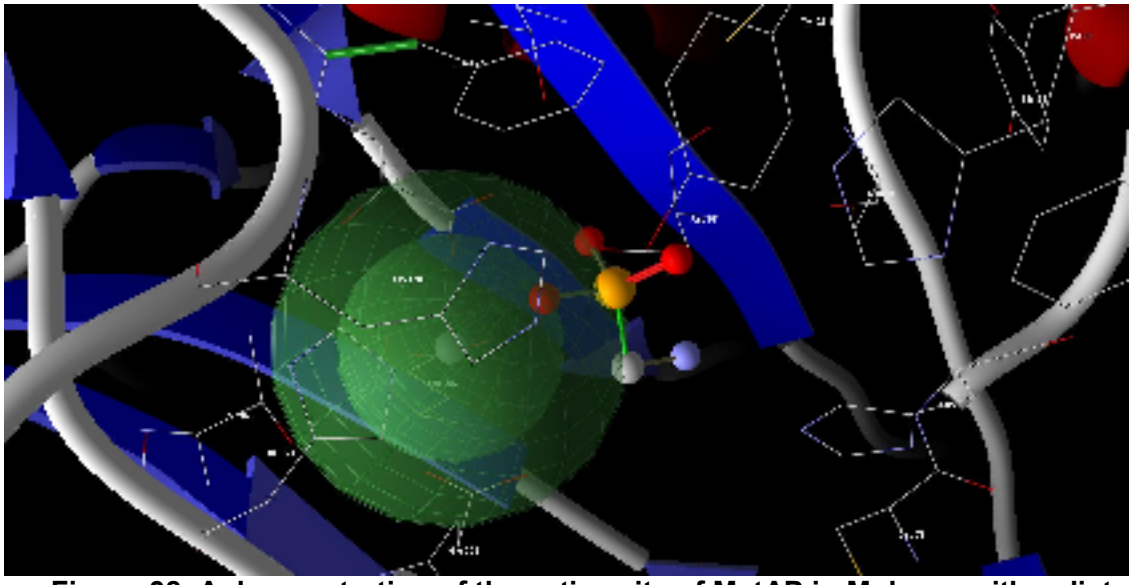


Figure 38. A demonstration of the active site of MetAP in Molegro with a distance constraint applied to the Mn(II) metal ion (constraint is represented by the green spheres surrounding it) insuring proximity to the sulfonic acid group of methylsulfonic acid

can recognize differences in compound functionality in a broad sense, and not assert that one, with molecular modeling and this form of *in silico* medicinal chemistry, is forcing results that are not at all in line with what occurs in reality. It was also intriguing to see that the energies for both 4NP and BNPP, the novel MetAP substrates and inhibitors, were so different with the newly imposed constraint considering their unique behaviour during Chapter 2's assay experiments. This perhaps implies that the new

constraint imposed upon these compounds, in spite of their still lowest binding energy, may not have been appropriate given that they do not seem to behave as the other Met-AMC substrate typically did in the kinetic assays. Most importantly, however, it is reassuring to see that in spite of this manually applied parameter, the results for the compounds used to study the functionality of specific groups in the active site of MetAP were not significantly affected by this seemingly large modification to the docking work space, such that the general trends observed in the results provided by Molegro could be rationally considered.

3.4.2. Active Site Waters

Based on the positive results supplied by adding the distance constraint for metal ions, we decided that the next step would be to incorporate the active site waters into the modeling of MetAP's active site and the docking of ligands within it. As mentioned previously, active site waters are typically left out of docking experiments, however, as the mechanism we have put forth demonstrates, as well as previous literature has established (6, 19), waters within the active site of MetAP play an integral role in NME hydrolysis of immature peptide substrates. Therefore, it was important to include them to confirm our mechanism, as well as further applying realistic phenomena to our computational modeling experiments.

3.4.2.1. Experimental

Initially, the PDB file 2GTX was imported into Molegro, but in contrast to previous docking experiments (as specified in Section 3.1.3.2.), all waters were included (268

waters exist in the dimeric form of this 2GTX structure of MetAP). The active site area was expanded, and the cavity in this region was automatically detected in Molegro. Then, with the cavity selected, waters were searched for in this region that would be proximal to compounds in this area, and would therefore be theoretically involved in catalysis. Two waters, 890 and 902 [Figure 39], were selected based on their presence

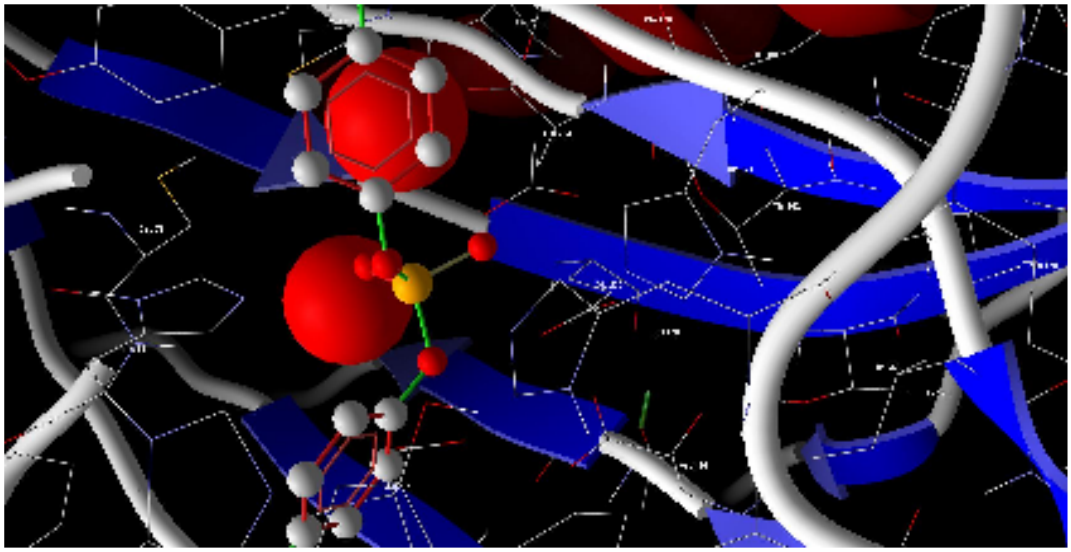


Figure 39. Two active site waters (890 and 902 in PDB code: 2GTX) (large red spheres – space filled versions of water molecules in Molegro) proximally located to the phosphonate group of BNPP

in the cavity-highlighted active site, their proximity to relevant residues (ie. His79), and their proximity to the included inhibitor. The file was then closed, and re-imported into Molegro, however, this time, only waters 890 and 902 were selected and imported into the Molegro work space. Docking was then performed as normal as per Section 3.1.3.2. However, when waters are contained in the Molegro work space, prior to performing the docking process, the program asks you whether to include them as a constraint, and to be treated as such. Similarly to the distance constraint imposed with the metal ions, Molegro treats the waters in kind, and bases its calculations of the

lowest energy poses with water molecules in the search space proximally to specified atoms of the compounds to be docked.

4.4.2.2. Results and Discussion

From viewing Table 10, one observes that there is quite a significant effect upon the LEPs from adding the waters to the active site of MetAP. As can be seen, there is an average increase in energy of best pose of 8-11% in kcal/mol per compound on average when the docking simulations are run with waters 890 and 902 included in the Molegro work space and treated as constraints. This increase in energy, and consequent assumed decrease in binding affinity of the compounds for MetAP's active

Compounds	LEP (kcal/mol)
Bis(4-Nitrophenyl)Phosphate	-88.204
4-Nitrophenyl Phosphate	-68.348
Iminobis(methylphosphonic) Acid	-67.476
Triethyl Phosphonoacetate	-66.458
Methyl-p-toluene sulfonate	-65.741
p-nitrobenzenesulfonic acid	-63.697
Methylene Diphosphonic Acid	-60.993
Pentane Sulfonic Acid	-58.723
Norleucine Aminophosphonate Diethyl Ester	-56.393
(Aminomethyl)-phosphonic Acid	-50.014
Aminomethane Sulfonic Acid	-48.918
Methylphosphonic Acid	-40.084

Table 10. LEP of phosphorus- and sulfur-based compounds docked with Molegro using PDB file 2GTX containing a single Mn(II) ion after active site waters included and treated as constraints.

site are most likely due to energetic penalties incurred by the compounds since they were required to be proximal to the water molecules, thereby losing other binding fostering interactions. Interestingly, this is in spite of the creation of hydrogen bonds

with the water molecules not previously found, generally since the compounds under study possess multiple oxygens with hydrogen bonding capability. One is actually capable of viewing these hydrogen bonds with a special feature of Molegro which allows the user to specifically view potential hydrogen bonds, which appear as broken green lines between donor and acceptor atoms [Figure 40]. This result is counterintuitive initially, but since the trends generally remain the same, with only minute changes in position with marginal energy differences by aminomethylphosphonic acid and aminomethane sulfonic acid, as well as iminobis(methylphosphonic) acid and triethyl phosphonoacetate, these results do offer insight into treatment of waters in the active site of MetAP as well as the general role of waters in the catalytic area of an enzyme.

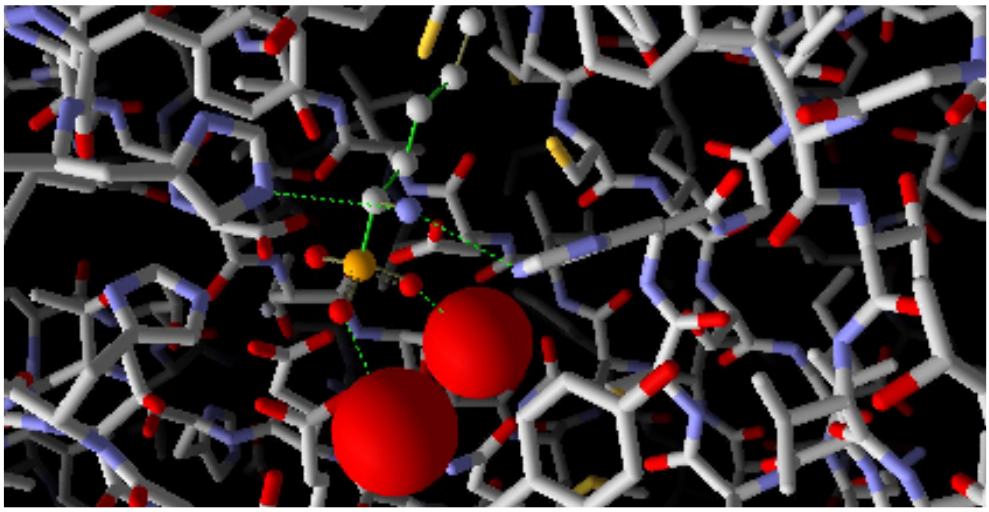


Figure 40. Active site waters (In red, space-fill for emphasis) in MetAP's active site with hydrogen bonds exhibited using green broken lines

4.4.3. Side Chain Flexibility

Another constraint available to the user in Molegro when docking ligands on a macromolecular complex, and the last one we felt relevant to these studies, is called "side chain flexibility". As one may guess, this refers to the idea that amino acid residue side chains are flexible, not rigid, and during catalysis, different portions of the residue

may occupy different points in space at any given time. This concept is certainly not novel, as most biochemical machinery has been noted to be in flux during activity (108), not the least enzymatic catalysis! Yet, many previous modeling publications have not included this concept in their virtual screens, nor have any of the articles that incorporate Molegro into their research published as of yet (108). Though, at the same time, various studies have mentioned that it is part of the reason virtual screening methods have continued to elicit false positive hits for lead candidates, as high as 33.3% (95). Therefore, as previously said, this would be the first study to date which actually utilizes the software's capabilities in incorporating phenomena other than structural biochemistry into *in silico* docking models. Certainly, in contrast to the previous two constraints, this is undoubtedly the most complex constraint and needs to be treated carefully. It is of course computationally demanding as many different arrangements need to be tested for docking interactions.

It can be difficult to state with conviction whether the movement of one area of a macromolecular complex will or will not affect the movement and structural conformation of another area of the larger assembly. It has definitely been previously noted in other enzymes, such as the classic example of triosephosphate isomerase, where movement in the sixth conserved loop of this glycolytic enzyme affects other neighbouring regions (109). Therefore, the original fear in using this constraint was that modification in one area should in theory affect other areas, and so therefore, one could not incorporate flexibility in one region without doing so in another. However, thankfully, in the case of MetAP, most structural movement has been previously identified in the

literature, and so therefore, we felt that if we incorporated the findings of major structural changes during catalysis, it was: a) better than maintaining rigidity, which would definitely not occur during actual ligand binding; and b) would still follow previous literature conclusions and therefore be faithful to what has been previously observed.

4.4.3.1. Experimental

Docking was performed as previously described in Section 3.1.3.2. However, the option for side chain flexibility was selected, and set up. Residues known to be involved in movement during catalysis were selected based on previous literature findings. Specifically, residues Tyr62, His63, Gly64, and Tyr65 (these four comprising what is known as "the YHGY loop"), His79, Asp97, Asp108, His178, and Trp221, all noted previously as being significantly mobile during ligand binding at the time of catalysis (6, 19, 21, 110), were selected as flexible side chains, with the individual settings left at default. The default setting meant that each side chain was to be rotated through the Molegro-assigned torsional degrees of freedom, unique for each amino acid, to find the most optimal torsional angles from which to assign bond energies and the overall energies of each LEP. More specifically, a histidine residue (we assigned side chain flexibility to three histidine residues in MetAP's active site) is given two degrees of freedom, meaning Molegro uses three specific side chain conformations for histidine in arriving at a torsional bond angle to be used in arriving at E_{Intra} based on the unique hybridization of the bonds in each conformation. For aspartate and tryptophan, two degrees of freedom are alloted, three for tyrosine, and only one for glycine. Once side

chain flexibility was set up, the selected residues' side chains become surrounded by a spherical net of varying blues and purples, depending on the settings selected, in the work space, denoting the region of potential movement possible with this setting enabled [Figure 41].

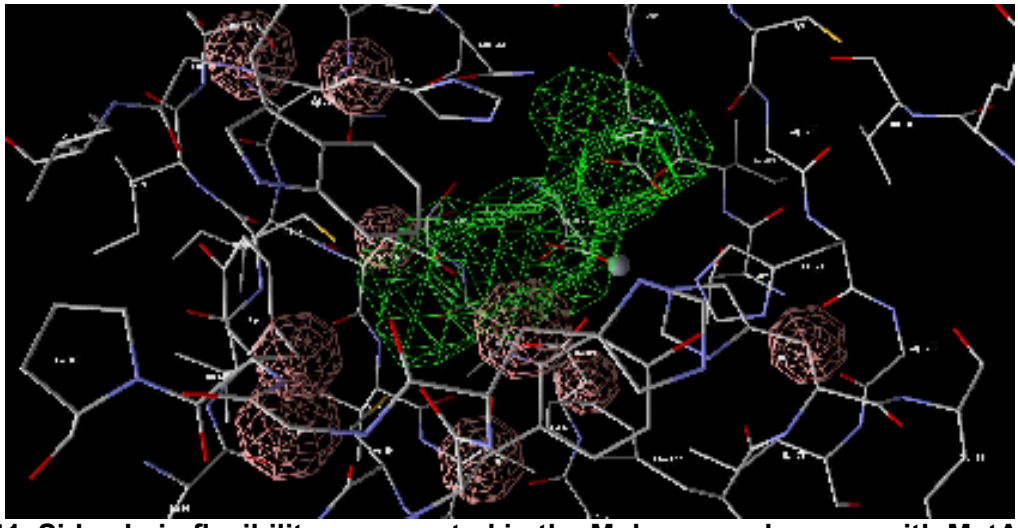


Figure 41. Side chain flexibility represented in the Molegro work space with MetAP active site residues selected *(refer to text)*.

3.4.3.2. Results and Discussion

Of all the constraints added, side chain flexibility elicited the greatest difference in lowest energy pose score of the compounds studied (Table 11). For example, in the

Compounds	LEP (kcal/mol)
Bis(4-Nitrophenyl)Phosphate	-74.327
4-Nitrophenyl Phosphate	-61.496
Triethyl Phosphonoacetate	-58.358
Iminobis(methylphosphonic) Acid	-60.428
Methyl-p-toluene sulfonate	-57.579
p-nitrobenzenesulfonic acid	-55.639
Methylene Diphosphonic Acid	-52.914
Pentane Sulfonic Acid	-53.436
Norleucine Aminophosphonate Diethyl Ester	-49.853
Aminomethane Sulfonic Acid	-48.372
(Aminomethyl)-phosphonic Acid	-47.852
Methylphosphonic Acid	-37.876

Table 11. LEP of phosphorus- and sulfur-based compounds docked with Molegro using PDB file 2GTX containing a single Mn(II) ion after the implementation of side chain flexibility.

'initial scan', from Section 3.3., and with the same basic settings of those studies utilizing Molegro from other research groups (96, 99, 102, 111, 112), BNPP had an energy of best pose of -117.9 +/- -4.215. Following the addition of distance constraints to the metal ion and active site waters, the calculated value became -88.204 +/- 3.996. However, with the addition of the side chain flexibility constraint following the previous constraint integration, the lowest energy pose score changes drastically to -74.33 +/- 2.473. This dramatic change is observed in all compounds docked with MetAP under this constraint. Nevertheless, the general trend in binding potency remained the same. There was only a single change in binding order between aminomethylsulfonic acid and aminomethylphosphonic acid, which does not change the picture much as they were already among the weakest binding compounds. With this, the LEP scores increased, effectively meaning that binding energy decreased as a result of this addition of this constraint.

Such a large adjustment in LEP scoring because of implementing side chain flexibility is interesting because of its positive direction. This is not entirely unexpected however. It is fitting that changing the relative mobility of residue side chains intimately involved in active site interactions with ligands taken up by the enzyme should have a considerable impact upon the energetics, and therefore the binding, of the ligand-enzyme complex. In an active site where the residue side chains remained static in position, it would be possible for the ligand, based on Molegro's iterating ever more favourable poses, to find a position where it remains flexed into a position that would be bound more tightly than if residues in this region were to become mobile themselves. Residue movement itself may, and the consequent steric obstructions this would introduce into the cavity region, disable movement of the ligand into simply any position that would satisfy the requirements for tightest binding possible. The legitimacy of this finding itself is substantiated by one of the preliminary experiments done earlier, in Section 3.2.3., which found residue mutations in the literature (His79, in this particular experiment) which affect enzymatic activity also influence the *in silico* binding ability of ligands to the active site. Though there were no mutations resulting in the decreased observed binding of the compounds, it is probably the case in this instance that the newly mobile residues selected, based on their participation in catalysis, render the simulation more realistic energetically speaking. Even still, though the scores changed, the trend has generally remained the same in terms of binding order, so this 'reality shift' induced by side chain flexibility may be just that – a shift to more realistic binding energies.

3.5. Correlation of Binding Assay Findings with Virtual Docking Results

The final test of molecular modeling reliability with respect to the virtual docking experiments performed in this chapter is through showing a definitive relationship between real biochemical assays that use the actual compounds and enzymes discussed rather than computer-simulated versions. So far, three of the nine papers that have performed virtual screening experiments with Molegro have done so in an effort to coordinate biochemical binding assays with the concordant results of Molegro. For instance, earlier this year (2009), Khan and colleagues at the University of Tromso, Norway, reported the discovery of novel thermolysin inhibitors from an NCI compound library of 273,000 molecules through Molegro, and correlated the biochemical binding affinities of the lead candidates with their *in silico* models (96). Their findings suggested that overall, Molegro possesses value as a tool to be used in what they called "virtual high throughput screening" (vHTS), and was reliable enough to use as a way to find lead candidates from a large, virtual compound library against a target (thermolysin) that had never been tested against inhibitors of the kind found in that library (3-nitrobenzamide compounds) (96). Moreover, they were able to correlate their binding assay results with the lead candidate compounds with those done *in silico,* with a fairly high confidence level (96). Based on the findings of this group, we decided to perform the same type of correlation with our phosphorus- and sulfur-based compounds and *E. coli* methionine aminopeptidase. Moreover, we decided to perform this correlation between the initial scan virtual docking results, and the results that had been acquired after applying the various constraints. In this way, it would give insight into whether the

addition of constraints improved docking accuracy relative to what was observed at the lab bench, or decrease it.

3.5.1. Experimental

The docking results (in the form of LEPs) from this chapter without any constraints (the initial scan), and all constraints added, as well as the IC_{50} values from Chapter 2 of the phosphorus- and sulfur-based compounds were imported into Microsoft Excel, where scatter plots were created using both sets of values and R^2 values obtained.

3.5.2. Results and Discussion

As can be seen from Figures 42 and 43, as well as taking into account the R^2 values of both sets of IC_{50}-LEP relationships, there is a positive correlation between the results of the virtual docking experiments and those of the actual experimental binding assays. Almost all of the 12 compounds assayed, except pentane sulfonic acid, and to some extent, 4NP, demonstrate a good correlation between their binding within Molegro and their actual IC_{50} values when performing the Met-AMC fluorogenic assays. The lack of perfect correlation is expected, especially in light of the fact that virtual screening often contains contains false positives, as well as what can be an error rate in a third of

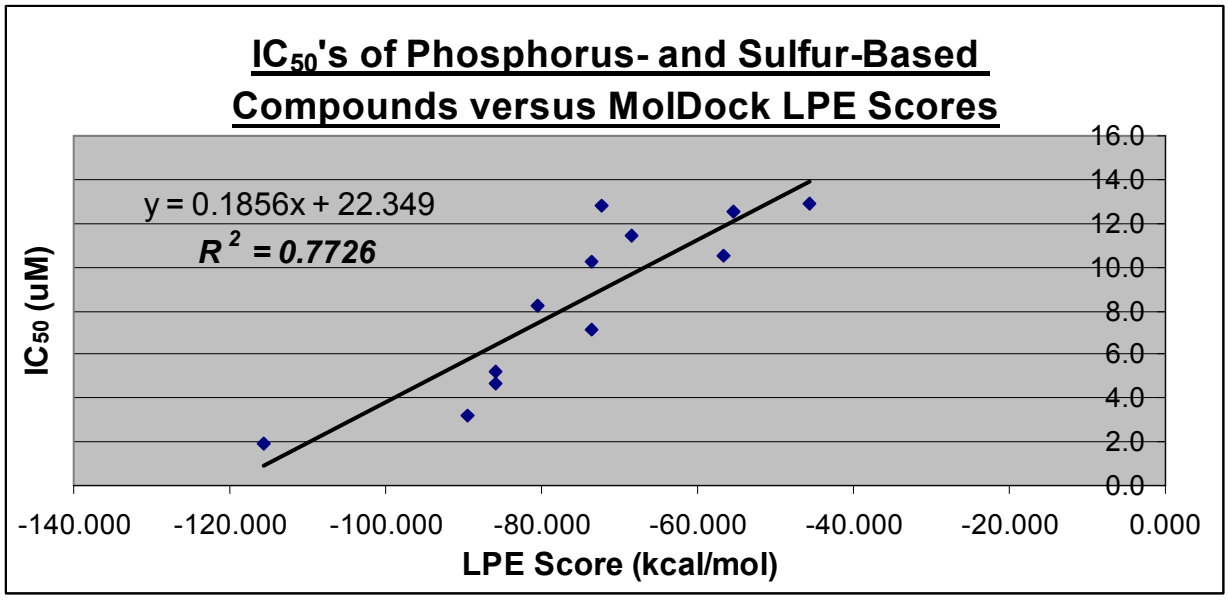


Figure 42. A correlation of IC_{50}'s of the phosphorus- and sulfur-based compounds under study with their MolDock LEP scores after the initial scan, $R^2 = 0.7726$.

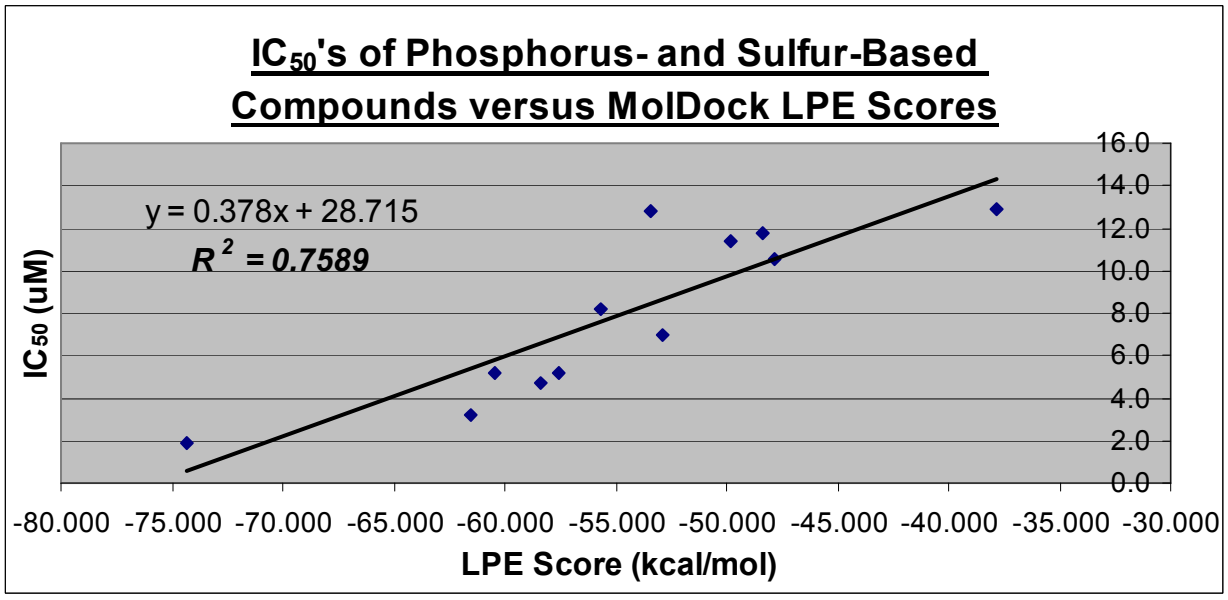


Figure 43. A correlation of IC_{50}'s of the phosphorus- and sulfur-based compounds under study with their MolDock LEP scores after the addition of constraints including distance, active site waters and side chain flexibility, $R^2 = 0.7589$.

compounds assayed (96). The correlations, both before and after constraint implementation, are 0.7726 and 0.7589, respectively. The reason for the slight variation is probably because of the change in position of various compounds, namely triethylphosphonoacetate, p-toluenesulfonic acid, aminomethylphoshonic acid, and

aminomethane sulfonic acid, albeit, their energy differences were minor and thus, elicited only minor changes in R^2. The R^2 values themselves are reasonable, their range supported by the previous literature on the subject, which show correlations of 0.67 (Khan *et al.*), 0.68 (Fogel *et al.*) and 0.87 (Sivaprakasam *et al.*), with their specific values being statistically considered to represent positive correlation (96, 102, 111). It is worth mentioning that even with the use of the various constraints, the correlation remains virtually unchanged. Therefore, our initial inquiry into whether the accuracy of Molegro would improve with their use has shown that our initial scan, along with those groups previously demonstrating its use in the literature, may be sufficient in determining some lead candidates within a group of compounds.

3.6. Summary

In summary, we have shown through a series of controlled docking experiments, the strengths and weaknesses in the use of the virtual docking software program, Molegro, and its potential use ability in performing virtual screening of molecular candidates for inhibitor discovery. Molegro appears to be a capable and reliable program to perform the molecular modeling of ligands docked on a macromolecular substrate (given the above described limitations!) , based on the program's ability to not only computationally support literature findings, but also to potentially replicate quite faithfully, in some instances, *in vitro* ligand and substrate behaviour. As well, we have demonstrated that although it is possible with Molegro to model phenomena in a fairly complex manner, moreso than has previously been done to date, it is also feasible to rely on the more simplified settings provided with Molegro by default to perform virtual

scans of compounds and identify, with a reasonable level of certainty, lead candidates in the drug design process. This is underscored by our correlation of our phosphorus- and sulfur-based compound binding assays with these molecular modeling results to a favourable degree of reliability, whether the constraints we applied were included in the correlations or not.

Chapter 4. Conclusions and Future Work

The objectives of this project were to explore the basis of MetAP catalytic functioning through probing its activity using various phosphorus- and sulfur-based probes, as well as novel fluorinated substrates. We also demonstrated the merits of a predictive *in silico* approach to the identification of lead candidates in the pursuit of potent enzyme inhibitors.

The ester phosphates, were most potent during our assaying of MetAP. This was probably because of their resemblance to MetAP's natural peptidic substrate's tetrahedral intermediate they are believed to exhibit during binding, hence their relatively low IC_{50} values and their concordant molecular modeling scores when docking in the enzyme's active site. However, all molecules assayed appeared to exhibit binding affinities, and this was dependant on their differences in chemical structure and how this affected each compounds interactions with the active site of MetAP. Interestingly, substrates known to be processed rapidly by MetAP, when modified with fluorine, as in the case of DFM-AMC and TFM-AMC, were demonstrably less prone to be catalyzed than without fluorine additions.

We also showed that MetAP may have a previously unknown function, being the removal of pNP from ester phosphate compounds not previously known to be catalytically acted upon by this enzyme. Though it is becoming more widely accepted that enzymes may have more than one natural substrate, this is novel in the case of MetAP and deserves further attention.

Further research is warranted in several areas. It would be interesting to assess the inhibitory potency of not only phosphonates with greater functional group diversity, taking cues from the phosphates studied in the substitutions on the oxygen groups. The low uM range observed in the ester phosphates we studied, and the more potent of the phosphonate group, could be enhanced potentially with functional group additions and general expansion in these areas.

More work is also worthy in examing fluorinated substrates like DFM-AMC and TFM-AMC. We have demonstrated that these compounds are both processed less readily by MetAP than the chromogenic Met-AMC. Though previous work in our laboratory has focused on this general area, little still remains in the field at large regarding unnatural amino acid substituted peptide substrate kinetics with other enzymes. Their unique position to be probed by means like ^{19}F NMR render them a valuable tool in biochemical studies, and it is hoped this will expand in the future.

Finally, we have developed a vHTS to be used in the drug design process, and showed that in spite of adding more realistic details to a virtual environment, the software's ability to pick out lead candidates amongst a larger group of possible inihibitors remained the same. The fact that compounds in the middle range were not so readily teased apart from one another, in addition to the previously observation and our own experience with the software docking ligands in inappropriate positions, means that work may need to be done in regards to refining the existing algorithms used in assessing active site interactions with potential therapeutics. Though MolDock appears

to be a reasonable molecular docking program, work still needs to be done in further rendering it capable of being more realistic to ever be truly valuable as a drug discovery tool. It should be added that all other docking software programs suffer from this weakness and that Molegro is not different in this regard.

Ultimately, the work presented in this thesis attempts to explore preliminary studies on understanding and inhibiting eMetAP. It is hard at times when performing experiments and analyzing results to keep in mind the bigger picture as the details under focus are so diminutive. As well, the arduous nature of laboratory research often obscures clarity and seems insurmountable. However, in spite of the toils of science, when the quest for knowledge to enrich all members of society is at hand, larger than each and every one involved, no detail is too minor to leave unrevealed, no sacrifice is too much to make, and no task is too difficult to perform when the rewards of these efforts are hope, health and vitality to all.

References

1. Campbell, K. M., Chambliss, G. H. (1997) *Mol. Gen. Genet*. 158(2):193-200
2. Blossom, D.B., McDonald, L.C. (2007) *Clin Infect Dis*. 45(2):222-7
3. D'Costa, V. M.; McGrann, K. M., Hughes, D. W. Wright, G. D. (2006) *Science*. 311(5759):374-377
4. J. H. Tran, G. A. Jacoby (2002). *Proc. Natl. Acad. Sci. U.S.A*. 99:5637
5. Neu, H. C. (1992) *Science 257*:1064-73
6. Vaughan, M. D., Sampson, P.B., Honek, J. F. (2002) *Curr. Med. Chem.* 9(3):385-409
7. Spencer, A. C., Heck, A., Takeuchi, N., Watanabe, K., and Spremulli, L. L. (2004) *Biochemistry. 43:* 9743-54
8. Adams, J. M. (1968) *J Mol Biol.* 33:571-89
9. Li, X., and Chang, Y. H. (1995) *Proc Natl Acad Sci U S A 92*:12357-61
10. Ben-Bassat, A., Bauer, K., Chang, S. Y., Myambo, K., Boosman, A., and Chang, S. (1987) *J Bacteriol. 169*:751-7
11. Frottin, F., Martinez, A., Peynot, P., Mitra, S., Holz, R.C., Meinnel, T. (2006) *Mol. Cell. Proteom*. 5(12):2336-2348
12. Lowther, W.T., Matthews, B.W. (2002) *Chem. Rev.* 102(12):4581-608
13. Brown, J.L. (1970) *Biochim. Biophys. Acta*. 221:480-488
14. Lowther, W.T., Matthews, B.W. (2002) *Chem. Rev*.102(12):4581-608
15. Addlagatta, A., Hu, X., Liu, J. O., Matthews, B. W. (2005) *Biochemistry*. 44:14741-14749
16. Arfin, S. M., R.L. Kendall, L. Hall, L.H. Weaver, A.E. Stewart, B.W. Matthews and R.A. Bradshaw. (1995) *Proc. Natl. Acad. Sci. USA*. 92:7714–7718
17. Bradshaw, R.A., W.W. Brickey and K.W. Walker. (1998) *Trends Biochem. Sci.* 23:263–267
18. Roderick, S. L., and Matthews, B. W. (1993) *Biochemistry. 32*:3907-12
19. Lowther, W. T., Orville, A. M., Madden, D. T., Lim, S., Rich, D. H., and Matthews, B. W. (1999) *Biochemistry 38*:7678-88
20. Schecter, I., Berger, Y. (1967) *Biochem. Biophys. Res. Comm.* 27(2):157-162
21. Mitra, S., Job, K.M., Meng, L.., Bennett, B., Holz, R.C. (2008) *FEBS J.* 275:6248-59
22. Mitra, S., Bennett, B., Holz, R.C. (2009) *Biochim Biophys Acta.* 1794:137-43
23. Walker, K.W., R.A. Bradshaw. (1998) *Protein Sci.* 7:2684–2687
24. Huang, Q.Q., Huang, M., Nan, F.J., Ye, Q.Z. (2005) *Bioorg. Med. Chem. Lett.* 15(24):5386-91
25. Wang, W. L., Chai, S. C., Huang, M., He, H. Z., Hurley, T. D., Ye, Q. Z. (2008) *J. Med. Chem.* 51:6110-6120
26. Deeth, R.J. (2008) *Inorg. Chem.* 47:6711-6725
27. Ye., Q. Z., Xie, S. X., Huang, M., Huang, W. J., Lu, J., Ma, Z. (2004) *J. Am. Chem. Soc.* 126:13940-13941
28. Li, X., Chang, Y.H. (1996) *Biochem. Biophys. Res. Commun.* 227:152–159
29. Qi-Zhuang, Y.,Sheng-Xue, X., Ze-Qiang, M., Min, H., (2006) *Proc. Nat. Acad. Sci. U.S.A*. 103(25):9470-9475

30. Ye, Q.Z., Xie, S.X., Huang, M., Huang, W.J., Lu, J.P., Ma, Z.Q. (2004) *JACS.* 126:13940-13941
31. D'Souza, V. M., Bennett, B., Copik, A. J., Holz, R. C. (2000) *Biochemistry.* 39:3817-26
32. Varshavsky, A. (1997) *Genes Cells* 2:13–28
33. Chang, S.Y., McGrary, E.C., Chang, S. (1989) *J. Bacteriol.* 171:4071–4072
34. Li, X., Chang, Y.H. (1995) *Proc. Natl. Acad. Sci. USA* 92:12357–12361
35. Miller, C.G., Kukral, A.M., Miller, J.L., Movva, N.R. (1989) *J. Bacteriol.* 171:5215–5217
36. Zuo, S. Q. Guo, Ling, C., Chang, Y.H. (1995) *Mol. Gen. Genet.* 246:247–253
37. Datta, B. M., Ray, K., Chakrabarti, D., Wylie, D.E., Gupta, N.K. (1989) *J. Biol. Chem.* 264:20620–20624
38. Datta, B. D. Chakrabarti, A.L.R., Gupta, N.K. (1988) *Proc. Natl. Acad. Sci. USA. 85:*3324–3328
39. Cutforth, T., Gaul, U. (1999) *Mech. Dev.* 82:23–28
40. Zhong, H., Bowen, J.P. (2006) *Curr. Med. Chem.* 13:849-62
41. Bernier, S. G., Taghizadeh, N, Thompson, C. D., Westlin, W. F., Hannig, G. (2005) *J. Cell*. *Biochem.* 95(6):1191-1203
42. Miller, C., G., Strauch, K.L., Kukral, A.M., Miller, J.L., Wingfield, P.T., Mazzei, G. J., Werlen, R.C., Graber, P., Movva, N.R. (1987*) Proc. Natl. Acad. Sci. U.S.A.* 84:2718-2722
43. Boutin, J. A. (1997) *Cell Signal 9:*15-35
44. Levitt, M. (1976) *J. Mol. Bio.* 104:59-107
45. Frottin, F., Martinez, A., Peynot, P., Mitra, S., Holz, R.C., Giglione, C. Meinnel T. (2006) *Mol. Cell. Proteom.* 5:2336-2349
46. Lowther, W.T., Zhang, Y., Sampson, P.B., Honek, J.F., Matthews, B.W. (1999) *Biochemistry.* 38:14810–14819
47. Douangamath, A., Dale, G.E., D'Arcy, A., Almstetter, M., Eckl, R., Frutos-Hoener, A. Henkel, B., Illgen, K., Nerdinger, S., Schulz, H., Mac, S., Thormann, M., Treml, A., Pierau, S., Wadman, S., Oefner, C. (2005) *J. Med. Chem.* 48(1):336
48. Kallander, L.S., Lu, Q., Chen, W., Tomaszek, T., Yang, G., Tew, D., Meek, T. D., Hofmann, G. A., Schulz-Pritchard, C.K., Smith, W. W., Janson, C.A., Ryan, M. D., Zhang, G., Johanson, K. O., Kirkpatrick, R. B., Ho, T.F., Fisher, P.W., Mattern, M. R., Johnson, R. K., Hansbury, M. J., Winkler, J.D., Ward, K.W., Veber, D.F., Thompson, S.K. (2005) *J. Med. Chem.* 48(18):5644-5647
49. Towbin, H., Bair, K. W., DeCaprio, J.A., Eck. M. J., Kim, S., Kinder, F. R., Morollo, A., Mueller, D. R., Schindler, P., Song, H.K., van Oostrum, J., Versace, R.W., Voshol, H., Wood, J., Zabludoff, S., Phillips, P.E. (2005) *J. Biol. Chem.* 278(52):52964-71.
50. Sheppard, G.S., Wang, J., Kawai, M., Fidanze, D., BaMaung, N. Y., Erickson, S. A., Barnes, D. M., Jason, S., Kolaczkowski, L., Vasudevan, A., Park, D. C., Wang, S., William, J., Mantei, R.A., Palazzo, F., Tucker-Garcia, L., Lou, P., Zhang, Q., Park, C.H., Kim, K. H., Petros, A., Olejniczak, E., Nettesheim, D., Hajduk, P., Henkin, J., Lesniewski, R., Davidsen, S.K., Bell, R. L. (2006) *J. Med. Chem.* 49(13):3832-3849

51. Jorgensen, A. T., Sorensen, M. D., Bjorkling, F., Liljefors, T. (2003) *J. Comp. Mol. Des.* 17(5-6):383-97
52. Baker, J. O., Wilkes, S. H., Bayliss, M. E., Prescott, J. M. (1983) *Biochemistry.* 22(9):2098-103
53. Sin, N. L., Meng, M.Q., Wang, J.J., Wen, W.G., Crews, C.M. (1997) *Proc. Natl. Acad. Sci. USA.* 94:6099–6103
54. Griffith, E.C., Su, Z., Turk, B.E., Chen, S., Chang, Y.H., Wu, Z., Biemann, J., Liu, J.O. (1997) *Chem. Biol.* 4:461–471
55. Hawkins, M.J. (1995) *Curr. Opin. Oncol.* 7:90–93
56. Zetter, B.R. (1998) *Annu. Rev. Med.* 49:407–424
57. Lev, E., Amar, Z. (2008) *J. Ethnopharmacol.* 119:24-40
58. Ekins, S., Mestres, J., Testa, B. (2005) *Br. J. Pharamcol*. 152(1):21-37.
59. Dailey, M.M., Hait, C., Holt, P.A., Maguire, J.M., Meier, J.B., Miller, M.C., Petraccone, L., Trent, J.O. (2009) *Exp. Mol. Pathol.* 86:141-50
60. Han, C.K., Ahn, S.K., Choi, N.S., Hong, R.K., Moon, S.K., Chun, H.S., Lee, S.J., Kijm, J.W., Hong, C.I., Deukjoon, K., Yoon, J. H., No, K.T. (2000) *Bioorg. Med. Chem. Lett.* 10:39-43
61. Evans, M.G., Polanyi, M. (1938) *Trans. Faraday Soc*. 34:11
62. R. Wolfenden. (1999) *Bioorg. Med. Chem.* 7:647–652
63. Smyth, T. P. (2004) *Bio. & Med. Chem*. 12(15):4081-4088
64. Alferov, K. V., Zhukov, Y. N., Khurs, E. N., Khomutov, M. (2003) *Mendeleev. Commun.* 243-244
65. Sauve, G., Rao, V.S., Lajoie, G., Belleau, B. (1985) *Can. J. Chem.* 63(11):3089-101
66. De Clercq., E. (2009) *Rev. Med. Virol.* 19(5):287-299
67. D'Souza, V.M., Holz, R.C. (1999) *Biochemistry*. 38(34):11079-11085
68. Zhou, Y., Guo, X. C., Yi, T., Yoshimoto, T., Pei, D. (2000) *Anal Biochem.* 280:159-65
69. Li, J. Y., Chen, L.L., Cui, Y.M., Luo, Q. L., Li, J., Nan, F. J., Ye, Q. Z. (2003) *Biochem. Biophys. Res. Commun.* 307(1):172-9
70. Brown, D.C., Collins, K.D. (1991) *J. Biol. Chem.* 266(3):1597-604
71. Chai, S.C., Wang, W.L., Ye, Q.Z. (2008) *J. Biol. Chem.* 283(40):26879-85
72. Guo, B., Yuan, Y., Wu, Y., Xie, Q., Yao, S. (2002) *Anal. Biochem.* 305(2):139-148
73. Murakami, K., Tsubouchi, R., Fukayama, M., Qiao, S., Yoshino, M. (2009) *Biol. Trace Elem. Res.* 130(1):31-8
74. Miguel, F., Augusto, A.C., Gurgueira, S.A. (2009) *Free Radic Res.* 43(4):340-7
75. Ghosh, M., Grunden, A.M., Dunn, D.M., Weiss, R., Adams, M.W. (1998) *J. Bacteriol.* 180(18):4781-9
76. D'Souza, V.M., Swierczek, N., Cosper, J., Meng, L., Ruebush, S., Copik, A. J., Scott, R. A., Holz, R. C. (2002) *Biochemistry*. 41(43):13096-13105
77. Atanassova, A., Sugita, M., Sugiura, M., Pajpanova, T., Ivanov, I. (2004) *Arch. Microbio.* 180(3):185-193
78. Yang, G., Kirkpatrick, R.B., Ho, T., Zhang, G.F., Liang, P.H., Johanson, K.O., Casper, D.J., Doyle, M.L., Marino, J.P. Jr, Thompson, S.K., Chen, W., Tew, D.G., Meek, T.D. (2001) *Biochemistry.* 40(35):10645-10654

79. Jeffery, C.J. (1999) *Trends Biochem Sci.* 24(1):8-11
80. Datta, B. (2009) *Biochim Biophys Acta.* Epub Ahead of Print (August 2009)
81. Datta, B., Majumdar, A., Datta, R., Balusu, R. (2004) *Biochemistry.* 43:14821-14831
82. Strasser, F., Pelton, P.D., Ganzhorn, A.J. (1995) *Biochem J.* 307:585-593
83. Han, R., Coleman, J.E. (1995) *Biochemistry* 34(13):4238 – 45
84. Gregan, F., Kettmann, V., Novomesky, P., Misikova, E. (1996) *Boll. Chim. Farm.* 135(4):229-31
85. Yadav, J.S., Reddy, B.V.S., Madan, U. (2001) *Synnlett.* 7:1131-1133
86. Shao, L.X., Shi, M. (2003) *Adv. Synth. Catalys.* 354(8):963-966
87. Sampson, P.B., Honek, J.F. (1999) *Org. Lett.* 1:1395-99
88. Honek, J.F. (2007) Fluorinated Methionines as Probes in Biological Chemistry, in *Current Fluoroorganic Chemistry: New Synthetic Directions, Technologies, Materials and Biological Applications* (Vadim, A. S., Mikami, K., Yamazaki, T., Welch, J.T., Honek, J.F., Ed.) pp 393-408, American Chemical Society, Washington, DC.
89. Duewel, H.S., Daub, E. Robinson, V., Honek, J.F. (2001) *Biochemistry*. 40(44):13167-76
90. Duewel, H.S., Daub, E., Robinson, V., Honek, J.F. (1997) *Biochemistry.* 36(11):3404-16
91. Jackson, J.C., Hammill, J.T., Mehl, R.A. (2007) *J. Am. Chem. Soc.* 129(5):1160-6
92. Budisa, N., Pipitone, O., Siwanowicz, I., Rubini, M., Pal, P.P., Holak, T.A., Gelmi, M.L. (2004) *Chem. Biodivers.* 1:1465-1475
93. Walasek, P., Honek, J.F. (2005) *BMC Biochem.* 6:21
94. Garner, D.K.,Vaughan, M.D., Hwang, H.J., Savelieff, M.G., Berry, S.M., Honek, J.F., Lu Y. (2006) *J. Am. Chem. Soc.* 128(49):15608-15617
95. Booth, R. J., Hodge, J.C. (1997) *J. Am. Chem. Soc.* 119:4882-84
96. Khan, M.T.H., Fuskevag, O.M., Sylte, I. (2009) *J. Med. Chem.* 52:48-61
97. Shoichet, B.K., McGovern, S.L., Wei, B., Irwin, J.J. (2002) *Curr. Opin. Chem. Biol.* 6(4):439-446
98. Honma, Y. Ishii, Y., Kasukabe, T., Okabe-Kado, J., Yamamoto-Yamaguchi, Y., Kakegawa, T., Awaya, A. (2001) *J. Med. Chem.* 44(26):4628-4240
99. Skone, G., Voiculescu, I., Cameron, S. (2009) *J. Comp. Mol. Des.* Epub ahead of print, July, 8, 2009.
100. Thomsen, R., Christensen, M.H. (2006) *J. Med. Chem.* 49:3315-3321
101. Storn, R., Price, K. (1997) *J. Global Optimiz.* 11:341-359
102. Fogel, G.B., Cheung, M., Pittman, E., Hecht, D. (2008) *J. Comp.Mol. Des. Epub 2007, Dec. 11.*
103. Hu, X., Zhu, J., Srivathsan, S., Pei, D. (2004) *Bioorg. Med. Chem. Lett.* 14(1):77-9
104. Rastelli, G., Degliesposti, G., Del Rio, A., Sgobba, M. (2009) *Chem. Biol. Drug Des.* 73(3):283-6
105. Attias, J., Bonnett, J.L. (1972) *Biochim Biophys Acta.* 268(2):422-30
106. Attias, J., Bonnett, J.L., Sauvagnargues, J.C. (1970) *Biochim Biophys Acta.* 212(2):315-321

107. Xie, S.X., Huang, W.J., Ma, Z.Q., Huang, M., Hanzlik, R.P., Ye, Q.Z. (2006) *Acta Crystallogr D Biol Crystallogr.* 62(Pt 4):425-32
108. Zhuravlev, P.I., Materese, C.K., Papoian, G.A. (2009) *J. Phys. Chem. B.* 113(26):8800-12
109. Sampson, N.S., Knowles, J.R. (1992) *Biochemistry.* 31(36):8482-7
110. Chiu, C.H., Lee, C.Z., Lin, K.S., Tam, M.F., Lin, L.Y. (1999) *J Bacteriol.* 181(15):4586-9
111. Sivaprakasam, P., Tosso, P.N., Doerksen, R.J. (2009), *J. Chem. Inf. Model.* 49(7):1787-96
112. Singh, N., Misra, K. (2009) *Bioinformation*. 3(6):255-62