7-10-2013

CATALYTIC MECHANISM AND FUNCTION EVOLVEMENT STUDIES OF PHOSPHATASES WITHIN HALOACID DEHALOGENASE SUPERFAMILY (HADSF)

Li Zheng

Li Zheng
Candidate

Chemistry and Chemical Biology
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Debra Dunaway-Mariano , Chairperson

Patrick S. Mariano

Wei Wang

Fu-Sen Liang

Karen N. Allen

CATALYTIC MECHANISM AND FUNCTION EVOLVEMENT STUDIES OF PHOSPHATASES WITHIN HALOACID DEHALOGENASE SUPERFAMILY (HADSF)

by

LI ZHENG

B.S., Chemistry, Sichuan University, China, 2002

M.S., Chemistry, Sichuan University, China, 2005

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Chemistry

The University of New Mexico
Albuquerque, New Mexico

January, 2013

DEDICATION

To

My parents Shifu Zheng and Changxiu Pang

and

My husband Min Wang

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation and gratitude to my research advisors Professor Patrick, S. Mariano and Professor Debra Dunaway-Mariano, not only for their scientific advice but also for their personal guidance. Their invaluable advice and everyday encouragement make all these years enjoyable and memorable. Their guidance and professional style will remain with me as I continue my career. Without them I could have never gone so far. They have the respect from the bottom of my heart.

I thank my committee members, Dr. Wei Wang, Dr. Karen N. Allen and Dr. Fu-Sen Liang for their valuable recommendations pertaining to this study and assistance in my professional development.

I am so grateful to have worked and collaborated with Dr. Karen N. Allen in Boston University. I am grateful to Dr. Kelly Daughtry for her great efforts in solving the crystal structure of YidA. I also thank Dr. Jeremiah Farelli for the high throughput screen for all the HAD phosphatases. Many thanks go to Dr. Yuriy V. Patskovsky in Albert Einstein College of Medicine for solving the structure of BT2542 phosphatase.

I would also like to thank all of the past and present members of Dr. Debra Dunaway-Mariano's research group for their help and friendship.

Finally, I would like to thank my husband, Min Wang who supports me greatly not only on my life, but also on my research. Without his support, friendship, love and understanding, none of this would be possible.

Catalytic Mechanism and Function Evolvement Studies of Phosphatases within Haloacid Dehalogenase Superfamily (HADSF)

by

Li Zheng

B.S., Chemistry, Sichuan University, China, 2002

M.S., Chemistry, Sichuan University, China, 2005

Ph.D., Chemistry, University of New Mexico, USA, 2013

ABSTRACT

The haloacid dehalogenase superfamily (HADSF) is one of the largest known enzyme superfamilies, having more than 32,000 nonredundant members. Enzymes in this superfamily catalyze a number of different chemical reactions including dehalogenation, phosphoryl transfer, and hydrolysis of phosphate esters and phosphates. The vast majority of HADSF members are phosphatases, which contain a highly conserved core domain that is constructed to catalyze hydrolysis reactions of phosphorylated sugars or nucleotides by using an active site Asp nucleophile and a Mg^{2+} cofactor. Most phosphatases possess a cap domain, which in association with the core domain assists active site desolvation and binds the region of the substrate that is displaced by water upon Asp nucleophilic attack at the phosphoryl group. In the studies described in this Dissertation, two nucleotidases (AphA and mdN) were chosen to examine how the core Rossmann fold of HADSF members stabilizes the transition state of phosphoryl transfer

reactions and conserves the stable catalytic scaffold under the pressure of evolutionary changes that expand substrate range. Another phosphatase, YidA, was subjected to studies in order to gain information about how accessory cap domains serve as sophisticated substrate recognition sites.

AphA is a nonspecific acid phosphohydrolase (NSAP) from *Escherichia coli*., which is secreted in the periplasmic space. Although AphA has the highest activity towards hydrolysis of nucleotide phosphates ($k_{cat}/K_m \sim 10^5\ M^{-1}s^{-1}$), it does not yet have an identified biological function. The crystal structures of the apo enzyme as well as complexes of AphA with products of nucleotide monophosphate hydrolysis and with substrate/intermediate analogs have been elucidated. Earlier analysis of these crystal structures shows that a specific hydrophobic substrate binding pocket, formed by Phe56 and Tyr193, exists in AphA. The results of solvent isotope studies carried out in the research effort described in this Dissertation, as well as investigations probing the effect on rates of alcohol additives suggest that the rate-limiting step of the AphA catalyzed hydrolysis reaction of ATP is the second partial reaction in which dephosphorylation of the phosphoenzyme (E-P) intermediate takes place. Observations made in an investigation of the linear free energy relationship between pKa of the alcohol that serves as the nucleophilic phosphoryl acceptor and the second order rate constant for hydrolysis of the E-P intermediate indicate that the mechanism for the rate limiting step of the reaction involving water (or added alcohol) as the nucleophile is highly dissociative. In addition, release of products in this process involves release of inorganic phosphate first followed by release of adenosine.

mdN is a mitochondrial deoxyribonucleotidase that is believed to function in a salvage pathways to regulate the sizes of deoxyribonucleotide pools. In Chapter 3 of the Dissertation, preliminary results of studies aimed at elucidating the detailed mechanism of mdN catalyzed hydrolysis reactions are presented. The findings show that mdN has activity towards substrates containing aromatic rings like those present in *para*-nitrophenyl phosphate, nucleotides and deoxynucleotides and that it is not active toward sugar phosphates. Like AphA, mdN contains a hydrophobic substrate-binding pocket formed by residues Phe75, Phe49, Trp76 and Trp96. The results of an investigation of mutli-turnover reactions of ^{14}C dUMP catalyzed by mdN revealed that a pre-steady state burst phase takes place followed by a slow steady-state phase. This observation leads to the conclusion that the substrate-binding step is not the rate-limiting step.

YidA is a HADSF phosphatase, isolated from *E.coli.*, that has a C2 cap. This enzyme is an effective catalyst for hydrolysis of several known phosphate metabolites, such as arabinose-5-phosphate, erythrose-4-phosphate, α-mannnose-1-phosphate and α-glucose 1-phosphate as well as two related carboxylic acid derivatives. The YidA gene is located in the same "neighborhood" as the *dgo* operon, whose genes encode enzymes involved in the galactonate to pyruvate and glyceraldehyde 3-phosphate degradation pathway. Studies described in Chapter 4 show that YidA has a moderate activity toward hydrolysis of the intermediate in this pathway, D-2-dehydro-3-deoxy-D-galactonate-6-phosphate ($k_{cat}/K_m = 6.4 \times 10^3\ M^{-1}s^{-1}$). Therefore, YidA is believed to play the role of a housekeeper whose purpose is to remove accumulated phosphorylated metabolites. Studies with a series of site-directed mutants, in which the 10 highly conserved cap residues Arg46, Leu111, His129, Glu130, Lys154, Met156, Ser184, Tyr187, Phe188 and Glu190 are

replaced by Ala, revealed that Glu130, Glu190 and Lys154 play important roles in substrate recognition and binding, while His129 and F188 are critical for binding C6 aldose/carboxylate phosphates.

Finally, a program targeted at protein function assignment is described in Chapter 5 of this Dissertation. Protein function assignment has always been a major goal in biology and, as the genome sequence database has rapidly expanded, this task has been significantly magnified. Unfortunately, using of bioinformatics techniques is unsuccessful in assigning the function of between 40-60% of the new enzyme sequences, and the much more demanding experimental methods have been applied to only a tiny fraction of the sequences. Therefore, a more reliable and efficient method for protein function assignment needs to be developed. The Enzyme Function Initiative (EFI), which is a large scale collaborative project that utilizes a multidisciplinary strategy to determine the functions of unknown proteins, was formed based on this consideration.

In Chapter 5, the general strategy employed for enzyme functional assignment in this initiative is described along with preliminary experimental results. From the results of high throughput screening, four HADSF phosphatases from different organisms, YigB (EFI-501262) and YbjI(501335) from *Escherichia coli*, BT2542 (501088) from *Bacteroides thetaiotaomicron* and Q9RUP0 (501193) from *Deinococcus radiodurans*, were identified to have high activities towards riboflavin 5'-monophoaphate (FMN). Further analysis of the kinetic parameters ($k_{cat}/K_m \sim 10^2$-10^4 $M^{-1}s^{-1}$) of these four proteins suggested the possibility that FMN may be their physiological substrate. A detailed study was performed with BT2542, which has a (Asp+2) - Gly instead of an (Asp+2) - Asp sequence motif. Observations made in this effort show that replacement of Gly with Ala

and Val does not significantly affect catalytic efficiency, while substitution with Asp causes a 20-fold decrease in activity.

Table of Contents

LIST OF FIGURES

LISTS OF TABLES

LIST OF ABBREVIATIONS

A	Alanine
AA	Amino Acid
ADE	Adenosine
ADP	Adenosine 5’-diphosphate
AMP	Adenosine 5’-monophosphate
AphA	Periplasmic acid phosphatase
ATP	Adenosine 5’-triphosphate
Arg/R	Arginine
Asn/N	Asparigine
Asp/D	Aspartic acid
ATCC	American Type Culture Collection

Bp	Base-Pair
Blast	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
BT	*Bacteroides thetaiotaomicron*
C	Cysteine
cDNA	Complementary Deoxynucleic acid
CDP	Cytidine diphosphate
CMP	Cytidine monophosphate
CTP	Cytidine triphosphate
Da	Dalton
dADP	Deoxy adenosine 5’-diphosphate
dAMP	Deoxy adenosine 5’-monophosphate
dATP	Deoxy adenosine 5’-triphosphate

DDT	Dichlorophenyltrichloroethene
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DI H2O	Distilled and deionized water
dGMP	Deoxy guanosine 5’-monophosphate
dGTP	Deoxy guanosine 5’-triphosphate
dTMP	Deoxy thymidine 5’-monophosphate
dTTP	Deoxy thymidine 5’-triphosphate
dUMP	Deoxy uridine 5’-triphosphate
3-D	Three dimensional
DTT	Dithiothreitol
E	Enzyme

E4P	Erythrose-4-phosphate
EC	Enzyme Commission
E.coli	*Escherichia Coli*
EDTA	Disodium ethylenediamine tetraacetate
E-P	Phosphoenzyme
ESI-MS	Electrospray ionization Mass spectrometry
EXPASY	Expert protein analysis system
F-6-P	Fructose-6-phosphate
FMN	Flavin mononucleotide
FPLC	Fast Protein Liquid Chromatography
g	Gram
G-1-P	Glucose-1-phosphate
G-6-P	Glucose-6-phosphate

Gln/Q	Glutamine
Glu/E	Glutamte
Gly/G	Glycine
GTP	Guanosine Triphosphate
h	Hour
HADSF	Haloalkanoic acid dehalogenase superfamily
H-bond	Hydrogen bond
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
His/H	Histidine
HPLC	High Performance Liquid Chromatography
Ile/I	Isoleucine
IPTG	Isopropylthio-β-galactoside

k	Rate constant
Kb	Kilo-base pair
k_{cat}	Enzyme turnover rate
K_d	Dissociation constant
K_m	Michaelis-Menten constant
K_i	Inhibition Constant
k_{obs}	Observed rate constant
kDa	Kilo-Dalton
KDPG	2-keto-3-deoxy-6-phosphogluconate
KDPGal	2-keto-3-deoxy-6-phosphogalaconate
L	Liter or Leucine
LB	Luria-Bertani
Lys/K	Lysine

M-1-P	Mannose-1-phosphate
M-6-P	Mannose-6-phosphate
mg	Milligram
min	Minute
ml	Milliliter
mmole	Milimole
MS	Mass spectrum
MW	Molecular Weight
Met/M	Methionine
ng	Nanogram
NCBI	National Center for Biotechnology Information
OD	Optical density

ORF	Open reading frame
pfu	Plaque forming units
PDB	Protein Data Bank
PCR	Polymerase Chain Reaction
PAGE	Polyacrylamide gel electrophoresis
PEG	Phosphoenolpyruvate
Phe/F	Phenylalanine
Pi	inorganic phosphate
pI	Isoelectric point
pNP	Purine nucleotide phosphorylase
PNPP	P-nitrophenyl phosphate
PPTase	Phosphopantetheinyl transferase
Pro/P	Proline

RNA	Ribonucleic acid
RT	Room Temperature
rpm	Rotation per minute
s	Second
S	Substrate/Serine
SDS	Sodium dodecyl sulfate
S_N	Nucleophilic substitution
Thr/T	Threonine
Tris	Tris[hydroxymethyl]aminomethane
Trp/W	Tryptophan
Tyr/Y	Tyrosine
U	Uridine

µM	Micromolar
UMP	Uridine monophosphate
UDP	Uridine diphosphate
UTP	Uridine triiphosphate
UV	Ultraviolet
Val/V	Valine
V_{max}	Maximum Velocity
V	Volume
WT	Wild type

CHAPTER ONE

1 Phosphatase Promoted Hydrolysis Reactions Have Diverse Catalytic Mechanism

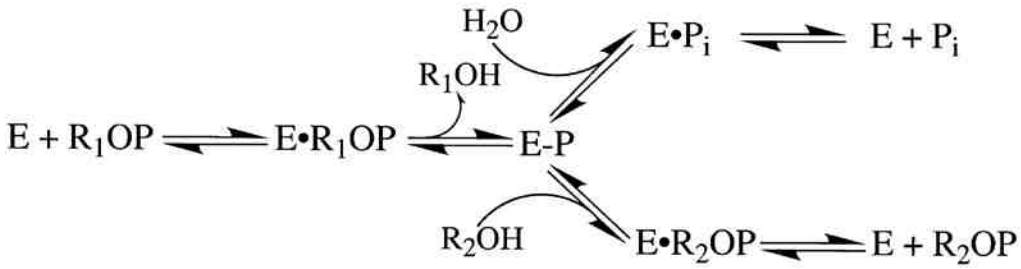
1.1 Phosphatases

Phosphatases are a large group of enzymes found in many different enzyme superfamilies that catalyze phosphoryl transfer reactions. In general, phosphatases can be arranged according to their substrates specificities, as exemplified by tyrosine specific phosphatases[(1)], serine/threonine specific phosphatases[(2)], dual specificity phosphatases[(3)], histidine phosphatases[(4)] and lipid phosphatases[(5)]. Owing to the response of their catalytic activity to the pH of the environment, these enzymes have been further classified as alkaline phosphatases[(6)], which are most active in an alkaline media, and acid phosphatases[(7)], which display higher activities under acidic reaction conditions.

1.1.1 Phosphatase Reaction Pathways Involving Phosphoenzyme Intermediates

Although many different types of phosphatases are known, with only a few exceptions they all rely on nucleophilic catalysis to promote phosphoryl transfer reactions. In 1961, Engstrom discovered that catalysis of phosphate ester hydrolysis by an *Escherichia coli* serine alkaline phosphatase proceeds through a pathway in which a phosphoseryl intermediate is generated [(8)]. Later work carried out by Lipmann, in which the serine phosphate intermediate was isolated after incubating alkaline phosphatase with inorganic phosphate, confirmed the earlier finding by Engstrom[(9)]. The results of experiments using a mixture of P^{32}-phosphate and P^{31}-glucose 6-phosphate enabled Lipmann to prove that a phosphoseryl intermediate is formed in the active site of the phosphatase.

Important insight into the mechanisms of phosphatase catalyzed reactions has come from studies probing the phosphorus stereochemical outcomes of the processes. Specifically, phosphatase catalyzed hydrolysis reactions of phosphate monoesters were shown to take place with retention rather than inversion of P-stereochemistry [(10)], an observation that demonstrates that the processes occur through two step routes, each of which takes place with inversion of P-stereochemistry. In the case of the serine phosphatase discussed above, the first mechanistic step involves substitution at phosphorus by the serine-102 hydroxyl group displacing the alcohol leaving group of the phosphate mono ester substrate (Scheme 1). The phosphoenzyme intermediate generated in this step then undergoes a displacement reaction, in which water or an alcohol serves as the nucleophile, to produce product and the regenerated active enzyme.



Scheme 1 General reaction mechanism of *E coli.* alkaline phosphatase catalyzed reaction.

Other nucleophilic active site residues play roles in phosphatase catalyzed reactions. For example, some acid phosphatases contain histidine imidazole groups in the active site that participate in forming phosphohistidine intermediates (Figure 1.1). In studies by Van Etten [(11)] trapping experiments were utilized to prove the existence of this intermediate in the prostatic acid phosphatase catalyzed reaction of ^{32}P-labeled p-nitrophenyl phosphate.

Specifically, Van Etten observed that quenching of the reaction before completion by addition of base led to isolation of a ^{32}P-labeled protein.

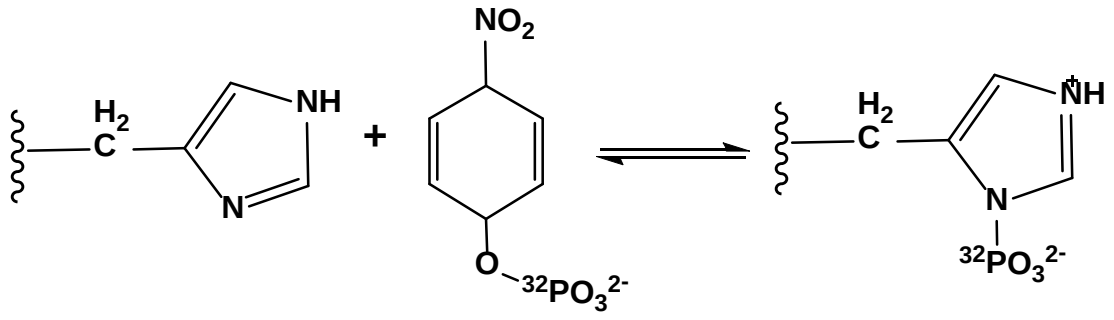


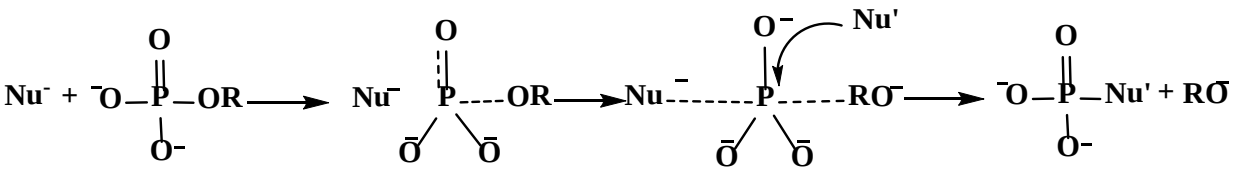
Figure 1.1 Phosphohistidine intermediate formed during the prostatic acid phosphatase catalyzed reaction.

As mentioned above, phosphoserine, phosphohistidine and phosphoaspartate intermediates have been identified as participants in the chemical mechanisms of phosphatase catalyzed phosphoryl transfer reactions. In addition, a cysteine phosphate[(12)] has been shown to serve as an intermediate in the hydrolysis reaction catalyzed by protein-tyrosine-phosphatase. An interesting feature of this process is that at least three distinctly different phosphoenzyme intermediates are formed in the reaction pathway.

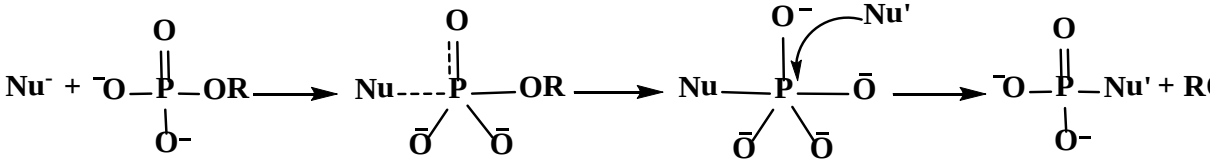
Additional information leading to a more detailed understanding of how phosphatases catalyze reactions has come from studies exploring the nature of the phosphoryl transfer process. The two limiting mechanisms that are typically considered operable for nonenzymatic solution phosphoryl-transfer reactions involve either a dissociative or an associative process [(13)]. In the transition state for the dissociative mechanism, the bond between phosphorus and the leaving group is broken either extensively or completely and

bond formation to the incoming nucleophile is either not well advanced or absent (Figure 1.2A). In contrast, in the associative pathway the respective bonds between the incoming nucleophile and phosphorus and the departing leaving group and phosphorus are formed and broken in a concerted (not necessarily synchronous) manner (Figure 1.2B).

In order to gain insight into which if any of these two mechanistic extremes operate in phosphatase promoted reactions, linear free energy relationship studies (LFERS) have been conducted to provide information needed to characterize the nature of intermediates/transition states in these processes. In 1986, Hall and Williams[14] carried out studies of the *Escherichia coli* alkaline phosphatase catalyzed hydrolysis reactions of para-substituted aryl phosphates and alkyl monophosphate esters. An evaluation of the k_{cat}/K_M data gave a $\beta_{leaving\ group}$ value of *ca.* -0.2, which is much less negative than that ($\beta_{leaving\ group}$ = -1.2[15] for the corresponding noncatalyzed solution reactions. This comparison led to the conclusion that possibility the enzymatic reaction possible follows an associative mechanism that differs from the dissociative pathway followed in the noncatalyzed solution process. In a later investigation, Hollfelder and Herschlag[13] obtained a much larger negative value ($\beta_{leaving\ group}$ = -0.8) for reactions of para-substitued aryl phosphorothioates promoted by the same phosphatase. This finding suggested that alkaline phosphatase can also catalyze reactions of certain substrates, especially ones that are more prone to forming metaphosphate intermediates[15], via a dissociative mechanisms. Zalatan and Herschlag[16] later demonstrated that alkaline phosphatase catalyzes the hydrolysis of phosphate monoesters through a loose dissociative transition state (Figure 1.3A) while it hydrolyzes phosphate diesters via a tighter, more associative transition state(Figure 1.3B).



(A)



(B)

Figure 1.2 Dissociative **(A)** and associative **(B)** reaction mechanisms of phosphoryl transfer reactions. In the dissociative mechanism, a metaphosphate intermediate is formed. In the associative mechanism, a pentavalent phosphorane intermediate is formed.

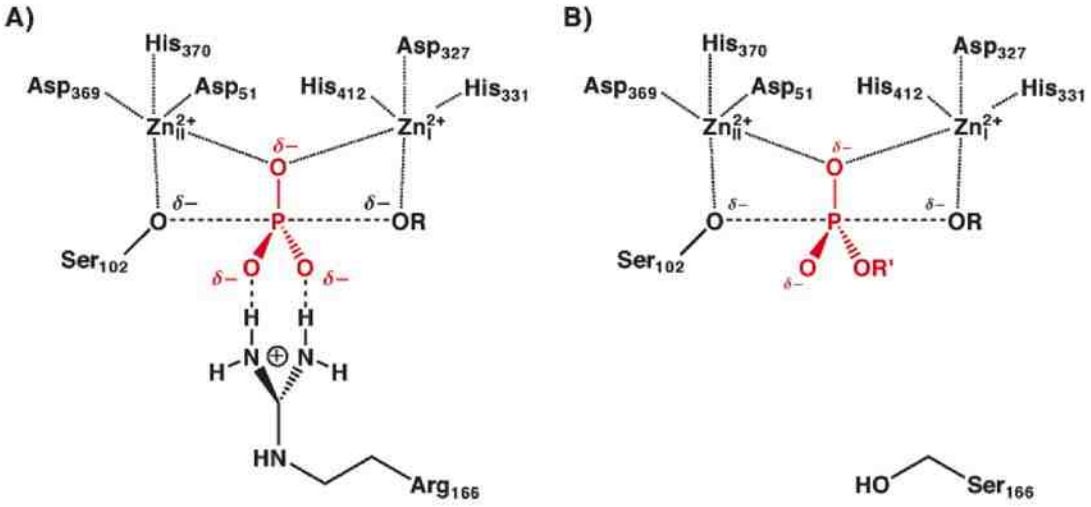


Figure 1.3 The interactions of alkaline phosphatase with an intermediate form by a phosphate monoester **(A)** and a phosphate diester substrate **(B)**. (from *Ref.*[16])

The dissociative and associative reaction pathways for tyrosine phosphatase promoted reactions have been studied by using density functional theory calculations[17], which yielded energy barriers and geometries for points along each reaction pathway. The energy barrier of the dissociative process with early proton transfer to the leaving group was predicted to be 9 kcal/mol, while that for the associative reaction with late proton transfer was predicted to be 22 kcal/mol. These results are in accord with experimental evidence[18] that showed that tyrosine phosphatase catalyzes phosphate ester hydrolysis of a dianionic substrate via a dissociative pathway.

Information about the intermediate of phosphoserine phosphatase (PSP) catalyzed reaction was obtained using x-ray crystallographic studies[19] with the BeF_3^- PSP complex as a model (Figure 1.4). The tight structure of the complex suggested that the intermediate in the phosphatase reaction pathway is highly associative.

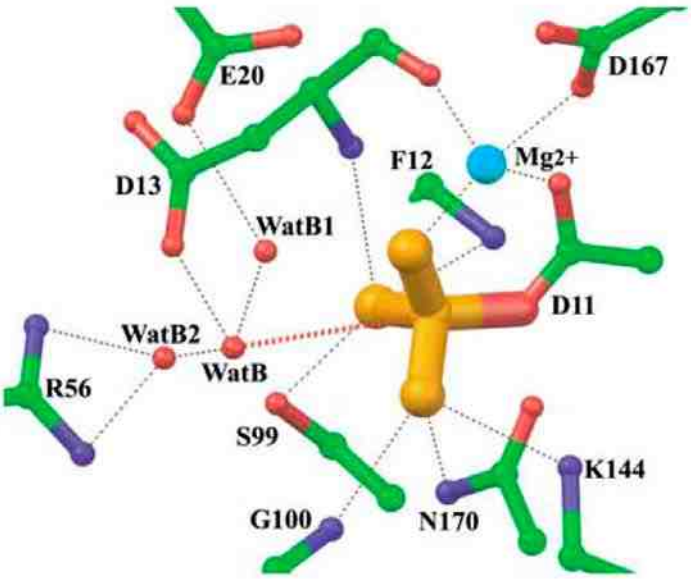


Figure 1.4 The structure of the PSP BeF_3^- complex as a mimic of the phospho-aspartyl enzyme intermediate.

1.1.2 Phosphatase Reaction Pathways Not Involving Phosphoenzyme Intermediates

Although nucleophilic 'double displacement' reaction mechanisms operate in most phosphatase catalyzed reactions, exceptions do exist. As mentioned above, one method of determining whether a covalent phosphoryl enzyme intermediate, formed by a nucleophilic residue, exists in the catalytic pathway involves the use of phosphorus stereochemistry. Specifically, when the reaction proceeds via a single-step, direct phosphory transfer mechanism, the process will take place with inversion of the P-configuration, while for a two-step mechanism, via a covalent phosphorylated intermediate, overall retention P-configuration will be observed. In 1984[(20)], the results of a study of the stereochemistry of yeast inorganic pyrophosphatase catalyzed reaction of chiral [γ-^{17}O,^{18}O]-ATPγS showed that the product, chiral inorganic thio-[^{17}O, ^{18}O]-phosphate was produced by inversion of P-configuration. Combined with observations made when carrying out a single turnover experiment utilizing $H_2{}^{18}O$, which gave 1 mol each of $^{18}O^{16}O_3P$ and $^{16}O_4P$ (Figure 1.5) per mol of inorganic pyrophosphate hydrolyzed, the findings strongly indicate that phosphoryl group transfer in the reaction catalyzed by pyrophosphatase takes place via a direct displacement mechanism, not involving a phosphorylated enzyme intermediate.

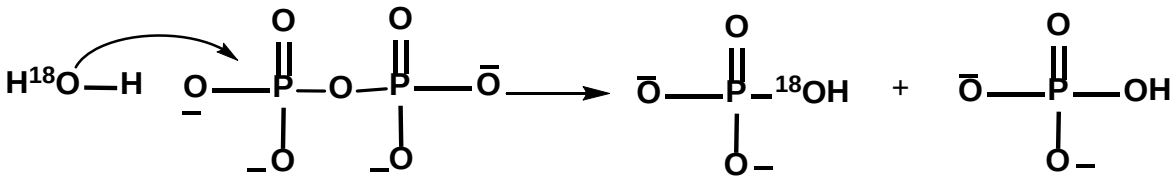


Figure 1.5 Proposed mechanism of yeast pyrophosphatase.

In a later effort, the stereochemistry of phosphoryl group transfer to water promoted by purple acid phosphatases was explored[21]. The observation that the P-configuration was inverted enabled the exclusion of a mechanism involving a phosphoenzyme intermediate and supported the direct transfer of a phosphoryl group to water.

Another interesting discovery[22] was made in studies of mutants of *Escherichia coli* alkaline phosphatase, an enzyme that utilizes an active site serine residue to promote nucleophilic catalysis. The S102A, S102L mutants of *Escherichia coli* alkaline phosphatase, in which the active-site nucleophile serine102 is not present, catalyze the hydrolysis of reactions of phosphate monoesters with K_m values that are similar to that of the wild type enzyme. Importantly, the mechanism of the hydrolysis reaction promoted by mutant enzymes changes from a two-step route via a phosphoryl-enzyme intermediate to a one-step, direct hydrolysis pathway without an intermediate.

1.1.3 Rate Limiting Step in the Phosphatase Reaction Pathway

It is generally accepted that the steps in phosphatase catalyzed phosphoryl transfer reactions involve substrate binding, a chemical step, and then products release (Figure 1.6).

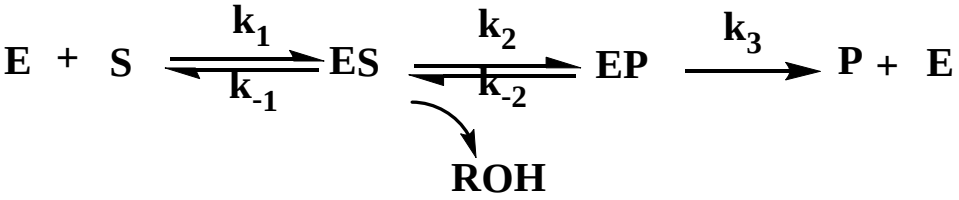


Figure 1.6 General mechanism for phosphatase promoted hydrolysis of phosphate monoesters.

Recently, kinetic analyses of the *Escherichia coli* alkaline phosphatase catalyzed reaction of methylumbelliferyl phosphate have been performed to determine which step in this pathway is rate limiting. The substrate was designed to enable detection of the highly fluorescent alcohol product of the phosphate hydrolysis reaction. At acid pH (5.3) utilizing excess substrate and the product, inorganic phosphate, a burst phase was shown to take place by using the stopped-flow kinetic technique[(23)]. In contrast, at alkaline pH no burst was observed and the rate limiting step was concluded to be dissociation of the product phosphate from the EP complex[(10)]. Moreover, the nature of the rate limiting step for reaction at pH 8.0 was found to vary according to the pKa of the leaving group[(24)]. Specifically, when the pKa of the leaving group is <10, the k_{cat} for the alkaline phosphatase was independent of the nature of alcohol group in the phosphate monoester, and dissociation of inorganic phosphate from the EP complex was much slower than dephosphorylation in ES. However, when the pKa of the leaving group is between 10 and 15, dephosphorylation in ES and dissociation of inorganic phosphate from EP both contribute to the rate. Finally, when the pKa of the leaving group is >15, dephosphorylation is the rate limiting step. A later investigation of the catalytic mechanism of human phosphatase Cdc25A also demonstrated that the rate limiting step of the process also depends on the pKa of the alcohol leaving group of the substrate phosphate monoester[(25)].

In 1991[(26)] the complete time course of the hydrolysis reaction of *p*-nitrophenyl phosphate catalyzed by the low molecular weight phosphotyrosyl protein phosphatase from bovine heart was analyzed in detail. The existence of a transient pre-steady-state “burst” phase, as well as the observation that the overall rate of hydrolysis is increased greatly while the

level of phosphoenzyme intermediate is constant when alcohol acceptors are present, are all consistent with a mechanism for this process involving rate limiting dephosphorylation of the phosphate ester in the ES complex.

Metal cations, present in almost all phosphatases, play important functional and/or structural roles. Divalent Zn^{2+} and Mg^{2+}, the most common metal ions found in these enzymes are utilized for substrate binding and catalysis.

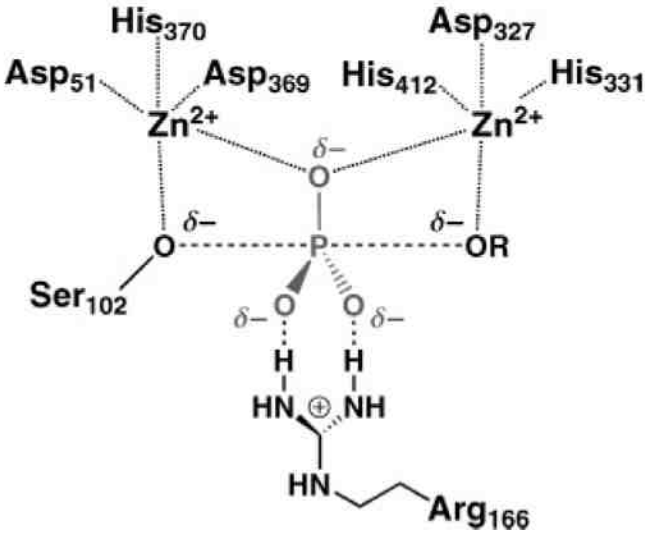


Figure 1.7 The Zn^{2+} binding sites in the transition state model for phosphoryl transfer catalyzed by alkaline phosphatase based on the structure with a bound vanadate ligand.

For example, fully active native alkaline phosphatase contains four Zn^{2+} and two Mg^{2+} ions[27] per each monomeric unit. Analysis of the crystal structure of the vanadate complex of this enzyme shows that two Zn^{2+} ions are located in the active site (Figure 1.7) to position the nucleophilic Ser102 hydroxyl group and to activate and bind to the oxygen

atom of the leaving group. In addition, one of the nonbridging phosphate oxygens interacts with both Zn^{2+} atoms. It is observed that the apoenzyme not containing these metal ions does not bind phosphate esters[28]. Moreover, the results of kinetic isotope effect (KIEs) studies with alkaline phosphatase demonstrate that electrostatic interaction between the two Zn^{2+} site and a nonbridging phosphate ester oxygen atom contributes to the large rate enhancement[29].

The activities of *Escherichia coli* pyrophosphatases are also strongly dependent on the presence of divalent metal ions[30]. These enzymes have four functional metal ions in the active sites, two of which are bound to the enzyme and serve as essential cofactors and the other two metal ions are complexed with the substrate PPi and forming product Pi[31].

1.2 Phosphatase Members of the Haloacid Dehalogenase Enzyme Superfamily (HADSF)

The Haloacid Dehalogenase (HAD) superfamily[32], named after the first family member to be structurally characterized[33], is a ubiquitous family of enzymes[32, 34]. Presently, *ca.* 50,000 deposited gene sequences have been identified that encode proteins in this superfamily. The HAD superfamily is highly evolvable and, thus, it is ideal to employ in formulating and testing theories about enzyme superfamily nucleation and growth.

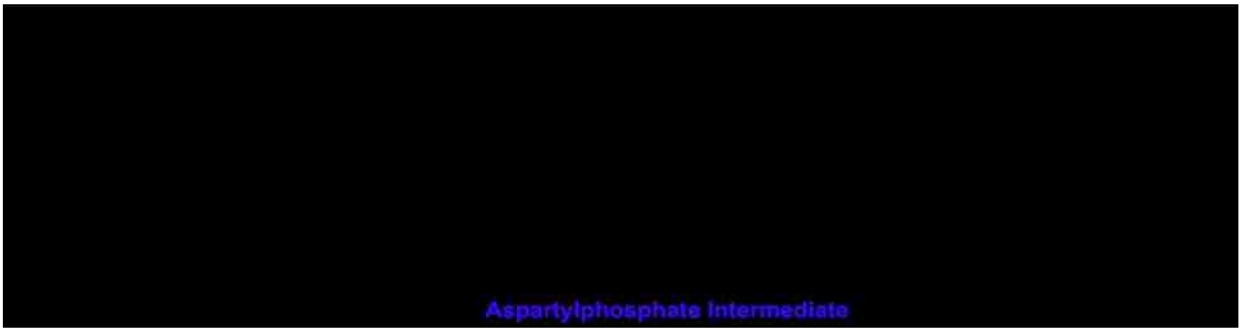


Figure 1.8 The general catalytic mechanism of HADSF phosphatase members.

A large subgroup of HAD superfamily (HADSF) members serve as phosphatases. All enzymes in this family possess a highly conserved core domain that contains amino acid side chains that participate in catalysis of phosphoryl group transfer processes. In addition, all members possess an Asp nucleophile residue in their active sites, which participates as a nucleophile in catalyzing the phosphoryl transfer reactions (Figure 1.8), and all members of the HADSF, except haloalkanoic acid dehalogenases, which catalyze a carbon group transfer reactions[(35)], utilize Mg^{2+} as a cofactor for catalysis [(36, 37)]. The catalytic scaffold of the core domain of phosphatase members of the HADSF is formed by 4-loops that position both the "core catalytic residues" and "diversification residues" (Figure 1.9A). The core residues include a loop 1 Asp that serves as a nucleophile and a loop 4 Asp that participates as a general acid/base and to bind Mg^{2+} ion. Additional catalytic residues in the active sites of HADSF members include loop2 Ser/Thr and loop 3 Arg/Lys side chains that serve to position the phosphate moiety of the substrate via hydrogen bonding interactions (Figure 1.9 A).

A B

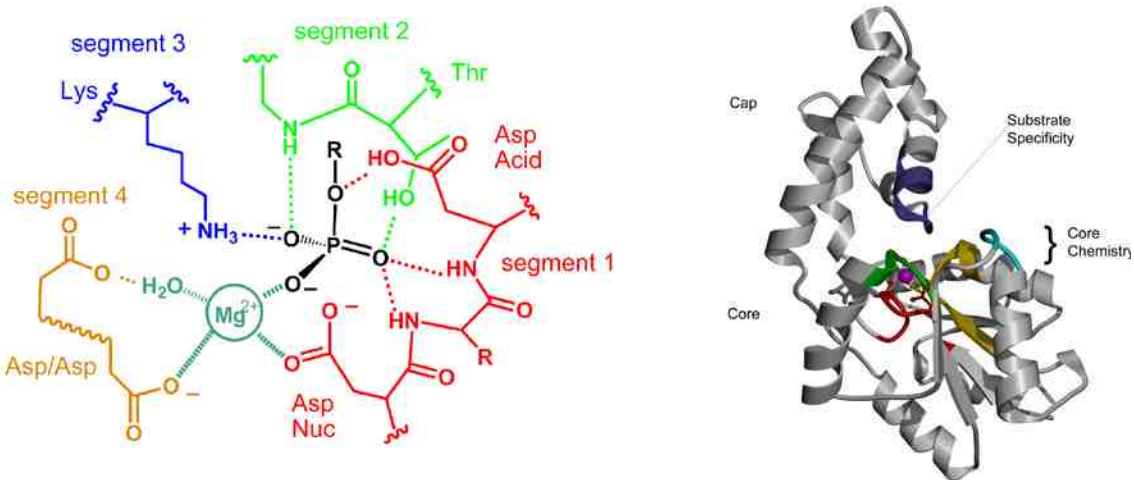


Figure 1.9 Catalytic loops comprising the active site of HADSF members shown as a chemical model to depict the conserved role of each loop **(A)** and in the context of the protein scaffold of phosphonatase with conserved loops colored as in A **(B)**.

Many HADSF members possess cap domains in addition to core domains[38-43] (see Figure 1.9B). Typically, the cap domains of these enzymes contain amino acid side chains that participate in binding to the leaving groups of phosphoryl transfer reactions and, as a result, contribute to substrate recognition. In addition, the cap domains play a role in separating solvent and substrate by functioning as a hinged "lid" over the substrate-filled active site of the core domain. The caps are connected at different sites of the enzyme backbone, a phenomenon that serves as the basis for subclassification of members of this family. These HAD subclasses are referred to as C0 caps, which have only small inserts in either of the two points of cap insertion, C1 caps, whose caps are defined as inserts occurring in the middle of a β-hairpin of the flap motif, and fold and C2

caps, which have inserts occurring in the linker positioned immediately after loop 2 (Figure 1.10).

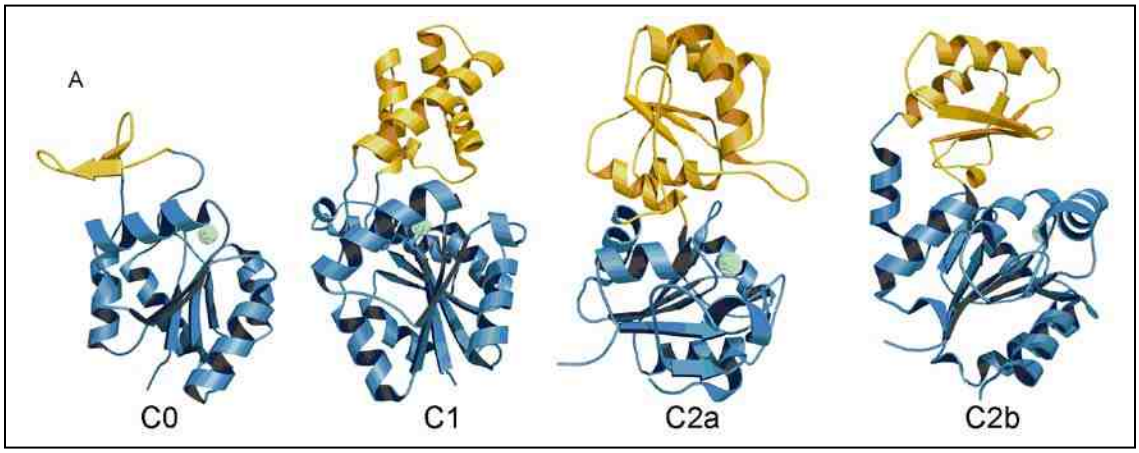


Figure 1.10 HADSF members C1 and C2 are distinguished by the locations of insertion points of the cap (gold) in the Rossmann fold (blue), including C0, C1 C2 and C2b.

1.2.1 Mechanism of HADSF Phosphatase Catalyzed Reactions

Approximately 80-90% of the HADSF members are phosphatases which catalyze organophosphate hydrolysis reactions[(44)]. Because approximately 40% of the bacterial metabolome is comprised of phosphorylated metabolites [(45)], it is not surprising that all cellular organisms depend extensively on reactions catalyzed by HADSF and other phosphatases.

Numerous studies have been performed to probe the detailed mechanism of HAD phosphatase hydrolyzed reactions. As mentioned above, all HADSF members have a highly conserved catalytic domain containing a nuclephilic Asp nucleophile and, many members a cap domain/loop that contains substrate recognition residues.

All HADSF members, except for 2-haloacid dehalogenases[(46)], have an absolute requirement for Mg^{2+} as a cofactor for catalysis. Mg^{2+} is coordinated in these enzymes to the oxygen atoms of the nucleophile Asp, the general acid/base Asp, loop 2 Tyr/Asp residues and three water molecules (Figure 1.11). In the reaction pathway, the positive charge of the bound enables transition state stabilization by binding to the negatively charged phosphate moiety of the substrate during reaction. Mg^{2+} cofactor shields the charge of phosphonyl group when it is approached by the negative-charged Asp nucleophile (Figure 1.12).

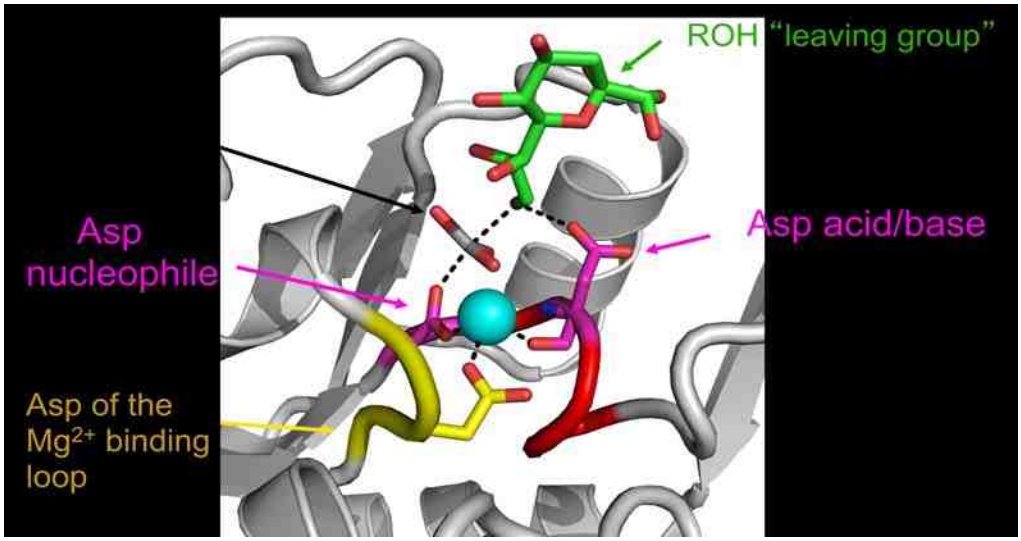


Figure 1.11 Model of the Mg^{2+} ion binding motif in the active site of phosphatase members of the HADSF. Mg^{2+} ion is coordinated to the nucleophilic Asp, general acid/base Asp, an oxygen atom of the main chain Thr/Asp and three water molecules (not shown here).

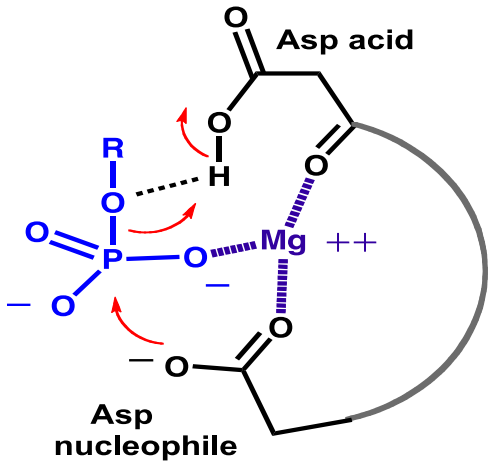


Figure 1.12 A depiction of the charge shielding of HADSF Mg^{2+} cofactor.

The HAD phosphatase reaction pathway consists of two partial reactions (Figure 1.8). The first is the transfer of the substrate phosphoryl group to the carboxylate group of the Asp nucleophile to form an enzyme aspartylphosphate intermediate and the second is phosphoryl transfer from the intermediate to a water molecule[(47)]. The Asp in motif 1 acts the nucleophile and the Asp in loop 4 acts as a general acid-base, which binds and protonates the substrate leaving group in the first step and deprotonates the nucleophile of the second step. Loop 2 Thr/Ser, loop 3 Lys and several catalytic scaffold backbone amide NHs stabilize the aspartylphosphate intermediate through hydrogen bonds and facilitate the nucleophilic attack [(48)] (Figure 1.13).

A recent crystal structure (Figure 1.14) of a pentacoordinate oxyphosphorane intermediate, formed in the reaction converting β-glucose-1-phosphate to β-glucose-6-phosphate by the HAD phosphatase β-phosphoglucomutase, was reported in 2003, indicating that the catalytic scaffold acts as an electrostatic mold for transition state stabilization. Thus discovery, provided an actual image of the phosphorus-aspartate

intermediate involved in the phosphoryl transfer process[49], which shows that a pentavalent phosphorus species is formed with the Asp8 OD1 and hexose C (1) O occupying apical positions and three P oxygen atoms occupying equatorial positions of a trigonal bipyramid (Figure 1.14).

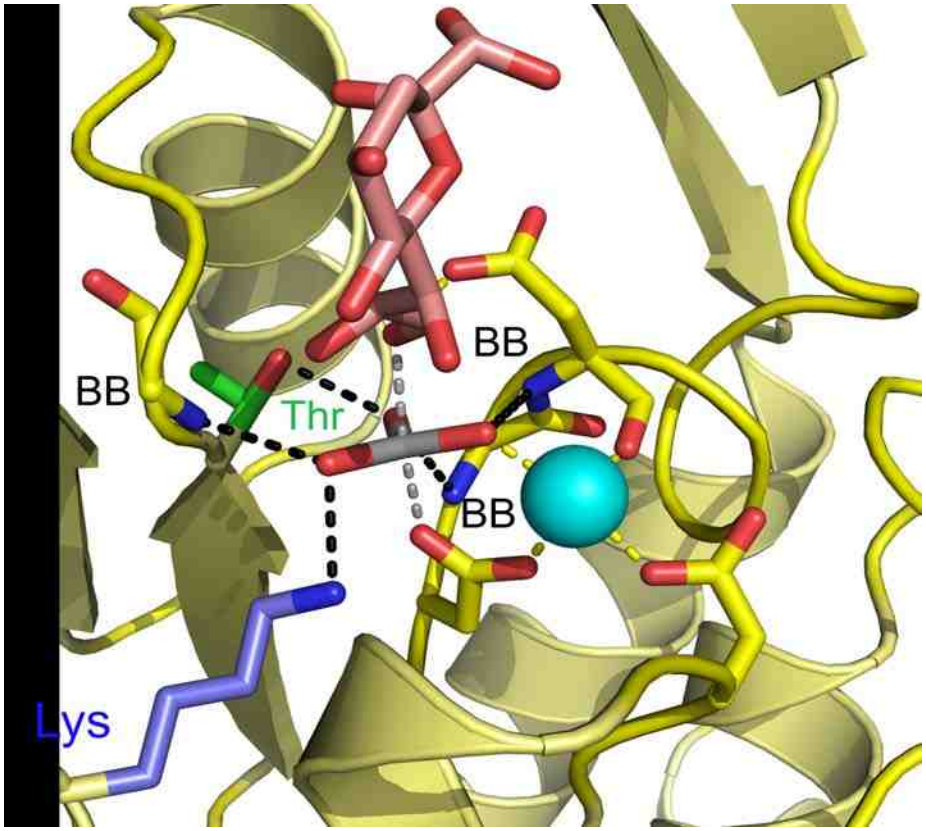


Figure 1.13 Snapshot of KDN-9-phosphate phosphatase bound with Mg^{2+} (cyan sphere), vanadate (red and grey) and neuramic acid (pink and red). The hydrogen bonds formed between vanadate and side chains of Thr (green), Lys (blue) and backbone amide NHs (BB) are shown as black dashed lines.

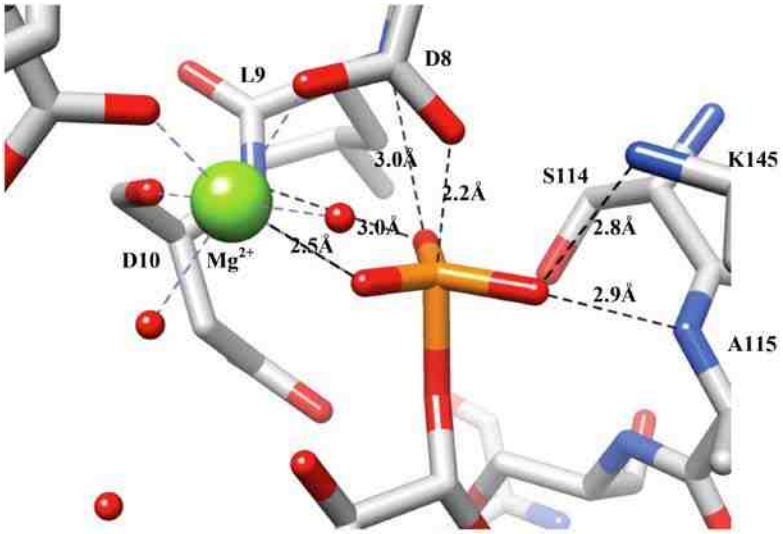


Figure 1.14 X-Ray crystal structure showing the intermediate in the active site of β-phosphoglucomutase catalyzed reaction of β-glucose-1,6-bisphosphate. The interactions between intermediate and the residues depicted as dashed lines.

1.3 Doctoral Research Goals

The doctoral research studies described in this dissertation focus on selected phosphatase members of the HADSF. The enzymes selected as targets in these studies have been chosen based on their interesting physiological roles, unknown functions and mechanistic features, and their ability to reveal information that will aid the biological function assignments of HADSF phosphatases. The specific phosphatases include AphA, a Class B acid phosphatase from *Escherichia coli*, a human mitochondrial deoxyribonucleotidase, the sugar acid phosphatase YidA from *Escherichia coli*, and YigB (EFI-501262) and YbjI(501335) from *Escherichia coli*, BT2542 (501088) from *Bacteroides*

thetaiotaomicron and Q9RUP0 (501193) from *Deinococcus radiodurans*. The major goal of the effort was to gain information about the nature of the physiological substrates for these phosphatases and how the enzymes promote hydrolysis reactions of phosphate esters.

1.3.1 Mechanistic Studies of AphA, a Class B Acid Phosphatase from Escherichia coli

AphA is a nonspecific acid phosphohydrolase (NSAP) from *Escherichia coli*. The first E coli. AphA was isolated in 1994[(50)] from the strain MG1655, belongs to the molecular class B of NASPs. AphA is a homotetrameric metalloenzyme, containing four identical subunits of about 25kDa each [(51)]. The first 25 amino acids, comprising a signal peptide region, function to export the protein into the periplasmic space. AphA is believed to be a HADSF phosphatase owing to the fact that it contains the four conserved signature loops present in members of the HADSF. No detail mechanistic studies have been carried out with AphA. Several complexes crystal structures (AphA with the substrate analog inhibitor 9-[(R)-2-(phosphono-methoxy)ethyl] adenine (PMEA), transition-state analog aluminum fluoride (AphA-Al) and intermediate analog beryllium fluoride (AphA-Be)[(52)]) have been determined that provide a direct insight of the catalytic site at each step of the reaction..

In Chapter 2 of this dissertation, experiments designed to define the energy profile of AphA-catalyzed reactions as well as the transition state structure are described. The effort has provided information about the detailed mechanism of AphA catalyzed hydrolysis reactions.

1.3.2 Mechanistic Studies of Human Mitochondrial Deoxyribonucleotidase

Cells need all the four deoxyribonucleotide triphosphates (dNTPs), including dADP, dCDP, dGDP and dTTP in order to replicate and repair DNA. Some specific diseases such as severe immune deficiency[(53)], apoptotic destruction of B and/or T cells[(54)] and neurogastrointestinal encephalomyopathy[(55)], are caused by inbalances in the dNTP pool. Several enzymes are involved in the synthesis of deoxyribonucleotides and, consequently, in regulating the supply of the four deoxyribonucleotide triphosphates.

The first known mammalian nucleotidase, mdN, was purified from human placenta and identified using isotope flow experiments [(56)]. The activity of this enzyme towards 5'-deoxyribonucleotides was found to be higher than towards 5'-ribonucleotides. Through the use of fluorescence experiments with a dN-green fluorescence protein fusion (GFP) protein in HEK 293 cells it was shown that mdN is located in mitochondria [(57)].

Sequence studies of the first identified mammalian nucleotidase, mdN[(58)], revealed that it contains the four signature loops that are highly conserved in HADSF members. Moreover, analysis of the crystal structures of mdN complex with bound phosphate and thymidine and with the intermediate mimic, beryllium trifluoride, led to a proposal for the general catalytic mechanism of mdN catalyzed hydrolysis of nucleotide monophosphates(6). It was suggested that, like with all other HADSF proteins[(59)], the reaction proceeds via an "in line" nucleophilic substitution pathway involving two nucleophiles. Asp41 serves as the first nucleophile, forming an aspartyl-phosphate intermediate and water is the second nucleophile involved in hydrolysis of the Asp41-mdN intermediate. However, no detailed studies have been conducted to elucidate kinetically and energetically the steps involved in this pathway or to determine the exact

structure of the enzyme intermediate. Thus, the studies I have carried out with this enzyme, described in Chapter 3, focus on these issues along with determining if mdN displays substrate specificity.

1.3.3 Structure Function Analysis of the Sugar Acid Phosphatase YidA from *Escherichia coli*

The haloacid dehalogenase superfamily (HADSF) is a ubiquitous family of enzymes [(34)]. HADSF members exist in all kingdoms of life with numerous proteins present in each organism and catalyze phosphate monoester. Now, more than 4,000 gene sequences encoded HADSF proteins are deposited in gene bank, while only a fraction of these have defined structure and/or function.

As mentioned above, the structural elements required for catalysis and substrate recognition physically separated into different domains in HADSF members. This phenomenon, which facilitates new function acquisition while conserving the catalytic scaffold, serves a possible reason for why the HADSF is highly evolvable. In Chapter 4, studies of the *Escherichia coli* HADSF phosphatase, YidA, which contains a C2 cap domain, are described. The aim of this effort was to elucidate how the cap domain facilitates the functional evolution and, specifically, to uncover the structural determinants that govern expansion of the HADSF and to formulate a unified model for the evolutionary process that has led to the ubiquitous HADSF.

1.3.4 Biological Function Assignment of HADSF Phosphatases

As the genome sequencing technology has become more sophisticated and efficient, greater numbers of protein sequences have been deposited in the database as reflected in the fact that approximately 200,000,000 sequence records exist in the GenBank as of

April 2011. The abundance of protein sequences aids efforts that focus on defining the metabolism and physiology of an organism and those that are aimed at gaining a more thorough understanding of the proteins functions. However, only a small fraction of the proteins, whose sequences have been determined, have been subjected to experimental studies. To cope with the dramatic increase in sequences, computational predictions have been developed to auto-annotate the functions of these proteins. This annotation method is based on the assumption that proteins having similar sequences will have similar functions. However, the results of several studies have demonstrated that the reliability of the computed functional assignments is questionable and, as a result, a number of new proteins have been mis-annotated. Therefore, a more reliable method for protein function assignment method is required.

The Enzyme Function Initiative (EFI)[(60)], a large collaborative and interdisciplinary project, has been established to remedy this problem by developing general methods to assign the functions of unknown proteins. The EFI's strategy is comprised of three steps, including, computational predictions based high throughput screening results to narrow the range of possible substrates, subsequent enzymological studies to test predictions, and biological investigations aimed at elucidating the in vivo functions of the proteins. The HADSF was chosen as a component of the EFI program because of its substrate diversity. In Chapter 5, the initial results of EFI related studies I have carried out, which combines experimental structural biology and computation-based experimental enzymology approaches to define the function of four members of the HAD superfamily, are described. In this effort, the question of diversified functional assignment within the HAD superfamily was addressed by comparing orthologs. For this purpose, four HADSF

phosphatases from different species of bacteria were chosen based on high through screening (HTPS) results which show all of these four have a high catalytically specificity towards riboflavin 5'-monophoaphate (FMN). The four proteins are YigB (EFI-501262) and YbjI (EFI-501335) from *Escherichia coli*, BT2542 (EFI-501088) from *Bacteroides thetaiotaomicron* and Q9RUP0 (EFI-501193) from *Deinococcus radiodurans*.

Reference

1. Zhang, Z. Y. (2002) Protein tyrosine phosphatases: Structure and function, substrate specificity, and inhibitor development, *Annu Rev Pharmacol 42*, 209-234.
2. Mumby, M. C., and Walter, G. (1993) Protein-Serine Threonine Phosphatases - Structure, Regulation, and Functions in Cell-Growth, *Physiol Rev 73*, 673-699.
3. Camps, M., Nichols, A., and Arkinstall, S. (2000) Dual specificity phosphatases: a gene family for control of MAP kinase function, *Faseb J 14*, 6-16.
4. Baumer, N., Maurer, A., Krieglstein, J., and Klumpp, S. (2007) Expression of protein histidine phosphatase in Escherichia coli, purification, and determination of enzyme activity, *Methods Mol Biol 365*, 247-260.
5. Sigal, Y. J., McDermott, M. I., and Morris, A. J. (2005) Integral membrane lipid phosphatases/phosphotransferases: common structure and diverse functions, *The Biochemical journal 387*, 281-293.
6. Millan, J. L. (2006) Alkaline Phosphatases : Structure, substrate specificity and functional relatedness to other members of a large superfamily of enzymes, *Purinergic signalling 2*, 335-341.
7. Bull, H., Murray, P. G., Thomas, D., Fraser, A. M., and Nelson, P. N. (2002) Acid phosphatases, *J Clin Pathol-Mol Pa 55*, 65-72.
8. Engstrom, L. (1961) Studies on calf-intestinal alkaline phosphatase II. Incorporation of inorganic phosphate into a highly purified enzyme preparation, *Biochim. Biophys. Acta 52*, 49-59.

9. Schwartz, J. H., and Lipmann, F. (1961) Phosphate Incorporation into Alkaline Phosphatase of E Coli, *P Natl Acad Sci USA 47*, 1996-&.

10. Gettins, P., Metzler, M., and Coleman, J. E. (1985) Alkaline phosphatase. 31P NMR probes of the mechanism, *J Biol Chem 260*, 2875-2883.

11. Van Etten, R. L. (1982) Human prostatic acid phosphatase: a histidine phosphatase, *Annals of the New York Academy of Sciences 390*, 27-51.

12. Guan, K. L., and Dixon, J. E. (1991) Evidence for protein-tyrosine-phosphatase catalysis proceeding via a cysteine-phosphate intermediate, *J Biol Chem 266*, 17026-17030.

13. Hollfelder, F., and Herschlag, D. (1995) The Nature of the Transition-State for Enzyme-Catalyzed Phosphoryl Transfer - Hydrolysis of O-Aryl Phosphorothioates by Alkaline-Phosphatase, *Biochemistry 34*, 12255-12264.

14. Hall, A. D., and Williams, A. (1986) Leaving Group Dependence in the Phosphorylation of Escherichia-Coli Alkaline-Phosphatase by Monophosphate Esters, *Biochemistry 25*, 4784-4790.

15. Varvoglis, A. J. K. a. A. G. (1967) A. J. Kirby and A. G. Varvoglis, *J. Am. Chem. Soc. 89*, 415-423.

16. Zalatan, J. G., and Herschlag, D. (2006) Alkaline phosphatase mono- and diesterase reactions: comparative transition state analysis, *Journal of the American Chemical Society 128*, 1293-1303.

17. Asthagiri, D., Dillet, V., Liu, T., Noodleman, L., Van Etten, R. L., and Bashford, D. (2002) Density functional study of the mechanism of a tyrosine phosphatase: I.

Intermediate formation, *Journal of the American Chemical Society 124*, 10225-10235.

18. Zhao, Y., and Zhang, Z. Y. (1996) Reactivity of alcohols toward the phosphoenzyme intermediate in the protein-tyrosine phosphatase-catalyzed reaction: Probing the transition state of the dephosphorylation step, *Biochemistry-Us 35*, 11797-11804.

19. Weiru Wang, H. S. C., Rosalind Kim, Jaru Jancarik, Hisao Yokota, Henry H. Nguyen, Igor V. Grigoriev, David E. Wemmer and Sung-Hou Kim. (2002) Structural Characterization of the Reaction Pathway in Phosphoserine Phosphatase: Crystallographic "snapshots" of Intermediate States, *J. Mol. Biol. 319*, 421–431.

20. Gonzalez, M. A., Webb, M. R., Welsh, K. M., and Cooperman, B. S. (1984) Evidence that catalysis by yeast inorganic pyrophosphatase proceeds by direct phosphoryl transfer to water and not via a phosphoryl enzyme intermediate, *Biochemistry-Us 23*, 797-801.

21. Mueller, E. G., Crowder, M. W., Averill, B. A., and Knowles, J. R. (1993) Purple Acid-Phosphatase - a Diiron Enzyme That Catalyzes a Direct Phospho Group Transfer to Water, *Journal of the American Chemical Society 115*, 2974-2975.

22. Butlerransohoff, J. E., Rokita, S. E., Kendall, D. A., Banzon, J. A., Carano, K. S., Kaiser, E. T., and Matlin, A. R. (1992) Active-Site Mutagenesis of Escherichia-Coli Alkaline-Phosphatase - Replacement of Serine-102 with Nonnucleophilic Amino-Acids, *J Org Chem 57*, 142-145.

23. Bloch, W., and Gorby, M. S. (1980) Catalytic mechanism of Escherichia coli alkaline phosphatase: resolution of three variants of the acyl-enzyme mechanism, *Biochemistry-Us 19*, 5008-5018.

24. Han, R., and Coleman, J. E. (1995) Dependence of the phosphorylation of alkaline phosphatase by phosphate monoesters on the pKa of the leaving group, *Biochemistry-Us 34*, 4238-4245.

25. McCain, D. F., Catrina, I. E., Hengge, A. C., and Zhang, Z. Y. (2002) The catalytic mechanism of Cdc25A phosphatase, *J Biol Chem 277*, 11190-11200.

26. Zhang, Z. Y., and Vanetten, R. L. (1991) Pre-Steady-State and Steady-State Kinetic-Analysis of the Low-Molecular-Weight Phosphotyrosyl Protein Phosphatase from Bovine Heart, *J Biol Chem 266*, 1516-1525.

27. Kim, E. E., and Wyckoff, H. W. (1991) Reaction mechanism of alkaline phosphatase based on crystal structures. Two-metal ion catalysis, *Journal of molecular biology 218*, 449-464.

28. Applebur.Ml, Johnson, B. P., and Coleman, J. E. (1970) Phosphate Binding to Alkaline Phosphatase - Metal Ion Dependence, *J Biol Chem 245*, 4968-&.

29. Zalatan, J. G., Catrina, I., Mitchell, R., Grzyska, P. K., O'Brien, P. J., Herschlag, D., and Hengge, A. C. (2007) Kinetic isotope effects for alkaline phosphatase reactions: Implications for the role of active-site metal ions in catalysis, *Journal of the American Chemical Society 129*, 9789-9798.

30. Welsh, K. M., Jacobyansky, A., Springs, B., and Cooperman, B. S. (1983) Catalytic Specificity of Yeast Inorganic Pyrophosphatase for Magnesium-Ion as

Cofactor - an Analysis of Divalent Metal-Ion and Solvent Isotope Effects on Enzyme Function, *Biochemistry-Us 22*, 2243-2248.

31. Baykov, A. A., Shestakov, A. S., Kasho, V. N., Vener, A. V., and Ivanov, A. H. (1990) Kinetics and Thermodynamics of Catalysis by the Inorganic Pyrophosphatase of Escherichia-Coli in Both Directions, *Eur J Biochem 194*, 879-887.

32. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D., and Aravind, L. (2006) Evolutionary genomics of the HAD superfamily: Understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes, *J Mol Biol 361*, 1003-1034.

33. Hisano, T., Hata, Y., Fujii, T., Liu, J. Q., Kurihara, T., Esaki, N., and Soda, K. (1996) Crystal structure of L-2-haloacid dehalogenase from Pseudomonas sp YL - An alpha/beta hydrolase structure that is different from the alpha/beta hydrolase fold, *J Biol Chem 271*, 20322-20330.

34. Koonin, E. V., and Tatusov, R. L. (1994) Computer-Analysis of Bacterial Haloacid Dehalogenases Defines a Large Superfamily of Hydrolases with Diverse Specificity - Application of an Iterative Approach to Database Search, *J Mol Biol 244*, 125-132.

35. Collet, J. F., Stroobant, V., Pirard, M., Delpierre, G., and Van Schaftingen, E. (1998) A new class of phosphotransferases phosphorylated on an aspartate residue in an amino-terminal DXDX(T/V) motif, *J Biol Chem 273*, 14107-14112.

36. Koonin, E. V., and Tatusov, R. L. (1994) Computer analysis of bacterial haloacid dehalogenases defines a large superfamily of hydrolases with diverse specificity. Application of an iterative approach to database search, *J Mol Biol 244*, 125-132.

37. Aravind, L., Galperin, M. Y., and Koonin, E. V. (1998) The catalytic domain of the P-type ATPase has the haloacid dehalogenase fold, *Trends Biochem Sci 23*, 127-129.

38. Morais, M. C., Zhang, G., Zhang, W., Olsen, D. B., Dunaway-Mariano, D., and Allen, K. N. (2004) X-ray crystallographic and site-directed mutagenesis analysis of the mechanism of Schiff-base formation in phosphonoacetaldehyde hydrolase catalysis, *J Biol Chem 279*, 9353-9361.

39. Zhang, G., Mazurkie, A. S., Dunaway-Mariano, D., and Allen, K. N. (2002) Kinetic Evidence for a Substrate-Induced Fit in Phosphonoacetaldehyde Hydrolase Catalysis, *Biochemistry 41*, 13370-13377.

40. Lahiri, S. D., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2002) Caught in the act: the structure of phosphorylated beta-phosphoglucomutase from Lactococcus lactis, *Biochemistry 41*, 8351-8359.

41. Wang, W., Cho, H. S., Kim, R., Jancarik, J., Yokota, H., Nguyen, H. H., Grigoriev, I. V., Wemmer, D. E., and Kim, S. H. (2002) Structural characterization of the reaction pathway in phosphoserine phosphatase: crystallographic "snapshots" of intermediate states, *J Mol Biol 319*, 421-431.

42. Kim, H. Y., Heo, Y. S., Kim, J. H., Park, M. H., Moon, J., Kim, E., Kwon, D., Yoon, J., Shin, D., Jeong, E. J., Park, S. Y., Lee, T. G., Jeon, Y. H., Ro, S., Cho, J. M., and Hwang, K. Y. (2002) Molecular basis for the local conformational

rearrangement of human phosphoserine phosphatase, *J Biol Chem 277*, 46651-46658.

43. Peeraer, Y., Rabijns, A., Verboven, C., Collet, J. F., Van Schaftingen, E., and De Ranter, C. (2003) High-resolution structure of human phosphoserine phosphatase in open conformation, *Acta Crystallogr D Biol Crystallogr 59*, 971-977.

44. Allen, K. N., and Dunaway-Mariano, D. (2004) Phosphoryl group transfer: evolution of a catalytic scaffold, *Trends Biochem Sci 29*, 495-503.

45. Kuznetsova, E., Proudfoot, M., Gonzalez, C. F., Brown, G., Omelchenko, M. V., Borozan, I., Carmel, L., Wolf, Y. I., Mori, H., Savchenko, A. V., Arrowsmith, C. H., Koonin, E. V., Edwards, A. M., and Yakunin, A. F. (2006) Genome-wide analysis of substrate specificities of the Escherichia coli haloacid dehalogenase-like phosphatase family, *J Biol Chem 281*, 36149-36161.

46. Ridder, I. S., and Dijkstra, B. W. (1999) Identification of the Mg2+-binding site in the P-type ATPase and phosphatase members of the HAD (haloacid dehalogenase) superfamily by structural similarity to the response regulator protein CheY, *Biochemical Journal 339*, 223-226.

47. Collet, J. F., Stroobant, V., and Van Schaftingen, E. (2002) Evidence for phosphotransferases phosphorylated on aspartate residue in N-terminal DXDX(T/V) motif, *Methods in enzymology 354*, 177-188.

48. Lu, Z., Wang, L., Dunaway-Mariano, D., and Allen, K. N. (2009) Structure-function analysis of 2-keto-3-deoxy-D-glycero-D-galactonononate-9-phosphate phosphatase defines specificity elements in type C0 haloalkanoate dehalogenase family members, *J Biol Chem 284*, 1224-1233.

49. Lahiri, S. D., Zhang, G. F., Dunaway-Mariano, D., and Allen, K. N. (2003) The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction, *Science 299*, 2067-2071.

50. Rossolini, G. M., Thaller, M. C., Pezzi, R., and Satta, G. (1994) Identification of an Escherichia-Coli Periplasmic Acid-Phosphatase Containing of a 27 Kda-Polypeptide Component, *Fems Microbiol Lett 118*, 167-173.

51. Calderone, V., Forleo, C., Benvenuti, M., Thaller, M. C., Rossolini, G. M., and Mangani, S. (2004) The first structure of a bacterial class B acid phosphatase reveals further structural heterogeneity among phosphatases of the haloacid dehalogenase fold, *J Mol Biol 335*, 761-773.

52. Leone, R., Cappelletti, E., Benvenuti, M., Lentini, G., Thaller, M. C., and Mangani, S. (2008) Structural Insights into the Catalytic Mechanism of the Bacterial Class B Phosphatase AphA Belonging to the DDDD Superfamily of Phosphohydrolases, *J Mol Biol 384*, 478-488.

53. Giblett, E. R., Ammann, A. J., Wara, D. W., Sandman, R., and Diamond, L. K. (1975) Nucleoside-phosphorylase deficiency in a child with severely defective T-cell immunity and normal B-cell immunity, *Lancet 1*, 1010-1013.

54. Cohen, A., Hirschhorn, R., Horowitz, S. D., Rubinstein, A., Polmar, S. H., Hong, R., and Martin, D. W. (1978) Deoxyadenosine Triphosphate as a Potentially Toxic Metabolite in Adenosine-Deaminase Deficiency, *P Natl Acad Sci USA 75*, 472-476.

55. Nishino, I., Spinazzola, A., and Hirano, M. (1999) Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder, *Science 283*, 689-692.

56. Rampazzo, C., Johansson, M., Gallinaro, L., Ferraro, P., Hellman, U., Karlsson, A., Reichard, P., and Bianchi, V. (2000) Mammalian 5 '(3 ')-deoxyribonucleotidase, cDNA cloning, and overexpression of the enzyme in Escherichia coli and mammalian cells, *Journal of Biological Chemistry 275*, 5409-5415.

57. Rampazzo, C., Gallinaro, L., Milanesi, E., Frigimelica, E., Reichard, P., and Bianchi, V. (2000) A deoxyribonucleotidase in mitochondria: Involvement in regulation of dNTP pools and possible link to genetic disease, *P Natl Acad Sci USA 97*, 8239-8244.

58. Rinaldo-Matthis, A., Rampazzo, C., Reichard, P., Bianchi, V., and Nordlund, P. (2002) Crystal structure of a human mitochondrial deoxyribonucleotidase, *Nat Struct Biol 9*, 779-787.

59. Allen, K. N., and Dunaway-Mariano, D. (2004) Phosphoryl group transfer: evolution of a catalytic scaffold, *Trends Biochem Sci 29*, 495-503.

60. Gerlt, J. A., Allen, K. N., Almo, S. C., Armstrong, R. N., Babbitt, P. C., Cronan, J. E., Dunaway-Mariano, D., Imker, H. J., Jacobson, M. P., Minor, W., Poulter, C. D., Raushel, F. M., Sali, A., Shoichet, B. K., and Sweedler, J. V. (2011) The Enzyme Function Initiative, *Biochemistry-Us 50*, 9950-9962.

CHAPTER TWO

2. Mechanistic Studies with AphA, a Class B Acid Phosphatase of *Escherichia coli*

2.1 Introduction

Nonspecific acid phosphohydrolases (NSAPs) are bacterial enzymes that have optimal catalytic activities at acidic to neutral pH values and that do not exhibit a marked substrate specificity in that they promote hydrolysis reactions of several structurally different phosphoesters. NSAPs are typically divided into three molecular classes, A, B and C, based on their sequence homologies. *S.enterica ser.typhimurium* AphA, a member of Class B, was one of the first bacterial NSAPs purified and characterized [(1)]. An E coli. AphA, isolated in 1994[(2)] from the strain MG1655, was found to belong to molecular class B of NSAPs. Interestingly, class B NSAPs belong to the DDDD superfamily of phosphohydrolases, members of which share some conserved sequence motifs including four highly conserved aspartate residues and which have an exceptionally large lineage with representatives in the bacteria, archaea and eukaryote kingdoms.

2.1.1 Structure of AphA

AphA enzyme is secreted in the periplasmic space of *E. coli* and as a homotetrameric metalloenzyme (Figure 2.1), which contains four identical subunits of about 25kDa each[(3)]. X-ray crystal structure analysis showed that the tetramer is comprised of intertwining N-terminal sequences from each monomer. In addition, the first 25 amino acids form a signal peptide that functions to export the protein and is then cleaved when the protein arrives in the periplasmic space.

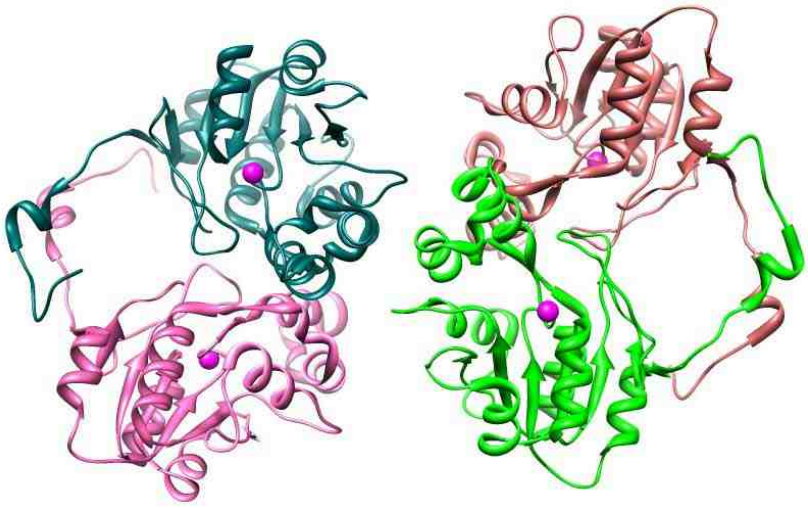


Figure 2.1 Quaternary structure of AphA. Each subunit is represented in different colors. The Mg^{2+} ions are represented as magenta spheres.

AphA also contains four conserved signature loops that are common among members of the HADSF (Figure 2.2). Inspection of the AphA apo crystal structure shows that a catalytic metal ion Mg^{2+} is coordinated to the carboxylate groups of Asp44, Asp167, the carbonyl oxygen of Asp46, and three water molecules with a facial arrangement (Figure 2.3A) forming an octahedral coordination sphere. In 2006[(4)] the crystal structure of AphA complexed with the adenosine and inorganic phosphate, the products of AMP hydrolysis, was determined. In this structure (Figure 2.3B) the two products are bound at different locations within the active site with inorganic phosphate replacing one of the three water molecules coordinated to Mg^{2+} in the apo-protein structure. The side chains of nucleophile Asp44, general acid/base Asp46, Try112 and Lys152 form several hydrogen bonds with phosphate. Separately, Asp44 constitutes motif 1 of HADSF, Asp46 constitutes motif 2, and Thr112 and Lys152 constitute motif 3 and motif 4 of HADSF.

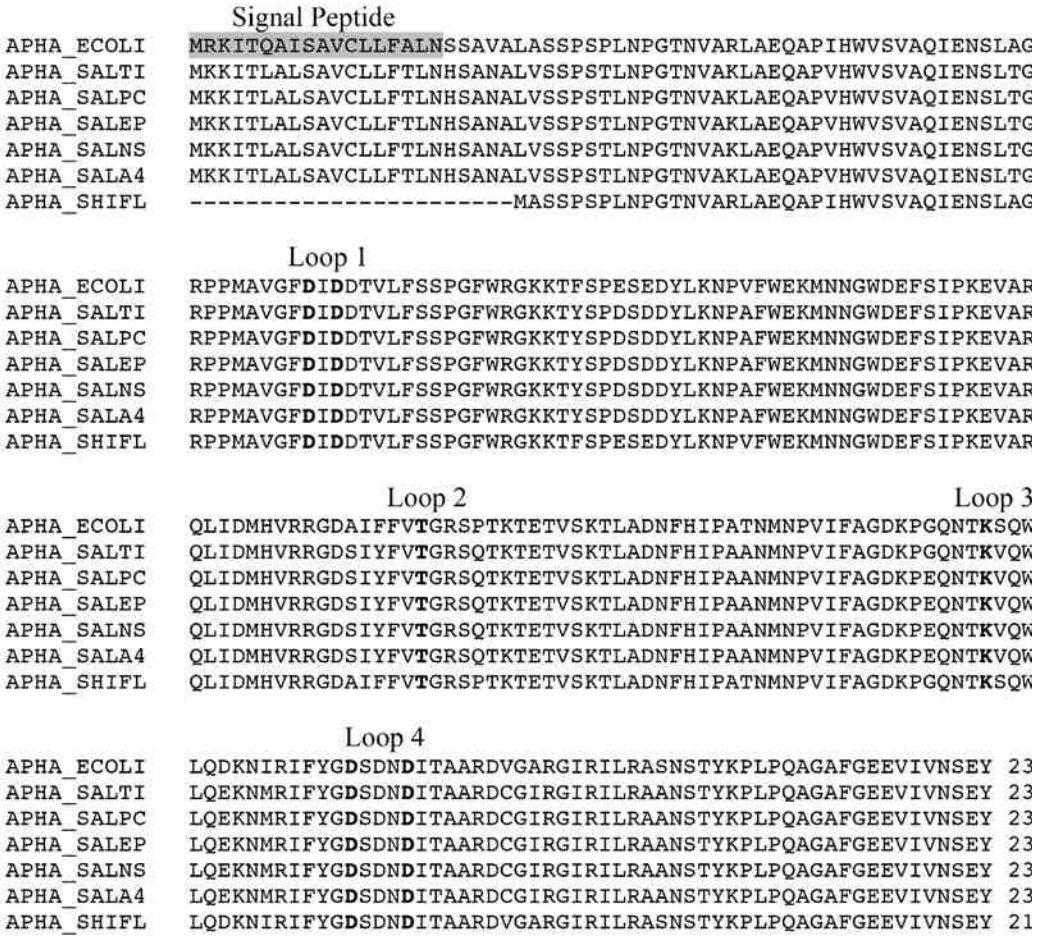
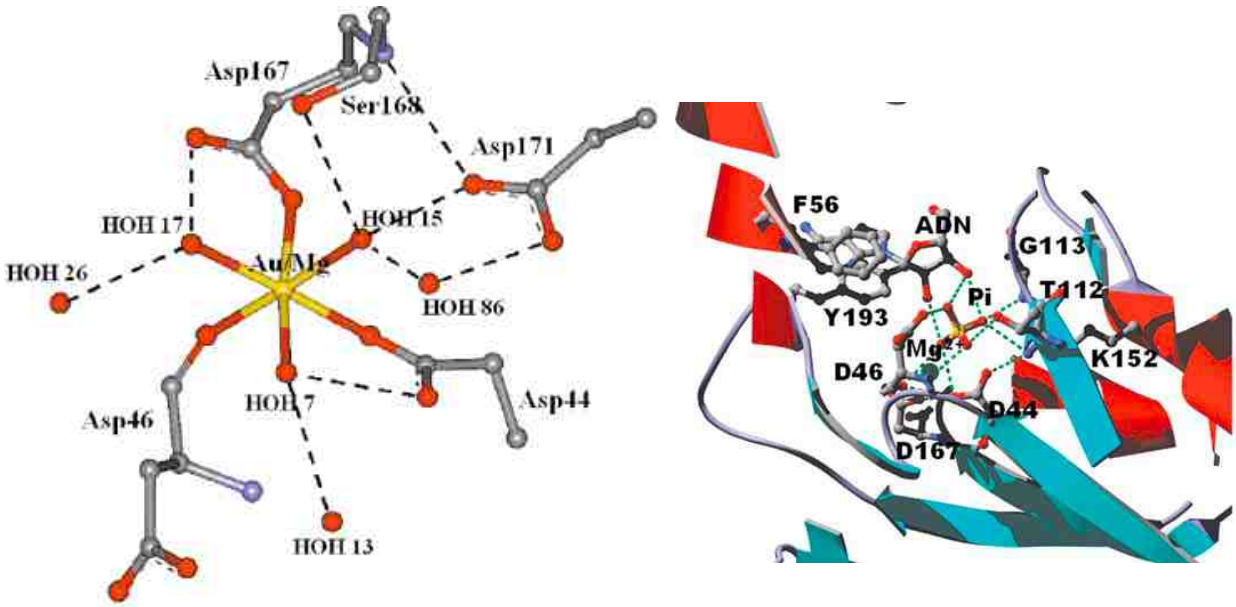


Figure 2.2 Sequence alignments of acid phosphatases on E coli AphA. The highlighted part is the signal peptide of AphA and the bold are the conserved residues in HADSF. APHA_ECOLI, Escherichia coli AphA; APHA_SALTI, Salmonella typhimurium AphA; APHA_SALPC, Salmonella paratyphi C; APHA_SALEP, Salmonella enteritidis PT4; APHA_SALNS, Salmonella newport; APHA_SALA4, Salmonella agona; APHA_SHIFL, Shigella flexneri AphA.

The OH2 and OH3 hydroxyl groups of the adenosine ribose ring are also H-bonded to phosphate (Figure 2.3 B). In addition, the adenosine aromatic ring resides in a hydrophobic cleft formed by Phe56, Tyr193, Tyr70 and Leu71. This hydrophobic binding site is also proposed to be the location of the aromatic moiety of the substrate AMP since the distance between the adenosine ribose ring and the cleaved phosphate is close. Finally, the presence of the hydrophobic binding cleft suggests that AphA prefers aromatic phosphate esters and nucleotides as its physiological substrates.



(A) (B)

Figure 2.3 (A) Schematic of the metal binding site of native AphA and (B) of hydrogen binding interactions between products adenosine, phosphate and the active site of AphA.

2.1.2 Biological Function of Acid Phosphatase AphA

The biological role played by *E coli.* AphA has not yet been fully determined. While in the periplasmic space, which lacks 5'-nucleotidases and alkaline phosphatases, AphA has

been proposed to function as a scavenging enzyme to transform 5'-UMP to pyrimidines[5]. Later a more detailed study[6] revealed that *Salmonella enterica* AphA is required in the nicotinamide mononucleotide (NMN) assimilation. Periplasmic AphA promotes removal of the phosphate group of NMN in its conversion to nicotinamide ribonucleoside (NmR). NmR is then imported into the cell by the PnuC transporter where it serves as a source of pyridine. *E coli.* AphA could play a similar role. *E coli.* AphA is also a DNA binding protein able to sequestrate transforming DNA penetrating from outside[7]. Another interesting observation is that *E coli.* AphA is involved in biosynthetic pathways for purine and thiamine synthesis[8]. Together with periplasmic L-asparaginase II (AnsB), AphA could be responsible for the production of phosphoribosylamine (PRA), the first intermediate in the purine and thiamine synthesis pathway.

2.1.3 Catalytic Mechanism of Acid Phosphatase AphA

Detailed mechanistic studies have not been conducted thus far with AphA. In spite of this deficiency, crystal structures of complexes of AphA with the substrate analog, inhibitor 9-[(R)-2-(phosphono-methoxy)ethyl]adenine (PMEA) (Figure 2.4), transition-state analog aluminum fluoride (AphA-Al) and intermediate analog beryllium fluoride (AphA-Be) have provided some preliminary information about the catalytic mechanism[9]. The crystal structure of AphA-PMEA (Figure 2.4) shows that PMEA, an acyclic adenosine derivative, which replaces the phosphoribosyl moiety of the nucleoside by an ethoxymethylphosphonic group, is involved in forming hydrogen bonds with a Mg^{2+} coordinated water and acts in second sphere binding to Mg^{2+}. Also analysis of this structure demonstrates that the hydrophobic pocket of AphA, defined by Leu71, Phe56 and Tyr193 residues, acts as the substrate specificity site. Inspection of the AphA-Be and

AphA-Al complexes indicates that the overall reaction catalyzed by AphA proceeds through a covalent acyl phosphate intermediate and involves an associative S_N2 type mechanism.

Studies I have carried out to gain more detailed information about the catalytic mechanism of AphA are described and discussed below.

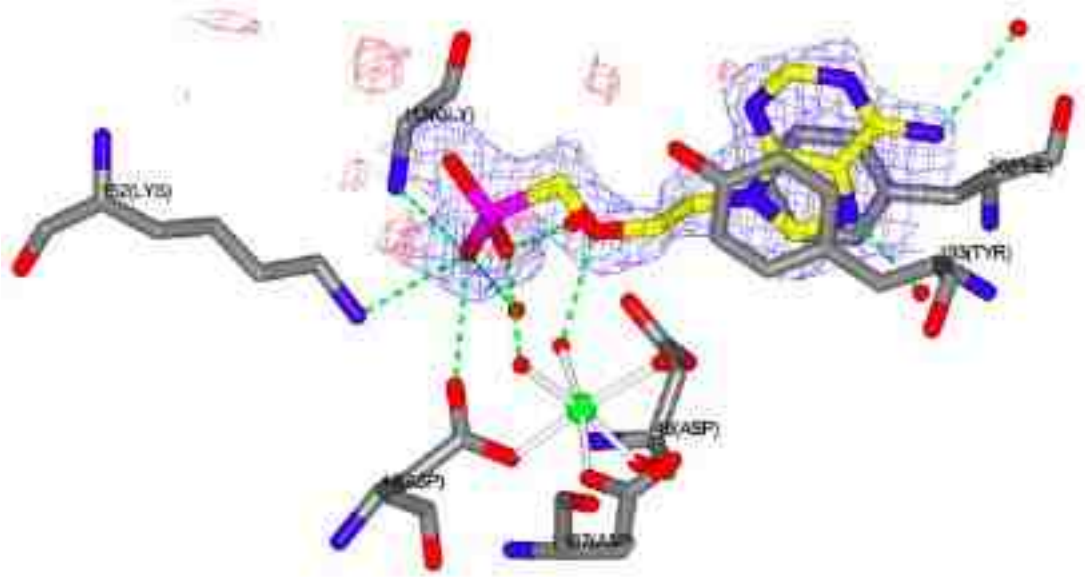


Figure 2.4 Crystal structure of AphA-PMEA active site. PMEA is represented by yellow carbon atoms. The Mg^{2+} is the green sphere, H2O is in red and hydrogen bonds are displayed by green dash lines.

2.2 Experimental

2.2.1 Materials

All chemicals were purchased from Sigma. Primers, T4 DNA ligase, and restriction enzymes were purchase from Invitrogen. *Pfu* polymerase, the pET23a vector kit and host

cells were from Stratagene. Genomic DNA from *Escherichia coli* k12 (ATCC 10798) was purchased from ATCC. Deuterium oxide was purchased from Cambridge Isotope Laboratories. ^{14}CAMP was purchased from American Radiolabeled Chemicals, Inc. 10k centrifugal filters were purchased from VWR.

2.2.2 Cloning, Expression, and Purification of AphA

The cDNA encoding the gene HAD protein AphA (NCBI accession NC000913, Swiss Pro accession: P0AE22) from *E coli* was amplified by PCR using the genomic DNA from *E. coli* K12 *(*ATCC 10798) and P*fu* DNA polymerase. The PCR product was cloned into the PET-23a vector, which was used to transform competent *E. coli*. BL21 (DE3) cells. Oligonucleotide primers (5'-AGGGAAACATATGCGCAAGA) and (5'-ACGCTCTCCCAAGCTTTCAGTATTCTGAAT) containing restriction endo- nuclease cleavage sites *NdeI* and *HindIII* were used in the PCR reactions. Plasmid DNA was purified using a Qiaprep Spin Miniprep Kit. The gene sequencing carried out at the Center for Genetics in Medicine, University of New Mexico School of Medicine.

Gene expression was performed using competent *E. coli* BL21 (DE3) cells. Following growth in 3 L of TB media (1x) containing 60 mM glucose and 50 μg/mL ampicillin at 37 °C for 6 h with mixing at 200 rpm, induction was initiated with the addition of 0.2 mM isopropyl α-D-thiogalactopyranoside (IPTG). After incubation for 12 h at 20 °C and 160 rpm, the cells were harvested by centrifugation (6500 rpm for 15 min at 4 °C), and then lysed by using a modified spheroplasting technique[(10)]. Accordingly, the cell pellet was suspended in 200 mM Tris-HCl (pH 8, 4 °C) (1 g wet cells/ 2.5 mL buffer). To this mixture an equal volume of 200 mM Tris-HCl, pH 8 containing 1 M sucrose was added. Then EDTA and lysozyme were added to reach a final concentration of 10 mM and 4

mg/g cells respectively. After adding an equal volume of deionized water the mixture was incubated at room temperature for 30 min, after which $MgCl_2$ was added to a concentration of 10 mM. The mixture was centrifuged at 20000 r.p.m for 45 min. $(NH_4)_2SO_4$ was added to the supernatent to 40% (w/v) and the precipitate was removed by centrifugation. The AphA was precipitated by increasing the $(NH_4)_2SO_4$ concentration to 50% and harvested by centrifugation. The pellet was dissolved in 10 mM Tris-HCl (pH 7.5, 4 °C) containing 5 mM $MgCl_2$ and then loaded onto a 5mL HiTrap DEAE FF column (GE healthcare) equilibrated with the same buffer. The flow-through fractions containing the target enzyme were pooled and dialyzed against 6 L 10 mM Na^+MES (pH 5.5, 4 °C), containing 5 mM $MgCl_2$ before loading onto a 1mL HiTrap CM FF column equilibrated with the dialysis buffer. The flow-through fractions were pooled and concentrated to ~2 mL and then dialyzed against 50 mM Na^+MES (pH 6, 4 °C) buffer containing 5 mM $MgCl_2$. The dialysate was chromatographed at 4 °C on a HiLoad 16/60 Superdex 200 column using the same dialyzed buffer as eluant. The fractions containing protein were pooled and concentrated.

2.2.3 Preparation of AphA Site-Directed Mutagenesis

The AphA site directed mutants, D71N, F81A, F81W, Y218A, Y218W and F81W/Y218W were prepared by site-directed mutagenesis. The pET-28a vector including AphA without the first 25 amino acids was used as the template. The mutagenesis primers used for mutants are listed in Table 2.1. Plasmid DNAs were purified using a Qiaprep Spin Miniprep Kit and following confirmation of the mutant gene sequence by commercial DNA sequencing (MClab).

Table 2.1 Primers used in AphA site-directed mutagenesis.

D71N	5’－GGGGTTTGATATCAATGACACGGTAC
F81A	5’－AGTCCGGGCGCCTGGCGCGGCAAAAAAACC
F81W	5’－TTCCAGTCCGGGCTTCTGGCGCGGCAAAAAA
Y218A	5’－TCCAACTCTACCGCCAAACCCTTGCCACAA
Y218W	5’－AACTCTACCTGGAAACCCTTGCCACAAG

Gene expression was performed using *E. coli* BL21 (DE3) competent cells. The transformed cells were grown at 37 °C with agitation at 200 rpm in Luria broth containing 50 μg/mL ampicillin for 4-6 h to an OD_{600nm} = 0.6-1.0 and induced for 6 h at 20 °C at a final concentration of 0.4 mM isopropyl α-D-thiogalactopyranoside (IPTG). The cells were harvested by centrifugation at 6500 rpm for 15 min at 4 °C to yield 3 g/L of culture medium. The cell pellet was resuspended in 1g wet cell/10 mL of ice-cold lysis buffer (20 mM Tris•HCl, 40 mM imidazole, 1 mM DTT and 500 mM NaCl, pH 7.5). The cell suspension was passed through a French press at 1,200 PSIG before centrifugation at 20,000 rpm and 4°C for 45 min. The supernatant was loaded onto a Ni-NTA column at a flow rate of 5 mL/min. The column was washed with 5CV (column volume) lysis buffer followed by 5CV of wash buffer (20 mM Tris•HCl, 500 mM NaCl, 500 mM imidazole and 1mM DTT at pH 7.5). The column fractions were analyzed by SDS-PAGE, and the desired fractions were combined and concentrated at 4 °C using a 10K Amicon Ultra Centrifugal filter (Millipore) and then stored at -80 °C.

2.2.4 Steady State Kinetic Constant Determination

The steady-state kinetic parameters (K_m and k_{cat}) of AphA were determined from initial

reaction velocities measured at varying substrate concentrations for reactions in 5 mM $MgCl_2$ in 50 mM Na^+MES assay buffer (pH 6.0) at 25 °C. Protein concentrations were determined by using the Bradford assay[(11)] and the product inorganic phosphate was determined by using the acidified ammonium molybdate method[(12)]. Data were fit using the SigmaPlot program to equation 2.1,

$$V_0 = V_{max} [S] / (K_m + [S]) \qquad \text{(eq. 2.1)}$$

where V_0 = initial velocity, V_{max} = maximum velocity, [S] = substrate concentration and K_m = Michaelis-Menten constant for substrate. The k_{cat} value was calculated from V_{max} and [E] according to the equation $k_{cat} = V_{max} /[E]$, where [E] is the free enzyme concentration.

The rate of *p*-nitrophenyl phosphate (pNPP) hydrolysis was determined by monitoring the increase in absorbance at 410 nm (ε = 1.86 mM^{-1} cm^{-1}) at 25°C. The 0.5 mL assay mixtures contained 50 mM Na^+MES, pH 6.0, 5 mM $MgCl_2$, and various concentrations of pNPP.

2.2.5 Determination of Inhibition Constants

Competitive and noncompetitive inhibition constants (K_i) of products inhibitors were determined by measuring the initial velocity of AphA catalyzed substrate hydrolysis as a function of substrate concentration (K_m to $5K_m$) and inhibitor concentration (0, $1xK_i$ and $2xK_i$ µM). The initial velocity data were fit using KinetAsyst (IntelliKinetics, PA) to equation 2.2 (competitive inhibition) and equation 2.3 (noncompetitive inhibition),

$$V = V_{max} [A] / ([A] + K_m (1 + [I]/K_i)) \qquad \text{(eq. 2.2)}$$

$$V = V_{max} [S] / [K_m (1+[I]/K_i) + [S] (1+[I]/K_{ii})] \qquad \text{(eq. 2.3)}$$

where [A] is the substrate concentration, V is the initial velocity, V_{max} is the maximum

velocity, K_m is the Michaelis constant, K_i and K_{ii} are the inhibition constant and [I] is the inhibitor concentration.

2.2.6 Single-Turnover Kinetic Experiment

The time course of the AphA catalyzed reaction converting [^{14}C]-AMP to [^{14}C]-adenosine was evaluated under single-turnover conditions using a KinTek rapid-mixing chemical-quench apparatus. The resulting mixture was passed through a 10-kDa filter to remove the enzyme, and then chromotographed using HPLC with a C18 reverse (4 mm × 250 mm) column and the following elution profile: mobile phase A was 0.1 mol/L KH_2PO4 (pH 5), and mobile phase B was 70% 0.01 mol/L KH_2PO4 (pH 3.5) and 30% methanol. 95% mobile phase A and 5% mobile phase B was maintained for 3 min, a 2-min linear gradient to 50% mobile phase A was initiated, 50% mobile phase A was maintained for 18 min, a 2-min linear gradient to 95% mobile phase A was initiated, and 95% mobile phase A was maintained for at least 10 min before the next sample was injected. Radioactivities, determined by using a β-RAM model 4 radio flow-through detector (IN/US System, Inc), were used to calculate the concentration of [^{14}C] AMP (retention time of 9 min) and [^{14}C] adenosine (retention time of 18 min) present in the reaction mixture at the time of quenching, by multiplying the initial AMP concentration by the fraction of the total radioactivity. The observed rate constants for the single turnover reactions were obtained by fitting the time course data to the first order equations 2.4 and 2.5 using the computer programs Kaleidagragh

$$[P]_t = [P]_{max}(1 - e^{-kt}) \qquad \text{(eq. 2.4)}$$

$$[S]_t = [S]_{max} - ([P]_{max}(1 - e^{-kt})) \qquad \text{(eq. 2.5)}$$

where *k* is the first-order rate constant; $[P]_t$ and $[S]_t$ are the product and substrate concentrations at time " *t* ", respectively; $[S]_{max}$ is the initial concentration of substrate; $[P]_{max}$ is the product concentration at equilibrium..

A typical experiment was carried out by mixing 16 µL buffer A (50 mM MES-NaOH, pH 6.0 containing 5 mM $MgCl_2$) containing 100 µM AphA and 16 µL buffer A containing 20 µM [^{14}C]-AMP. The reaction was quenched after a specified period of time with 212 µL 0.2 N HCl (if the reaction time was shorter than 75 ms, then additional 0.2 N HCl was added to adjust the total volume of the reagent to 212 µL). The enzyme was removed by using a 10-kD filter and then the reactants and products were separated by HPLC as described above.

2.2.7 Multi-turnover Reaction of AphA with [^{14}C(U)]-AMP

Multi-turnover reactions were performed at 25 °C using a rapid quench instrument from KinTek Instruments. Each reaction was carried out by mixing 16 µL buffer A (50 mM MES (pH 6.0) containing 5 mM $MgCl_2$) containing 40 µM AphA and 16 µL buffer A containing 200 µM [^{14}C(U)]-AMP. The reaction was quenched after a specified period of time with 212 µL 0.2 N HCl. The mixture was passed through a 10-kDa filter to remove the enzyme, and then subjected to HPLC using a C18 reverse (4 mm × 250 mm) column. The column was eluted at flow rate of 1 mL/min by using the same elution program used for the single turnover reaction (see above) and a β-RAM model 4 radio flow-through detector (IN/US System, Inc.) was used to measure the radioactive contents of each species. The observed rate constants and k_{cat} for steady state multi turnover reactions were obtained by fitting the time course data using the computer programs Kaleidagragh to the burst equations[(13)] 2.6 and 2.7,

$$[P] = A * (1 - e^{-k_{obs} t}) + k * t \qquad \text{(eq. 2.6)}$$

$$k = k_{cat} * [E]_0 \qquad \text{(eq. 2.7)}$$

where [P] is the observed concentration of product adenosine, A is the burst amplitude (A_0) times the active enzyme concentration ($[E]_0$), k_{obs} is the burst rate constant, k is the steady-state turnover rate (approaching k_{cat} at saturating substrate) times the active enzyme concentration ($[E]_0$).

2.2.8 Stopped-Flow UV-*vis* Absorption Kinetic Experiments Measured under Multiple Turnover Conditions

A DX.17MV sequential stopped-flow spectrometer (Applied Photophysics, Leatherhead, U. K.) with a light path of 10 mm and a dead time of 2 ms was used for the determination of transient rate constants for AphA catalyzed hydrolysis of *para*-nitrophenyl phosphate (pNPP). The light source was a 150-W xenon lamp with a slit width of 1 mm. The drive syringes were driven by a pneumatic actuator driven by compressed nitrogen.

Solutions of AphA (40 to 120 μM after mixing) and pNPP (500 μM after mixing) in 50 mM Na^+MES buffer containing 5 mM $MgCl_2$, pH 6 were mixed in the stopped flow apparatus, while monitoring absorbance changes at 410 nm. The observed steady state rate constants and k_{cat} for multi turnover reactions were obtained by fitting the time course data using the computer programs Kaleidagragh to equation 2.8, where

$$OD = \varepsilon E_0 [A_0(1-e^{-k_{obs}t}) + k_{cat}t] \qquad \text{(eq. 2.8)}$$

OD is the absorbance at time t, ε is extinction coefficient at 410 nm, E_0 is the concentration of enzyme active sites, k_{obs} is the first order rate constant of the fast and slow reaction steps in the reaction pathway, k_{cat} is the steady-state rate constant.

2.2.9 Measurement of AMP Binding Constant to Wild Type AphA and the Y218W Mutant

Reactions were performed using an Applied Photophysics SX20 stopped-flow spectrophotometer. The excitation wavelength was 280 nm. The observed first-order rate constants k_{obs} was obtained by fitting the data using the software provided by Applied Photophysics to equation 2.9,

$$F_t = F_\infty + P \exp(-k_{obs}t) \qquad \text{(eq. 2.9)}$$

where F_t = the fluorescence at time t, F_∞ = fluorescence at infinite time, P = amplitude, k_{obs} = the observed first-order rate constant.

Time courses for fluorescence quenching at different ligand concentration (2-40 μM after mixing) and fixed enzyme concentration (5 μM Y218W and 1 μM Apha after mixing) were determined to obtain association rate constants (k_{on}) and dissociation rate constants (k_{off}) for AMP binding to wild type AphA and the Y218W mutant. The resulting k_{obs} values were plotted against the ligand concentrations ([L]) and the data were fitted to equation 2.10 in order to define the k_{on} and k_{off} values.

$$k_{obs} = k_{on}[L] + k_{off} \qquad \text{(eq. 2.10)}$$

2.2.10 HPLC Assay of the Phosphotransferase Activity of AphA

In the AphA phosphotransferase assay, pNPP was used as phosphate donor and (^{14}C/U) adenosine was the phosphate accepter. Reactions, carried out in 50 mM Na^+MES buffer containing 5 mM $MgCl_2$, pH 6.0, at room temperature, were stopped by adding concentrated HCl after 10 min. The precipitated protein was removed using 10 kDa centrifugal filter (VWR), and the pH of the filtrate was adjusted to 7 by addition of 1 N NaOH. The solutions were subjected to HPLC analysis using a Shimadzu LC-20AB

binary solvent delivery unit with a SPD-20AV UV detector, equipped with Ultra Aqueous C18 column (RESTEK, 5 μM particle size, 250 x 4.6 mm). The elution program was the same as that desribed above, the flow rate was 1mL/min, and the absorbance of the eluent was monitored at 260 and 280 nm. In addition, eluent radioactivity was monitored by using a β-RAM model 4 radio flow-through detector (IN/US System, Inc).

2.2.11 Proton Inventory Measurement

The proton inventory study of AphA catalyzed hydrolysis of pNPP was conducted using a series of H_2O and D_2O buffer mixtures, in which the deuterium fractions varied from 0.0 to 1.0. For this purpose, stock solutions of 50 mM Na^+MES with 5 mM $MgCl_2$ in pure H_2O (pH = 6.0) or pure D_2O (pD = 6.0) were prepared and then mixed to obtain the proper deuterium fractions. The pL (L = H or D) for each H_2O/D_2O buffers, determined using the equation $(\Delta pH)_n = 0.076n^2+0.3314n$, where n is molar fraction of D_2O[(14)], was 6.0. Initial velocities at various substrate concentrations covering 0.5 K_m-5 K_m were measured in each H_2O/D_2O mixed buffer system and k_{cat} and K_m values were calculated as described above. The concentration of pNPP was varying from 50 to 800 μM, and the concentration of AphA was 0.0488 μM.

2.2.12. Alcohol Effect on AphA Hydrolysis of pNPP

The initial velocities of AphA-catalyzed pNPP were measured by monitoring the absorbance change at 410 nm, caused by the appearance of the p-nitrophenolate chromophore ($\Delta\varepsilon$=1.86 mM^{-1} cm^{-1}), in the presence of different concentrations of alcohols (0-1.0 M) in 50 mM Na^+MES buffer, containing varying concentrations (0.5 – 5 K_m) of pNPP (50-1600 μM) and 5 mM $MgCl_2$ (pH 6.0 and 25°C). The concentration of AphA used was varying from 0.017 μM to 0.034 μM.

The alcohols used were ethanol, 1-propanol, 2-methoxyethanol, 2-chloroethanol, allyl alcohol, 2,2,2-trifluoroethanol, 3-hydroxypropionitrile and propargyl alcohol. The k_{cat} value for each alcohol concentration was obtained by using SigmaPlot program as described in section 2.2.4. An intercept of a linear plot of k_{cat} against alcohol concentration is the rate constant of hydrolysis. And the slope of this line is the second order rate constant for the reaction of E-P with the alcohol[15].

$$k_{cat} = k_3 + k_4[ROH] \qquad \text{(eq. 2.11)}$$

2.2.13 pH/pD Rate Profile Determination

Initial velocities of *E.coli.* AphA (0.008 to 0.024 μM) catalyzed hydrolysis reaction of AMP (1.67 μM to 300 μM) at 25 ^{o}C were determined by measuring the amount of released inorganic phosphate using the acidified ammonium molybdate method[12]. For pD profile the reaction of AphA (0.0488 μM) catalyzed pNPP (50-1000 μM) was used. And the initial velocities were monitored by measuring the absorbance increase at 410nm ($\Delta\varepsilon$ = 1.86 mM^{-1} cm^{-1}) caused by the production of pNP at 25°C. The buffer system consisting of reaction solutions containing 50 mM 2-(N-morpholino)ethanesulfonate (MES) at pH/pD 5.0-6.0, 50 mM HEPES at pH/pD 6.5-7.0 with 5mM $MgCl_2$ was used to measure the pH dependence of the catalyzed reaction. The K_m and v_{max} were acquired by fitting the rates into equation 2.1 using SigmaPlot program, and k_{cat} was calculated from the ratio of v_{max} and the enzyme concentration, which was determined using Bradford assay[11].

2.2.14 Mass Spectrometric Analysis of the AphA Promoted Hydrolysis Reaction of AMP in Presence of Ethanol

The reaction was initiated by adding 10 mM AMP and 0.1425 μM AphA in 10 mM NH_4Ac, pH 6.0 buffer containing 1 mM $MgCl_2$, 1M ethanol. After 1 h, the reaction was quenched by adding 1 M NaOH. After removing the protein by using a 10 kDa centrifugal filter (VWR), the filtrate was acidified by adding 5 μL of 4 N HCl. Lypholization gave a solid that was subjected to MS analysis (University of the New Mexico mass spectral facility).

2.2.15 Steady-State Multi-Turnover Time Course Measurements

Steady-state reactions under multi-turnover conditions were performed using different concentrations of AMP (1.5, 2.6 4.6 μM) and 0.006 μM wild type AphA in 50 mM Na^+MES buffer containing 5 mM $MgCl_2$, pH 6.0, at 25°C. Product formation was monitored by adding the coupling enzyme, adenosine deaminase, and detecting the absorbance decrease at 265 nm caused by the conversion of product adenosine to inosine ($\Delta\varepsilon = 8.4\ mM^{-1}\ cm^{-1}$).

2.3 Results and Discussions

2.3.1 Purification of Wild Type AphA

The gene encoding AphA, a periplasmic HAD class B acid phosphatase from *E. coli (K-12),* was cloned for over expression in *E. coli* BL21 (DE3) cells transformed with the AphA/pET-23a plasmid clone. The cell pellet was collected 24 h after induction, and the cells were lysed and subjected to purification. The purity of protein was found to be ~90% as determined by SDS-PAGE analysis. The protein concentration was determined by

using the Bradford method and by measuring the absorbance at 280 nm. The final yield of AphA is 4 mg protein/g wet cells.

The mass spectrum of purified AphA showed it contained two different polypeptides of molecular masses of 23528.6 and 23712.8, corresponding to two different cleavage sites (23 and 25) of the signal peptide from the AphA precursor[16]. This finding is in close agreement with the theoretical masses of 23529.4and 23713.6, calculated by using ExPASy.

2.3.2 AphA Substrate Specificity

The steady state kinetic constants, k_{cat} and K_m, for AphA catalyzed hydrolysis of several phosphoesters were determined (Table 2.2). Inspection of the k_{cat}/K_m values (or specificity constant) shows that E coli. wild type AphA has highest activities towards mononucleotides ($k_{cat}/K_m \sim 10^5$) except for 5'-dCMP and CMP, which has a specificity constant 10-fold lower ($\sim 10^4$). In contrast, this enzyme has low or no activity towards dinucleotides and trinucleotides, like ADP, ATP, ITP, UTP, dTTP, GDP, UDP and IDP. The aryl monophosphate pNPP serves as a very good substrate ($k_{cat}/K_m = 1.8 \times 10^5$) even though the overall binding affinity is relatively low (K_m = 80 μM). Coenzyme A (CoA) and β–NADP, both of which contain an adenosine ring, are also substrates ($k_{cat}/K_m \sim 10^4$). The activity of AphA towards o-phosphotyrosine and riboflavin-5-phosphate, which have an aryl group, is relatively high ($k_{cat}/K_m \sim 10^4$). Sugar phosphates and glycerol-2-phosphate are poor substrates with high K_m values (>2 mM), and K_m of AphA hydrolyzing ADP and D,L-glyceraldehyde-3-phosphate are higher than 50 mM. Finally AphA displays no detectable activity towards ATP, ITP, UTP, dTTP, GDP, UDP and IDP.

Table 2.2 The steady-state kinetic constants k_{cat} and K_m for AphA hydrolyzing various substrates are measured in 50 mM Na^+MES buffer containing 5mM $MgCl_2$, pH 6.0, 25 ^oC.

Substrate	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
5'- GMP	2.9 ± 0.2	4.2 ± 0.1	1.4×10^6
3'- AMP	11.0 ± 0.5	6.9 ± 0.1	6.3×10^5
5'- AMP	9.7 ± 0.6	2.6 ± 0.1	2.7×10^5
5'- IMP	28.0 ± 0.7	6.5 ± 0.1	2.3×10^5
2'- AMP	6.7 ± 0.6	1.3 ± 0.1	2.0×10^5
pNPP	80.0 ± 5.7	14.6 ± 0.5	1.8×10^5
5'- TMP	52.0 ± 2.0	5.8 ± 0.1	1.1×10^5
5'- UMP	25.8 ± 2.2	2.3 ± 0.1	8.7×10^4
β-NADP	100.1 ± 2.8	6.7 ± 0.1	6.7×10^4
CoA	17.0 ± 0.6	$(9.0 \pm 0.1) \times 10^{-1}$	5.5×10^4
5'- dCMP	145.7 ± 8.1	3.8 ± 0.1	2.6×10^4
O-Phospho-Tyrosine	$(1.6 \pm 0.2) \times 10^2$	13.5 ± 0.5	8.5×10^4
Riboflavin-5-phosphate*	$(6.5 \pm 0.3) \times 10^3$	$(9.5 \pm 0.2) \times 10^1$	1.5×10^4
D-Fructose-6-phosphate	$(4.9 \pm 0.4) \times 10^3$	47.2 ± 1.6	9.5×10^3
Ribose-5-phosphate	$(4.4 \pm 0.2) \times 10^3$	37.1 ± 1.0	8.4×10^3
5'- CMP	$(3.7 \pm 0.1) \times 10^2$	2.1 ± 0.1	5.7×10^3
D-Glucosamine-6-phosphate	$(6.0 \pm 0.3) \times 10^3$	14.6 ± 0.4	2.5×10^3
Glycerol-2-phosphate	$(2.9 \pm 0.1) \times 10^3$	$(30.0 \pm 0.5) \times 10^{-2}$	1.2×10^2

* Measured in ammonium acetate buffer, pH 6.0.

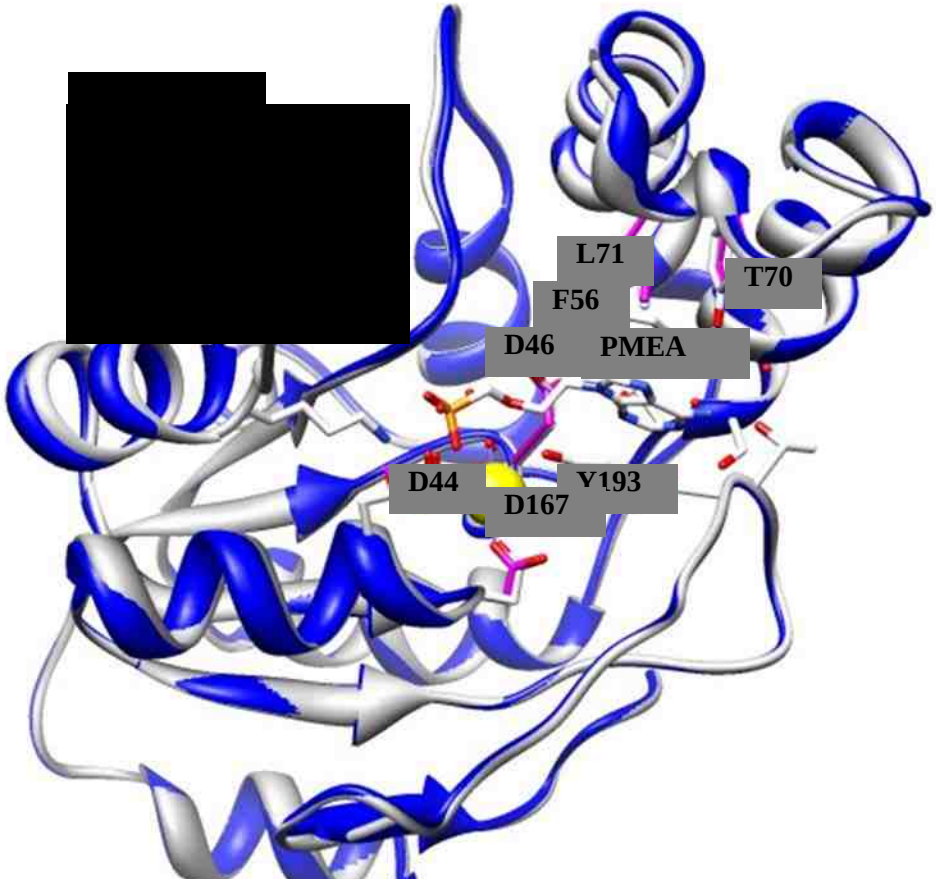


Figure 2.5 Comparison of the crystal structure of native AphA (PDB code: 1N8N) (blue) and the complex of AphA with its substrate analog inhibitor PMEA (PDB code: 2G1A) (grey). The arrows point out the different loop positions in the two structures, and Mg^{2+} is shown as a yellow sphere.

2.3.3 Construction of Substrate Binding Site Mutants and Determination of Their Steady-State Rate Constants

The structural characterization of wild type AphA with its substrate analog inhibitor 9-[(R)-2-(phosphono-methoxy) ethyl] adenine (PMEA) (Figure 2.4) by the Mangani [(9)] group provided a direct insight into the substrate leaving group binding site. Like all the

Table 2.3 Steady-state kinetic constants k_{cat} and K_m for wild type and mutant AphA catalyzed hydrolysis of AMP in 50 mM Na^+MES buffer containing 5mM $MgCl_2$, pH 6.0, 25 ^oC.

Protein	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
Wild Type Apha	9.7 ± 0.6	2.6 ± 0.1	2.7×10^5
D71N	4.0 ± 0.3	$(3.0 \pm 0.1) \times 10^{-3}$	7.9×10^2
F81A	190.1 ± 9.7	$(1.9 \pm 0.5) \times 10^{-1}$	1.0×10^3
F81W	72.0 ± 7.5	4.7 ± 0.2	6.5×10^4
Y218A	220.2 ± 9.8	$(4.0 \pm 0.1) \times 10^{-2}$	1.7×10^2
Y218W	11.8 ± 0.7	3.2 ± 0.1	2.7×10^5
F81W/Y218W	20.3 ± 1.0	5.8 ± 0.1	2.9×10^5

other HADSF members the crystal structure of AphA[17] shows that the protein consists of two parts, including a cap and core domain [3]. In addition, all four highly conserved residues are located in the core domain and the substrate recognition residues are located in the cap domain (Phe56, Leu71 and Tyr70). The structure of the AphA – PMEA complex is almost identical to that of native AphA, (Figure 2.5). The only difference is found in a loop comprised of residues 143-151. The PMEA purine ring is stacked with the aryl rings of Phe56 and Tyr193. This specific hydrophobic binding pocket for the alcohol leaving groups of the phosphoester substrates explains the substrate preference of AphA shown by the data in Table 2.2. The movement of the loop (143-151) may serve

the same function as the cap domain movement in other HADSF members[17], which is to provide an "open" conformation allowing substrates to enter the active site and a "closed" conformation after the substrate is bound to desolvate the active site.

Two residues, Phe81 and Tyr218, appear to play very important roles in the hydrophobic binding cleft of AphA. When the aromatic side chains of these residues are replaced by aliphatic side chains as is the case for mutants F81A and Y218A in Tables 2.3 and 2.4, the catalytic efficiencies are dramatically decreased especially for the native substrate AMP where the efficiencies decrease from 2.7×10^5 $M^{-1}s^{-1}$ to 1.0×10^3 $M^{-1}s^{-1}$ for the F81A mutant and to 1.7×10^2 $M^{-1}s^{-1}$ for the Y218A mutant. The efficiency decreases are caused by increases of K_m as well as decreases of the catalytic turnover rates. Both replacement result in a ~10 fold increase in K_m and a 10-100 fold reduction in k_{cat}, which means that creation of the Ala mutants not only lead to disrupted binding but also to changes in the orientation of the substrate that locate the PO_4^{3-} moiety at an inappropriate location in the reaction site. In contrast, when the side chains in Phe81 and Tyr218 are replaced by another aromatic moiety, as in the F81W and Y218W mutants, only small changes in the catalytic efficiencies take place. Another key residue is Asp71, which is located in the catalytic site and suggested to serve in a general acid/base catalytic role. Replacement of this residue by Asn greatly influences k_{cat} but it does not greatly effect K_m.

Table 2.4 Steady-state kinetic constants k_{cat} and K_m for wild type and mutant AphA catalyzed hydrolysis of pNPP in 50 mM Na^+MES buffer containing 5mM $MgCl_2$, pH 6.0, 25 ^oC.

Protein	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
Wild Type Apha	$(8.0 \pm 0.6) \times 10^1$	14.6 ± 0.5	1.8×10^5
D71N	$(2.1 \pm 0.1) \times 10^1$	$(2.5 \pm 0.1) \times 10^{-2}$	1.2×10^3
F81A	$(7.0 \pm 0.1) \times 10^2$	26.5 ± 0.2	3.8×10^4
F81W	$(2.6 \pm 0.2) \times 10^2$	$(12.4 \pm 0.5) \times 10^1$	5.0×10^5
Y218A	$(3.4 \pm 0.3) \times 10^1$	$(2.6 \pm 0.1) \times 10^{-1}$	7.5×10^3
F81W/Y218W	$(3.1 \pm 0.1) \times 10^2$	56.5 ± 0.8	1.8×10^5
Y218W	$(3.5 \pm 0.5) \times 10^2$	13.8 ± 0.6	3.9×10^4

2.3.4 Optimum pH for AphA Catalysis

The k_{cat} and k_{cat}/K_m pH profiles of E coli. AphA were determined at varying AMP concentration and a saturating Mg^{2+} concentration as a function the pH of the reaction solution. The log k_{cat} and log k_{cat}/K_m pH profiles shown in Figure 2.6 define an optimum pH of 6.0. The kinetic constants are greatly reduced when the pH is higher than 6.0 while the differences below pH 6.0 are small.

The pH dependence of k_{cat} monitors the ionization constants of enzyme-substrate complex ES, whereas the dependence of k_{cat}/K_M monitors the ionization of the uncomplexed enzyme and substrate. The pH profile data suggest that a key ionizable

group loses its function upon deprotonation. While the identity of the residue that is ionized is unkown, it is likely Asp46 that acts as a general acid/base in the overall reaction. And whose pK_a is slightly elevated by the environment of the active site making it exist in a protonated form at slightly acidic pH [(4)].

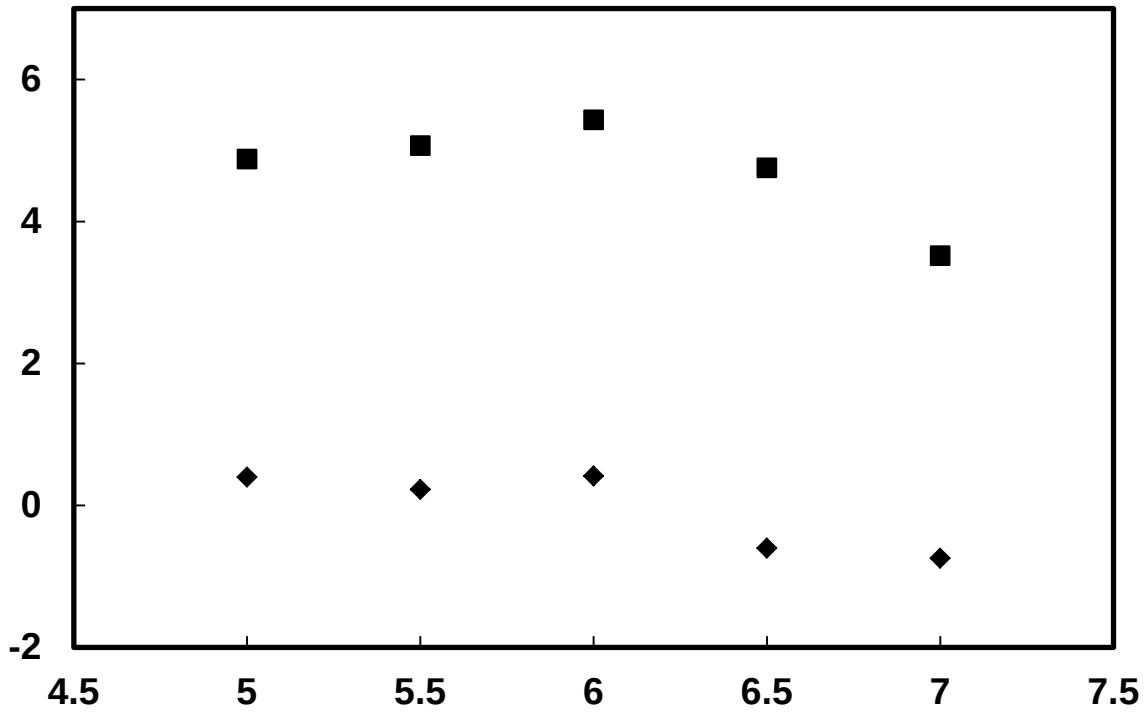


Figure 2.6 The pH dependence of Apha catalysis was measured using AMP as the substrate. The buffer system consisting of reaction solutions were buffered at pH 5.0-6.0 with 50mM MES, pH 6.5-7.0 with 50 mM HEPES containing 5mM $MgCl_2$. ((■) $lg(k_{cat})$, (◆) $\log (k_{cat}/K_m)$).

2.3.5 Deuterium Solvent Isotope Effect on *AphA* Catalyzed pNPP Hydrolysis

The solvent deuterium isotope effect for the AphA catalyzed reaction of PNPP was determined by measuring the k_{cat} and K_m values of the reaction in 50 mM MES, pH/pD

6.0 containing 5 mM $MgCl_2$. The results are summarized in Table 2.5. Because the solvent isotope effect could be caused by either an effect on the reaction rate or perturbations of equilibrium constants of ionizable groups of the substrate or enzyme[(14)], pH/pD rate profiles were determined first. Thus, to avoid equilibrium effects, the isotope effects were determined in plateau regions (Figures 2.6 and 2.7).

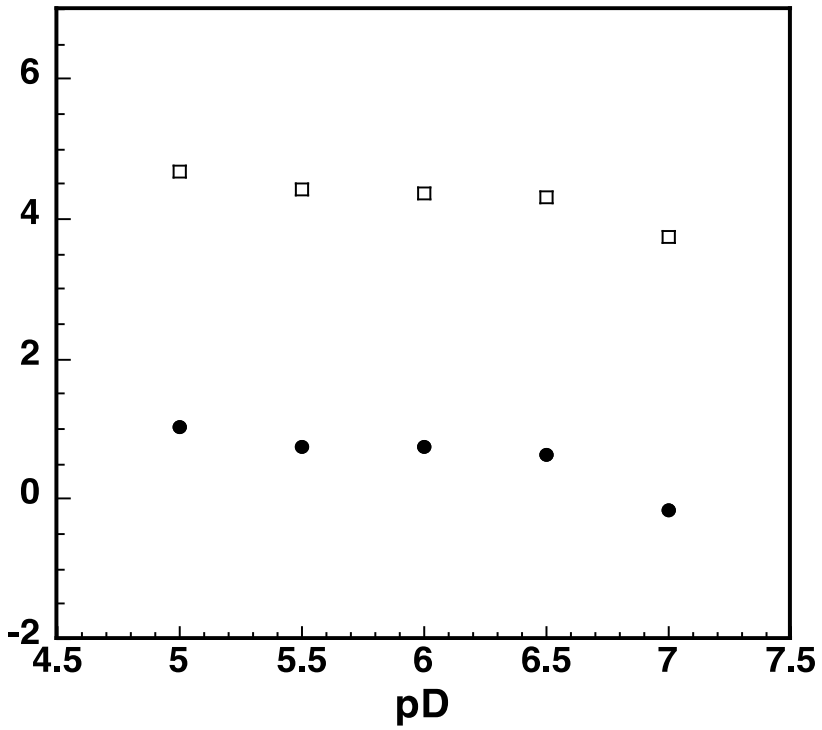


Figure 2.7 The pD dependence of AphA catalyzed reaction of pNPP. The buffer system consisting of reaction solutions were buffered at pH 5.0-6.0 with 50m MES, pH 6.5-7.0 with 50 mM HEPES containing 5mM $MgCl_2$. ((•) log (k_{cat}), (□) log (k_{cat}/K_m))

A significant D_2O solvent isotope effect on k_{cat} was observed (k_{cat}^{H}/k_{cat}^{D} = 4.78) in the pH-independent region while the effect on k_{cat}/K_m was small ($(k_{cat}/K_m)^{H}/(k_{cat}/K_m)^{D}$ = 1.68). The existence of the large solvent isotope effect on rate indicates that proton

transfer is involved in the catalytic step. In addition, the small D_2O solvent isotope effect on k_{cat}/K_m is a consequence of the large decrease of the K_m value, which is associated with an effect on substrate binding.

Table 2.5 Deuterium solvent isotope effect of wild-type AphA catalyzed pNPP hydrolysis at pL (pH or pD) 6.0, 25°C.

k_{cat}^{H} (s)	k_{cat}^{D} (s)	K_m^{H} (μM)	K_m^{D} (μM)	k_{cat}^{H}/k_{cat}^{D}	$(k_{cat}/K_m)^{H}$ /$(k_{cat}/K_m)^{D}$
8.6 ± 0.1	1.8 ± 0.1	(7.0 ± 0.2) x 10^1	24.6 ± 0.8	4.8	1.7

2.3.6 Proton Inventory Study of *AphA* Catalyzed pNPP Hydrolysis

The proton inventory experiment was performed with pNPP as a substrate using five different concentrations of D_2O buffers containing 5 mM $MgCl_2$, pL 6.0, 25 °C. The plot of V_n/V_1 vs the deuterium atomic fraction n was found to give a straight line with a least squares correlation coefficient of 0.999 (Figure 2.8).

$$V_n/V_1 = -2.55\ n + 4.93\ (R^2 = 0.999)$$

Consistent with the results from D_2O solvent isotope effect experiment, the observation made in the proton inventory experiment confirm that proton transfer is involved in the rate determining step of the reaction and that a single proton is in flight in the transition state of this step.

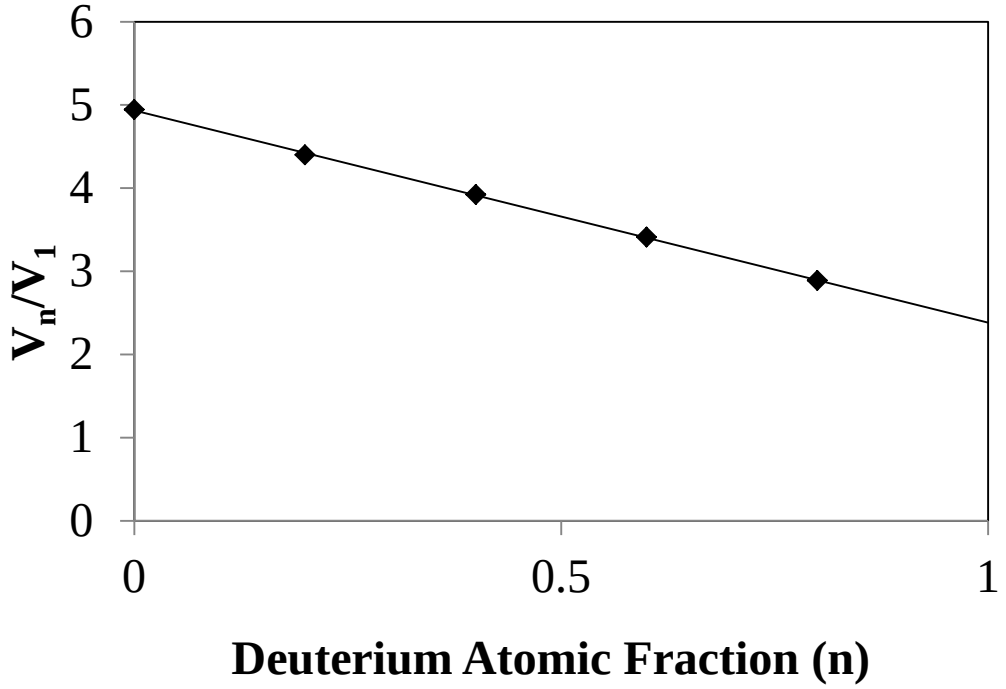


Figure 2.8 Proton inventory graph of V_{max} for the hydrolysis of PNPP at pL = 6.0, 25°C**.** The line is generated by using a linear regression method and the correlation coefficient is 0.99.

2.3.7 HPLC Assay of the Phosphotransferase Activity of AphA

The phosphothansferase activity of *E coli.* AphA was determined by using HPLC to monitor the conversion of (^{14}C/U)-adenosine to ^{14}C –AMP in reactions carried out in the presence of pNPP in 50 mM Na^+MES buffer containing 5mM $MgCl_2$ pH6.0 at room temperature (Figure 2.9). The radioactivity content of (^{14}C/U)-adenosine and ^{14}C –AMP was determined by using a β-RAM model 4 radio flow-through detector (IN/US System, Inc).

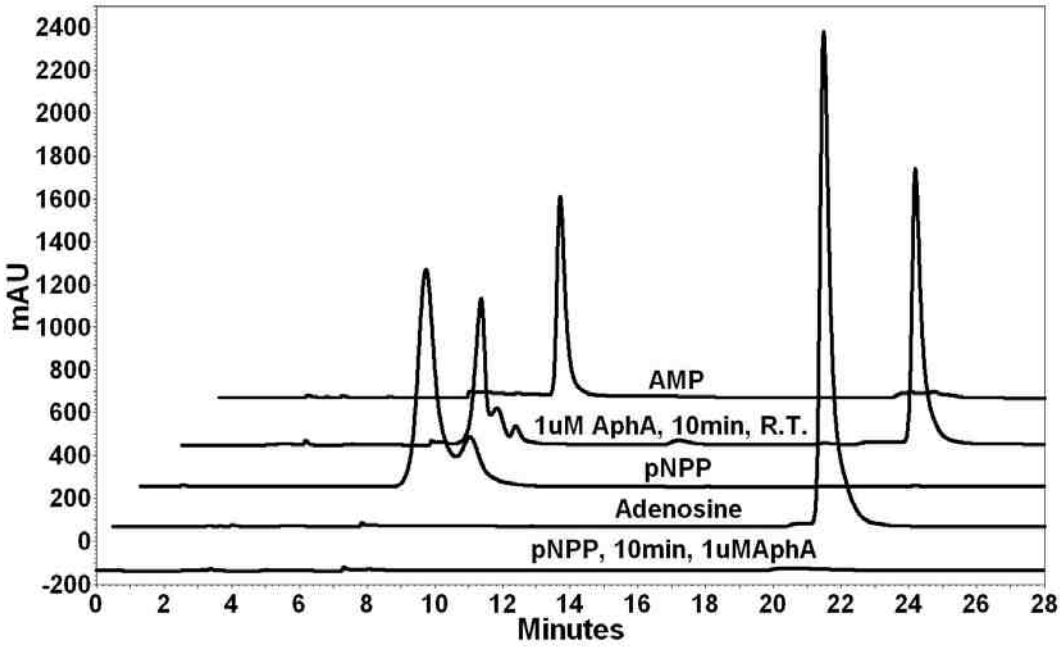


Figure 2.9 HPLC assay of the phosphotransferase activity of AphA with pNPP as phosphate donor and adenosine as the phosphate acceptor in 50 mM MES buffer, pH 6.0, room temperature. The reaction was initiated by adding 1 μM AphA and quenched after 10 min.

The results, shown in Figure 2.9, indicate that no new peak is observed in the HPLC trace of the product mixture after 10 min reaction time when AphA promotes hydrolysis of pNPP in the absence of adenosine. This is expected because the product *p*-nitrophenol (pNP) does not absorb at the wavelength (260 nm) used in the HPLC system. Because the k_{cat} of pNPP hydrolysis by AphA is 14.6 ± 0.5 s^{-1}, after 10 min reaction with 1 μM AphA we expect that all of the substrate pNPP (5mM) was consumed. When adenosine was

included in the reaction mixture, a new peak is formed with a retention time of *ca.* 10 min, which is ssociated with AMP formation. The fact that the AMP peak is associated with its radioactivity content (Figure 2.10) confirms that ^{14}C-AMP is produced form ^{14}C-adenosine. Therefore, *E. coli.* AphA serves as a phosphotransferase catalyst that is capable of transferring a phosphate group from an organic phosphate monoesters to the free 5'-hydroxyl group of adenosine. The decreased reaction rate observed for this phosphory transfer process suggests that in addition to being a phosphate acceptor adenosine is also an inhibitor of the AphA.

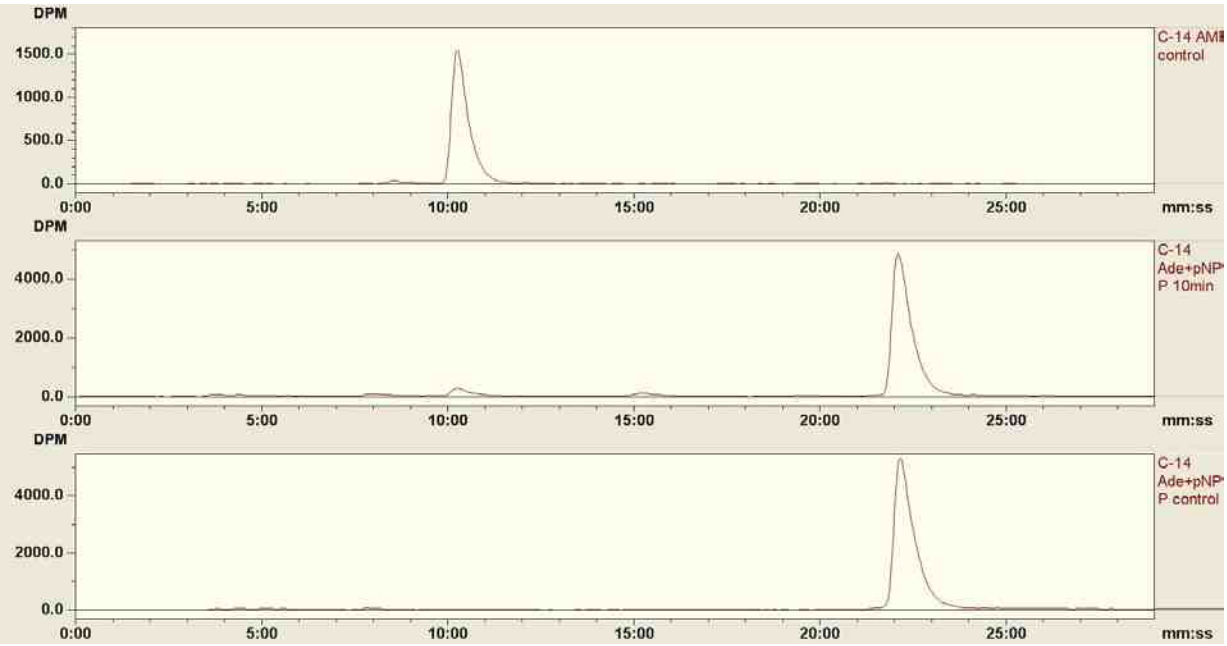
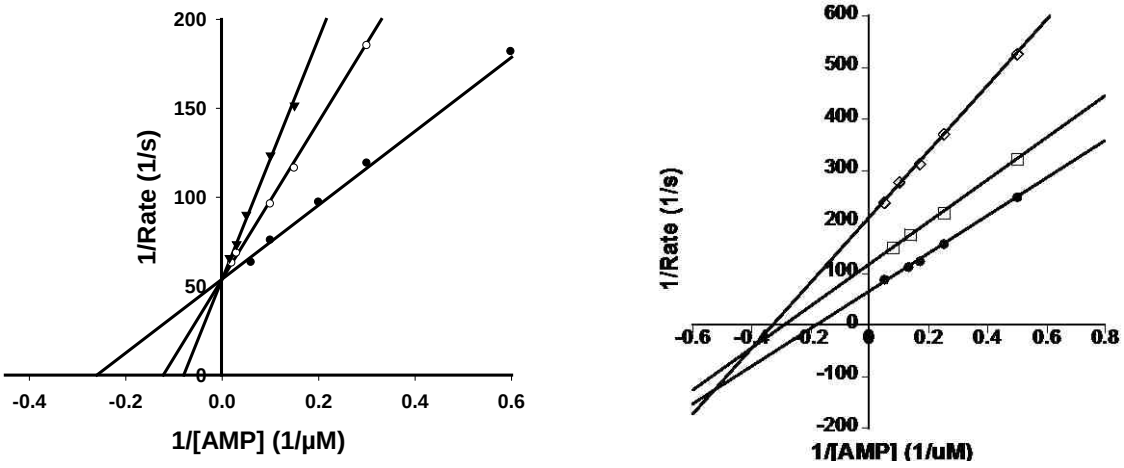


Figure 2.10 The radioactivity of products from HPLC assay. Upper: AMP control; middle: products of AphA catalyzed pNPP in the presence of ^{14}C adenosine; bottom: ^{14}C adenosine control.

2.3.8 Products inhibition studies



(A) (B)

Figure 2.11. Lineweaver-Burk plots showing adenosine products inhibition of AphA catalyzed AMP hydrolysis in 50 mM MES (pH 6.0, 25 ^{o}C) containing 5 mM $MgCl_2$. (A) Adenosine as an inhibitor. (•) 0 μM adenosine; (o) 8.75 μM adenosine; (▼) 17.5 μM adenosine. (B) Inorganic phosphate as an inhibitor. (•) 0 μM PO_4^{3-}; (□) 10 mM PO_4^{3-}; (◇) 20 mM PO_4^{3-}.

Product inhibition studies of AphA catalyzed hydrolysis reactions of AMP and pNPP were performed at 25 °C in the assay buffer (50 mM Na^+MES, 5mM $MgCl_2$, pH 6.0) using substrate concentrations in the range of 0.5-10 K_m. Inhibition by inorganic phosphate of the AphA catalyzed reaction of pNPP was determined by monitoring the absorbance change associated with pNP formation at 410nm, adenosine inhibition of AMP hydrolysis was determined by using the acidified ammonium molybdate assay for inorganic phosphate generation, and inorganic phosphate inhibition of the AphA-catalyzed AMP reaction was determined by using a coupling enzyme method employing

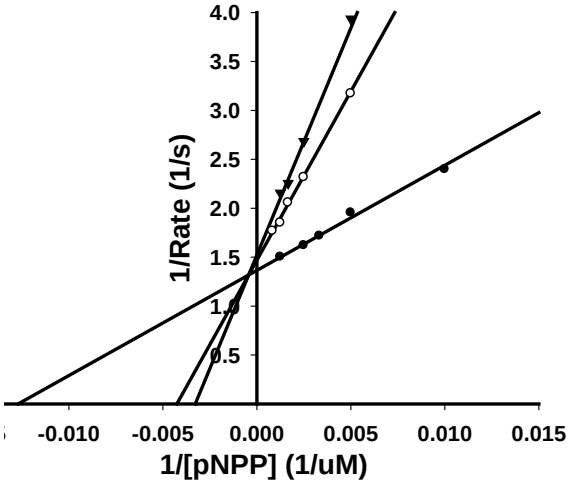


Figure 2.12. Inorganic phosphate inhibition of AphA-catalyzed hydrolysis of pNPP at pH 6.0, 25°C. (•) 0 mM PO_4^{3-}; (o) 20 mM PO_4^{3-}; (▼) 30 mM PO_4^{3-}.

Table 2.6 The rate of AphA hydrolysis of pNPP in the presence of different concentrations of *p*-nitrophenol in the assay buffer, at 25°C.

4-Nitrophenol (mM)	v (μM/s)
0	0.20
1.67	0.22
5	0.24
10	0.24
15	0.23
20	0.23
25	0.23
31.75	0.21

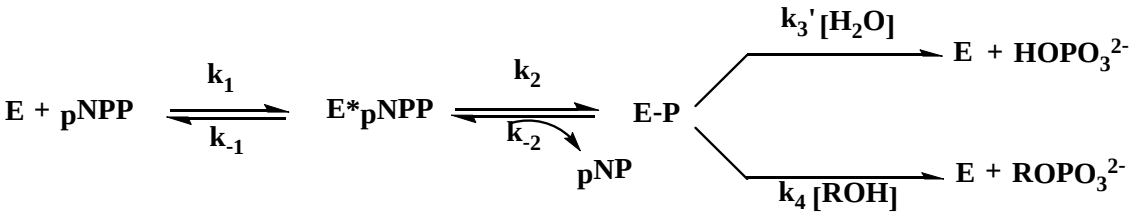
adenosine deaminase and monitoring the absorbance decrease at 265nm caused by the conversion of adenosine to inosine.

The results of studies of product inhibition of the AphA promoted AMP hydrolysis reaction showed that adenosine acts as a strong competitive inhibitor with a K_i = 7.8 ± 0.5 μM and that inorganic phosphate acts as a weak noncompetitive inhibitor with a K_i = 15.3 ± 1.4 mM (Figure 2.11). In contrast, inorganic phosphate is a weak noncompetitive inhibitor (Ki = 9.0 ± 0.9 mM) (Figure 2.12) and para-nitrophenol has no obvious effects on AphA catalyzed hydrolysis reaction of pNPP (Table 2.6).

AphA-catalyzed hydrolysis of monophosphoesters is a one-substrate/two products (ROH and PO_4^{3-}) type reaction. The inhibition results suggest that in the hydrolysis reaction of AMP the adenosine product is released first followed by release of inorganic phosphate, while in the AphA catalyzed pNPP reaction release of the two products does not follow a defined sequence. However, due to the high Ki of PO_4^{3-} to AphA hydrolyzed AMP reaction and the similar structures of adenosine and AMP, the products inhibition test may not reflect the real products releasing sequence.

2.3.9 Alcohol Effect on AphA-Catalyzed Hydrolysis of pNPP

As mentioned in Section 2.3.6, *E coli.* AphA serves as a phosphotransferase. Therefore, when another nucleophile in addition to water is present in the system, the possibility exists for formation of another monophosphoester by a phosphoryl transfer process (Scheme 2.1). To probe this expectation, a series of alcohols were included in reaction mixture used for AphA catalyzed hydrolysis of pNPP. The rates of these processes were monitored by measuring product pNP formation and the product distributions were determined by using mass spectrometric analysis.



Scheme 2.1 Steps in the AphA catalyzed hydrolysis reaction of pNPP in the presence of an alcohol ROH.

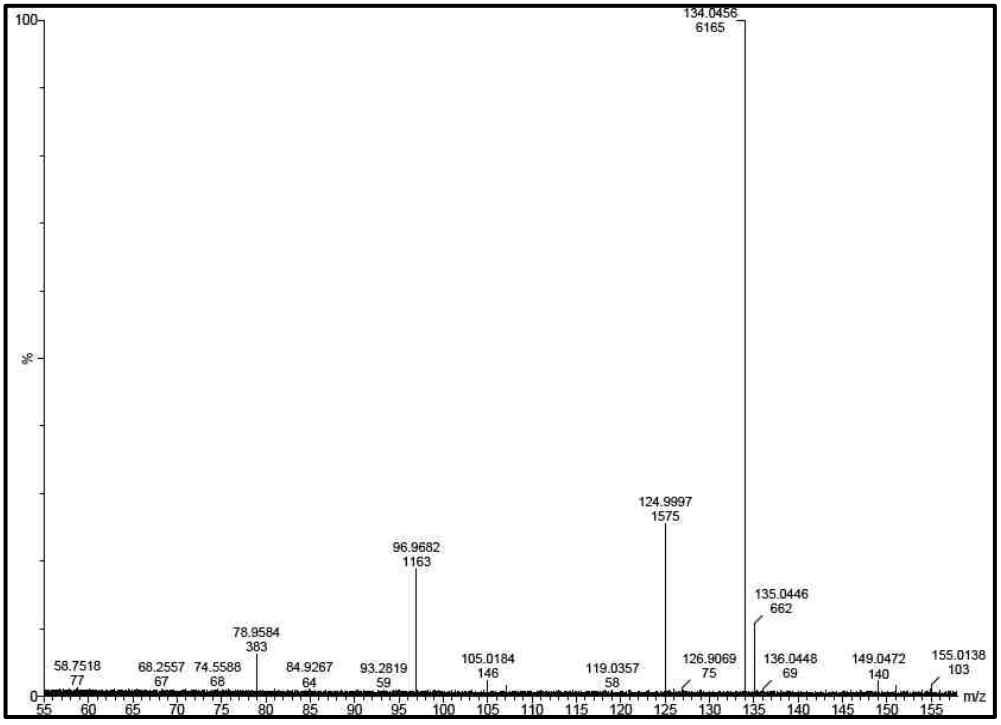


Figure 2.13 Mass spectrum of products from the AphA-catalyzed pNPP reaction in the presence of ethanol.

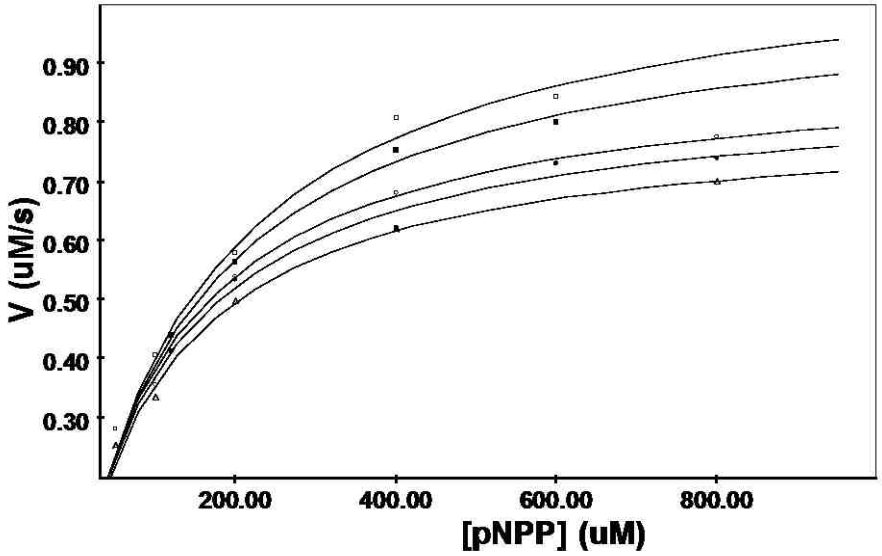


Figure 2.14 Michaelis-Menten plots of AphA hydrolysis of pNPP in the presence of ethylene glycol in assay buffer, 25 °C. The concentrations of ethylene glycol from bottom to top are 0, 0.1, 0.2, 0.3 and 0.4 M, respectively.

The production of ethyl phosphate in the AphA catalyzed reaction of pNPP carried out in the presence of ethanol was demonstrated by using mass spectrometric analysis of the product mixture (Figure 2.13). Specifically, the peak at 124.99 m/z seen in the spectrum is associated with ethyl phosphate while the peak at 96.97 m/z belongs to PO_4^{3-}.

The observation that the presence of ethylene glycol and ethanol in the mixture used for AphA catalyzed hydrolysis of pNPP greatly alters the Machaelis-Menten kinetic parameters (Table 2.7 and Figure 2.14 and 2.15) provides mechanistic information about the process. Because these alcohols are neither substrates nor products of the AphA-catalyzed pNPP reaction, their presence can only alter the chemical steps of the process. The two chemical steps involved in the overall reaction are phosphorylation of the enzyme forming an acyl-phosphate intermediate and nucleophile promoted

dephosphorylation of the phosphoenzyme intermediate. If the formation of phosphoenzyme intermediate is rate limiting, the addition of alcohol acceptors that only

Table 2.7 Kinetic parameters of AphA hydrolysis of pNPP in the presence of ethylene glycol.

Ethylene glycol (M)	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
0	130 ± 13	16.7 ± 0.6	1.28×10^5
0.1	131 ± 10	17.6 ± 0.6	1.33×10^5
0.2	141 ± 13	18.6 ± 0.6	1.32×10^5
0.3	162 ± 15	21.1 ± 0.7	1.30×10^5
0.4	180 ± 8	23.1 ± 0.4	1.29×10^5

react with the intermediate after the rate-limiting step will not alter the overall reaction rate of disappearance of the starting phosphate ester and formation of the *p*-nitrophenol product. However, if dephosphorylation of the phospho-enzyme (E-P) intermediate is the rate-limiting step and if the added alcohol is a better nucleophile than water, then the presence of ethylene glycol should increase the rate of the overall reaction rate (*i. e.,* the rate of PNP product formation). This was observed to occur in the AphA catalyzed hydrolysis reaction of pNPP as shown by the data given in Figure 2.14 and 2.15, and Table 2.7. Specifically, the rate of formation of pNP increased dramatically in the presence of this diol. Furthermore, the increase of the rate was proportional to the concentration of the added ethylene glycol.

Thus, the presence of a substance that is a better nucleophile than water increases the rate of breakdown rate of the E-P intermediate. This observation suggests that the dephosphorylation of the intermediate is rate-limiting in the AphA catalyzed hydrolysis reaction of PNPP. Finally, even though the addition of alcohol leads to a significant increase of the rate of the catalyzed reaction, it does not influence the enzyme catalytic efficiency (the k_{cat}/K_m value remains at 1.30 x 10^5 $M^{-1}s^{-1}$).

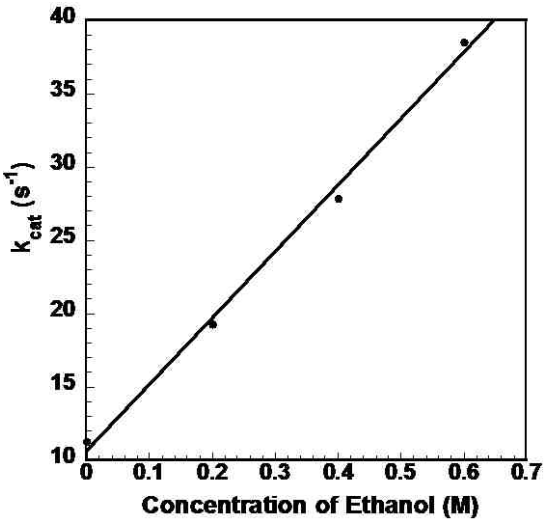


Figure 2.15 Effects of ethanol concentration on k_{cat} of AphA hydrolyzed pNPP in assay buffer, at 25°C.

Because the rate-limiting step of the AphA catalytic reaction involves dephosphorylation of the E-P intermediate, the overall rate of substrate turnover, k_{cat}, is determined by the decomposion of E-P according to equation 2.11, derived from the pathway shown in Scheme 2.1 above. In this eqation

$$k_{cat} = k_3' [H_2O] + k_4 [ROH] \qquad \text{(eq. 2.11)}$$

$k_3' [H_2O]$ is the rate of hydrolysis E-P and $k_4 [ROH]$ is the rate of E-P alcoholysis.

The fact that the alcoholysis/hydrolysis of the E-P intermediate is the rate-limiting step a determination of how the nucleophilicity of the added alcohol effects rates can help provide information about the nature of the transition state for the E-P breakdown step. The general method typically used for this purpose is termed a linear free energy relationship (LFER) study in which reactions rates are correlated to a parameter that reflects the properties of the substituent or reactant being varied[18]. In particular, a Brϕnsted plot of the log of the rate of the AphA catalyzed reaction of PNPP in the presence of various alcohols as a function of the pKa of the alcohols (reflecting their nucleophilicity) will give a Brϕnsted coefficient, $\beta_{nu:}$ (obtained from the slope of the line) that reflects the degree of charge build-up on the hydroxyl oxygen and, as a result, degree of bond formation bond formation between the alcohol nucleophile and the phosphate phosphorus of the E-P intermediate. Although often used for enzymatic reactions [19-22], LFER results are sometimes difficult to interpret because the changes being made could influence other facets of the reaction being observed. Based on knowing that care must be taken in the choice of alcohols used in this study[23], eight β-substituted primary ethyl alcohols, covering a pKa range of 12.0-16.0, were chosen to minimize the influence of steric effects. The rate constants of alcoholysis of the phosphoenzyme intermediate obtained with these alcohols are listed in Table 2.8 and a Brϕnsted plot of the data is given in (Figure 2.16). Analysis of the plot gives a slope of $\beta_{nu:} = 0.14 \pm 0.05$ ($r = 0.76$), which shows that nucleophilicity of the alcohol only slightly influences reactivity. The small dependence of the alcoholysis second order rate constant k_4 on the basicity of the ethyl alcohols suggests that reaction of the E-P intermediate in the AphA-catalyzed reaction is highly dissociative.

Table 2.8 Second-order rate constants for AphA-hydrolyzed pNPP in presence of a series of β-substituted ethanol.

β-substituted ethanol	pK_a	k_4 ($M^{-1}s^{-1}$)
CF_3CH_2OH	12.37	28.32 ± 2.7
$HC\equiv CCH_2OH$	13.55	32.75 ± 3.0
$N\equiv CCH_2CH_2OH$	14.03	19.50 ± 1.3
$ClCH_2CH_2OH$	14.31	38.44 ± 3.3
$CH_3OCH_2CH_2OH$	14.82	43.30 ± 4.5
$CH_2=CHCH_2OH$	15.52	62.75 ± 3.0
CH_3CH_2OH	15.90	46.50 ± 2.4
$CH_3CH_2CH_2OH$	16.10	123.50 ± 1.4

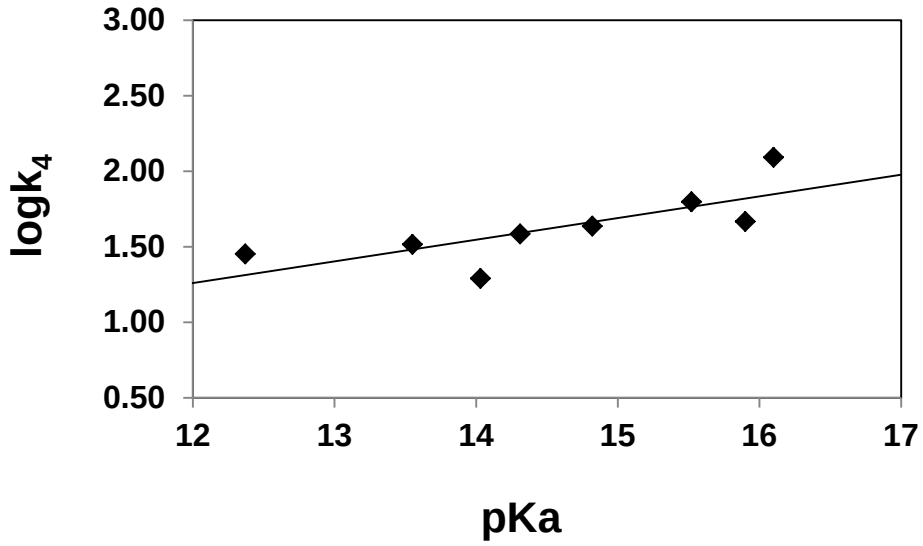


Figure 2.16 Bronsted plot of the second order rate constants for the alcoholysis of E-P. reaction is highly dissociative in nature with a large degree of P-O bond cleavage and a small degree of P-O bond formation taking place in the transition state.

2.3.10 Binding Constant of AMP to Wild Type AphA and Y218W Mutant Measured by Using a Stopped-Flow Fluorescence Technique

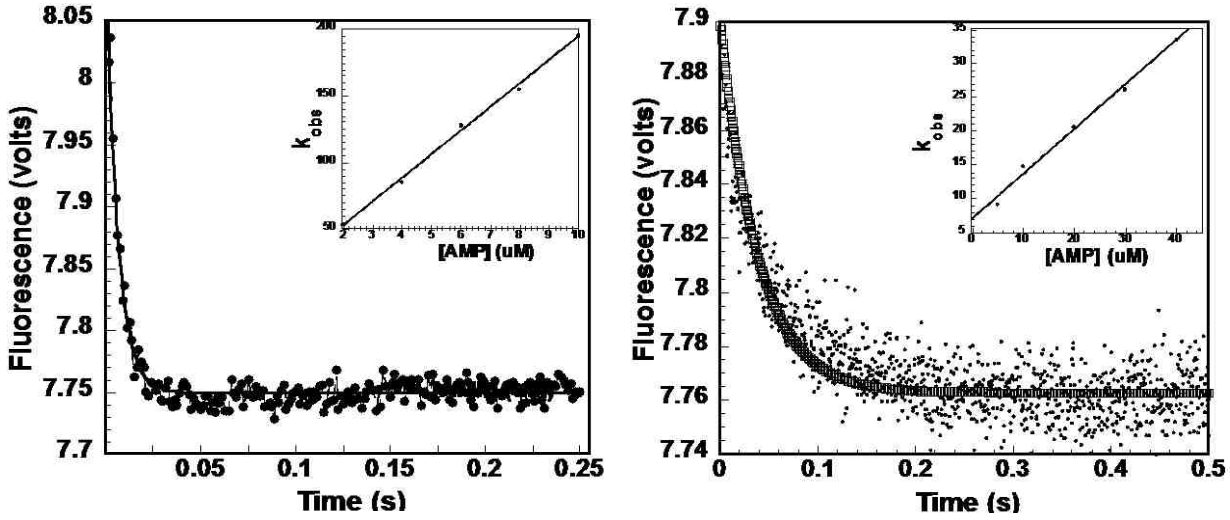


Figure 2.17. Time-dependent fluorescence changes associated with the binding of **(A)** 10 μM AMP and 1 μM Apha wild type in 50 mM Na^{+}MES (pH 6.0, 25 ^{o}C) and **(B)** 5 μM Y218W and 40 μM AMP. Insets: plots of the k_{obs} vs AMP concentration (A) from 2 - 10 μM, (B) from 5 – 40 μM.

Five tryptophan residues are present in AphA and, as shown by the crystal structure, only one (Tyr77) is located near the active site [(3)]. As such, this residue can be used to probe changes taking place, such as substrate binding, in active site cavity. Accordingly, the fluorescence emission from Tyr77 was monitored during the course of substrate binding, catalysis and product release by using the stopped-flow method. Time-dependent fluorescence quenching of the tryptophan residue in wild-type AphA (1 μM) by AMP (10

μM) binding was monitored under pseudo-first-order conditions, defined by the AMP to enzyme concentration ratio of 10:1 as shown in Figure 2.17 (A). The stopped-flow time course shows that a continuous decrease in fluorescence intensity occurs, implying that a single step involved in the substrate binding (E + S = ES). The observed first-order rate constant, k_{obs}, for binding was obtained by fitting the data to equation 2.9, which describes a single-exponential decay, to give a rate constant of 195 ± 8 s^{-1}. Various ligand concentrations (2, 4, 6, 8, 10 μM) were then used to measure the association rate constant (k_{on} = 18 ± 0.5 $\mu M^{-1}s^{-1}$) and dissociation rate constant (k_{off} = 17 ± 4 s^{-1}) of AMP binding to wild-type AphA. For comparison, the AMP binding constant to mutant Y218W was also measured. The time course of fluorescence quenching for the binding of 40 μM AMP to 5 μM Y218W, shown in Figure 2.17 (B), gave a second order rate constant k_{obs} = 33.5 ± 2.2 s^{-1}. The variation in the k_{obs} value with AMP concentration gave the plot shown as an inset to Figure 2.17 (B). The association (k_{on}) and dissociation (k_{off}) rate constants, determined by fitting the data to equation 2.2, were found to be 0.66 ± 0.04 $\mu M^{-1}s^{-1}$ and k_{off} = 7 ± 0.9 s^{-1}, respectively.

2.3.11 Single-Turnover Kinetic Analysis of Wild Type AphA Hydrolysis of AMP

The time courses for consumption of [^{14}C]-AMP and formation of [^{14}C]-adenosine in single turnover reactions, carried out with excess protein (37.5 μM) and limiting substrate (10 μM), are shown in Figure 2.18. The first order rate constants k_{obs}, which were obtained by fitting the data to equations 2.4 and 2.5 are 4.6 ± 0.66 s^{-1} for [^{14}C] AMP consumption, and 4.7 ± 0.41 s^{-1} for [^{14}C] adenosine formation.

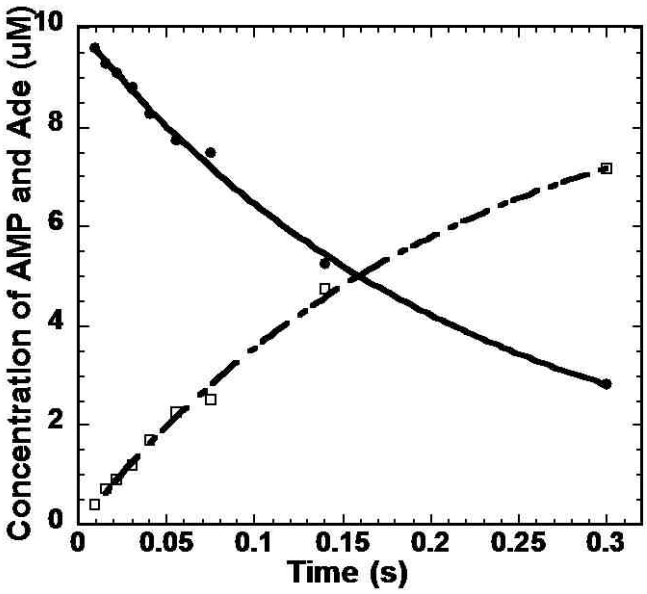


Figure 2.18 The time course for single turnover reaction of 37.5 μM wild-type AphA with 10 μM [^{14}C]-AMP in 50 mM Na^+MES containing 5 mM $MgCl_2$, pH 6 (•) AMP; (□) adenosine. The curves show fits to the first order rate equation 2.6.

2.3.12 Kinetic Analysis of Wild Type AphA Catalyzed Reaction of [^{14}C/U]-AMP under Multi-Turnover Conditions by Using the Stopped Flow Method

A multi-turnover AphA catalyzed reaction of AMP was run in 50 mM Na^+MES buffer containing 5 mM $MgCl_2$, pH 6.0, 25 °C with 20 μM enzyme and excess substrate (100 μM). At fixed time intervals (3-1500 ms), the reactions were quenched, and the amount of adenosine was determined. A plot of adenosine formation versus time (Figure 2.19), displays a pre-steady-state burst of product formation with a fast initial rate followed by a

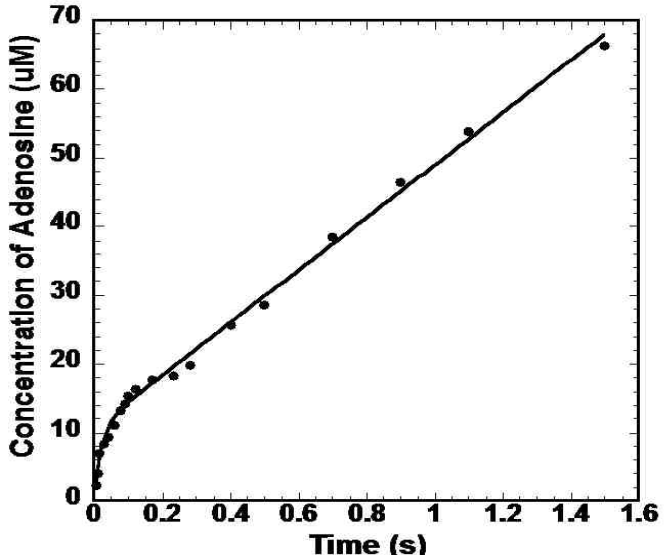


Figure 2.19 The time course of the multi-turnover reaction of 20 μM AphA with 100 μM AMP in 50 mM Na^+MES buffer containing 5mM $MgCl_2$, pH 6.0, and at 25°C. (•) experimental data; the curve shows fit to burst equation.

slower steady-state rate of product formation. The phenomenon of burst kinetic indicates that the substrate-binding step is not the rate-limiting and that chemical reactions or product release may be the rate-limiting step. This finding is consistent with the observation that the rate-limiting step is dephosphorylation of EP intermediate.

Fitting the rapid-quench data to the burst equation 2.6 yielded a burst rate constant of 45.1 ± 8.0 s^{-1} and the steady-state k_{cat} of 2.0 ± 0.03 s^{-1}. The k_{cat} value was obtained by dividing observed rate constant by the enzyme concentration. The k_{cat} value obtained in this manner is in good agreement with the steady-state k_{cat} value of 2.6 ± 0.06 s^{-1}.

2.3.13 Kinetic Analysis of Wild Type AphA-Catalyzed pNPP under Multi-Turnover Conditions by Using Stopped-Flow Method

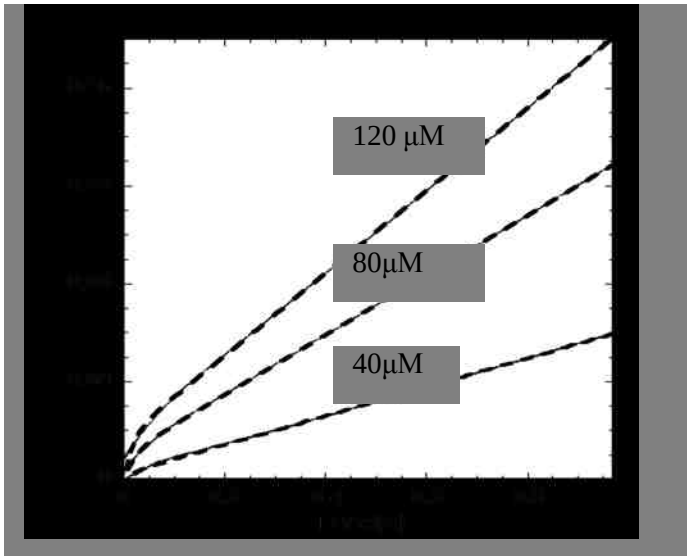


Figure 2.20 The time course for multi-turnover reaction of 500 μM pNPP catalyzed by wild-type AphA (40, 80 and 120 μM) in 50mM Na^{+}MES containing 5mM $MgCl_2$, pH 6.

A pre-steady state kinetic analysis of wild-type AphA catalyzed pNPP hydrolysis was carried out under multiple-turnover conditions with substrate in excess (500 μM) to ensure the saturation of the enzyme (40 - 120 μM). As is evident from the data obtained for the process (Figure 2.20), a burst of product production occurs, marked by the increase in absorption at 410 nm, followed by a slow linear phase. The amplitudes of the bursts were dependent on enzyme concentration. The rate constant for each different enzyme concentration was obtained (Table 2.9) by fitting the stopped-flow data to equation 2.7.

Table 2.9 The rate constants of wild type AphA catalyzed pNPP under multi-turnover reaction conditions.

[AphA] (μM)	k_{obs} (s^{-1})	k_{cat} (s^{-1})
40	60.0 ± 1.8	7.9 ± 0.0023
80	43.2 ± 0.5	8.2 ± 0.0022
120	56.9 ± 1.0	7.6 ± 0.003
Average	53.4	7.9

2.3.14 Time Course of Steady-State Reaction under Multi-Turnover Conditions

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P$$

Scheme 1 Kinetic model representing AphA-catalyzed AMP reaction under steady-state multi-turnover condition.

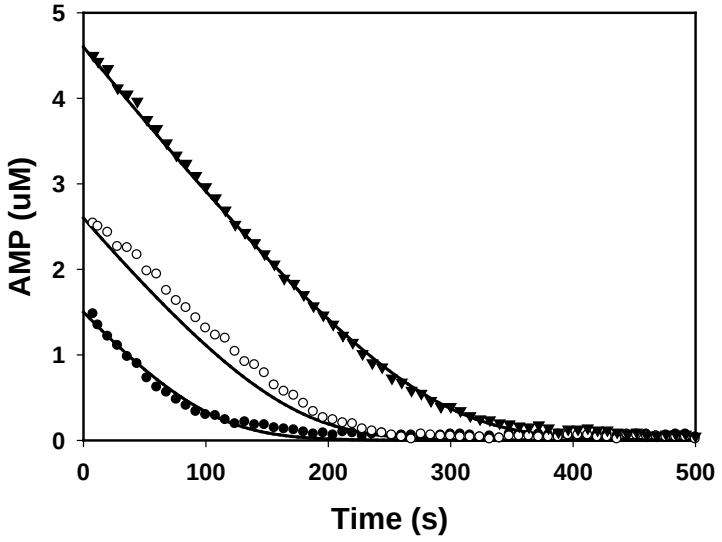


Figure 2.21 Time courses for the steady-state multiple-turnover reactions of 1.5 (●), 2.6 (○), or 4.6 (▼) μM AMP catalyzed by 0.006 μM wild type AphA in 50 mM Na^+MES (pH 6.0, 25 °C). The traces of absorbance decrease at 265 nm are shown in the context of the calculated concentrations of AMP ($\Delta\varepsilon$= 8.4 mM^{-1} cm^{-1}) present in the reaction mixture (y-axis) as a function of reaction time (x-axis). The simulated curves (solid lines) were generated using the kinetic model shown in Scheme 1.

Steady-state multi-turnover reaction time courses were measured by mixing 0.006 μM AphA with 1.5, 2.6 and 4.6 μM AMP (Figure 2.21). The reaction process was analyzed by fitting the data to the simple model shown in Scheme 1 using KinTek Explorer. This treatment gives k_1 = 5.95 $\mu M^{-1}s^{-1}$, k_2 = 3.36 s^{-1} and k_{-1} =1.19 s^{-1}. From these results k_{cat} ($k_{cat} = k_2$) and K_m ($K_m = (k_{-1}+ k_2)/k_1$) were calculated to be 3.36 s^{-1} and 0.76 μM, respectively.

2.4 Conclusions

$$\mathbf{E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EI \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} EP \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} E + P}$$

Scheme 2 Kinetic model for the AphA catalyzed reaction of AMP under steady state multi-turnover conditions. (Abbreviations: E, wild type AphA; S, substrate AMP; Q, inorganic phosphate; P, adenosine.)

In the studies described above, a series of kinetic and mechanistic experiments were performed in an attempt to uncover the detailed mechanism for catalysis of AMP hydrolysis by AphA. Previous structural studies of complexes of AphA-products and AphA-osmate[4] revealed that two different binding sites P1 and P2 exist for the two different hydrolysis products, phosphate and adenosine, and that the existence of E-P intermediate was possible. A later study [9] proved the P2 adenosine binding site was also the AMP substrate binding pocket.

The results obtained from the product inhibition studies described above show that, in the absence of phosphate, adenosine acts as a strong competitive inhibitor with Ki = 7.8 ± 0.5 μM. This finding is consistent with the fact that AMP and adenosine share the same hydrophobic binding cleft. The ability of AphA to catalyze phosphate transfer was obseved and is regarded as further evidence that the phosphatase-catalyzed hydrolysis reaction involves a phosphoenzyme intermediate (E-P). The process involving phosphoryl transfer to alcohols was used to identify the rate-limiting step of the process[24]. The observation that addition of alcohol leads to a significant increase in the overall reaction rate and that the increase is dependent on the concentration of alcohol, shows that E-P dephosphorylation of E-P intermediate is the rate-limiting step. Furthermore, the presence of a solvent deuterium isotope effect (k_{cat}^{H} /k_{cat}^{D} = 4.78) indicates that proton transfer is involved in the rate-limiting step of the catalytic process and the results of the proton inventory study demonstrate that only one proton is in flight during the catalysis reaction. Taken together, these results are consistent with the operation of a mechanism for AphA catalyzed phosphate ester hydrolysis that is the same as those generally followed by all HADSF phosphatases[25]. Specifically, the reaction

proceeds by intial displacement by an active site aspartate nucleophile of the substrate phosphate group and that a second Asp residue functions as a general acid/base to protonate the leaving group in the phosphorylation step and deprotonate the water/alcohol nucleophile in the dephosphorylation step.

AphA catalyzed dephosphorylation reactions that proceed through formation of a stable intermediate can follow two limiting pathways[(26)], involving a dissociative or an associative mechanism. Linear free energy relationship studies, which correlated the pKa of nucleophilic alcohol phosphoryl acceptors and first order rate constants k_4, show that the process of dephosphorylation of the E-P intermediate has a small $\beta_{nu:}$ value (0.14 ± 0.05). This finding indicates that the basicity of phosphate acceptor does not greatly influent the rate of E-P dephosphorylation and that, therefore, the process is highly dissociative in nature.

Time courses of transient state, multi-turnover, single-turnover reactions as well as substrate binding were determined to define the microscopic rate constants for the steps involved in the AphA catalytic pathway. Based on the results of these efforts, a kinetic model was generated (Scheme 2) to explain the process. The overall reaction of AphA hydrolyzed AMP involves one substrate-binding step, two chemical steps and one product-releasing step. The simulation-based fitting curves of the four sets of time course data were shown in Figure 2.22. 2.23 and 2.24. k_1 and k_{-1} from binding experiment are obtained first to initiate the simulation process (Figure 2.22).

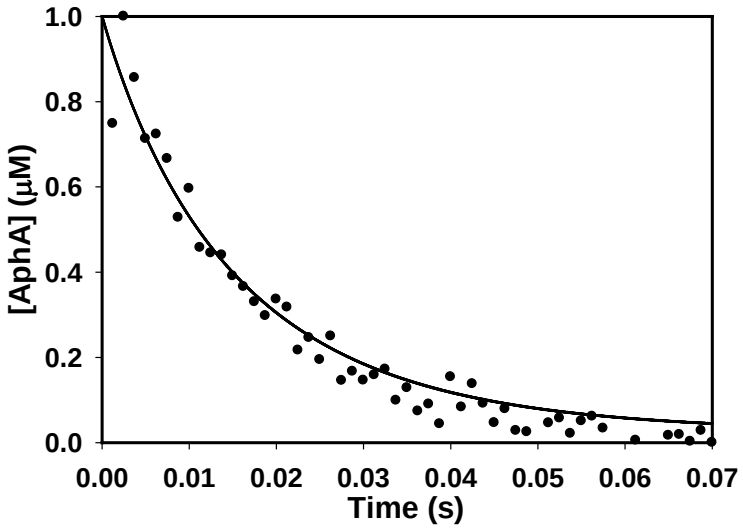
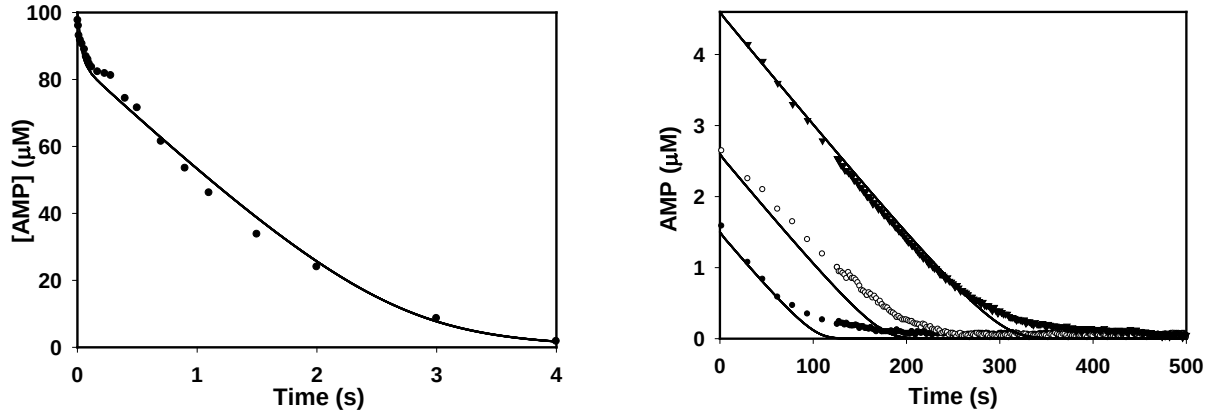


Figure 2.22 A plot of the concentration of unbound AphA vs time for the time course for the binding reaction between 4.0 μM AMP and 1.0 μM AphA in 50 mM Na^+MES (pH 6.0, 25 °C). The simulated curve was generated using the kinetic model shown in Scheme 2 and rate constants k_1 (19.9 μM^{-1} s^{-1}) and k_{-1} (25.8 s^{-1}).



(A)

(B)

Figure 2.23 Plot of experimental and simulated time courses for multiple-turnover reactions of (A) the transient-state (100.0 μM AMP and 20.0 μM AphA in 50 mM

Na^+MES, pH 6.0, 25 °C and (B) steady-state 0.06 μM AphA and 1.5 (●), 2.6 (◯), or 4.6 (▼) μM AMP in 50 mM Na+MES, pH 6.0, 25 °C. The simulated curve was generated using the kinetic model shown in Scheme 2 and rate constants k_1 = 17.0 μM^{-1} s^{-1}, k_{-1} = 18.0 s^{-1}, k_2 = 27.7 s^{-1}, k_{-2} =18.1 s^{-1}, k_3 =3.8 s^{-1}, k_{-3} = 0 s^{-1}, k_4 = 165.5 s^{-1}, k_{-4} = 59.6 s^{-1}.

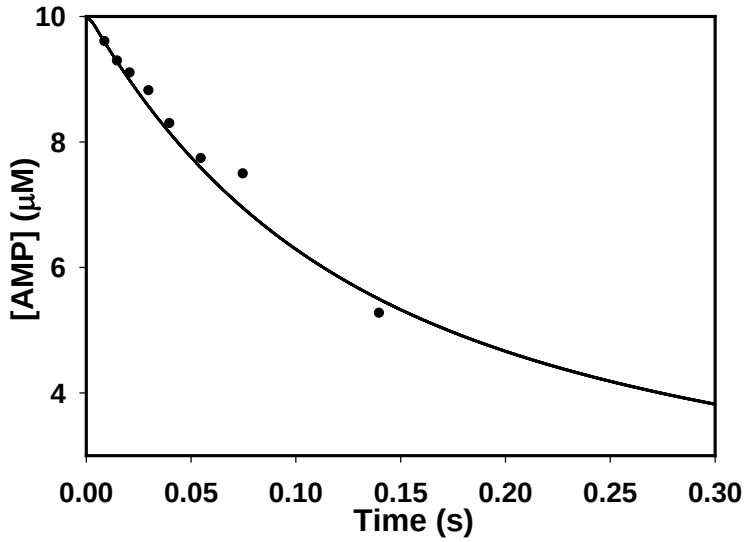


Figure 2.23 Plot of experimental and simulated time courses for the transient-state single-turnover reactions of 10.0 μM AMP and 37.5 μM AphA in 50 mM Na^+MES, pH 6.0, 25 °C. The simulated curve was generated using the kinetic model shown in Scheme 2 and rate constants k_1 = 18.4 μM^{-1} s^{-1}, k_{-1} = 17.3 s^{-1}, k_2 = 7.2 s^{-1}, k_{-2} =6.1 s^{-1}, k_3 =3.1 s^{-1}, k_{-3} = 0 s^{-1}, k_4 = 175.0 s^{-1}, k_{-4} = 53.6 s^{-1}.

Inspection of the data shows that the chemical step governs the rate of the reaction with a rate constant k_3 = 3.1 ~ 3.8 s^{-1}. This value is close to the experimental k_{cat} (2.6 s^{-1})

determined by using steady-state experiments. The difference between the two rate constants is likely caused by the imprecise assumption that product inhibition is negligible, which is typically used to estimate initial velocities in conventional steady-state analysis. This result is consistent with the rate-limiting step is the dephosphorylation of E-P intermediate.

References

1. Kier, L. D., Weppelman, R., and Ames, B. N. (1977) Resolution and purification of three periplasmic phosphatases of Salmonella typhimurium, *Journal of bacteriology 130*, 399-410.
2. Rossolini, G. M., Thaller, M. C., Pezzi, R., and Satta, G. (1994) Identification of an Escherichia-Coli Periplasmic Acid-Phosphatase Containing of a 27 Kda-Polypeptide Component, *Fems Microbiol Lett 118*, 167-173.
3. Calderone, V., Forleo, C., Benvenuti, M., Thaller, M. C., Rossolini, G. M., and Mangani, S. (2004) The first structure of a bacterial class B acid phosphatase reveals further structural heterogeneity among phosphatases of the haloacid dehalogenase fold, *J Mol Biol 335*, 761-773.
4. Calderone, V., Forleo, C., Benvenuti, M., Thaller, M. C., Rossolini, G. M., and Mangani, S. (2006) A structure-based proposal for the catalytic mechanism of the bacterial acid phosphatase AphA belonging to the DDDD superfamily of phosphohydrolases, *J Mol Biol 355*, 708-721.
5. Uerkvitz, W., and Beck, C. F. (1981) Periplasmic phosphatases in Salmonella typhimurium LT2. A biochemical, physiological, and partial genetic analysis of three nucleoside monophosphate dephosphorylating enzymes, *The Journal of biological chemistry 256*, 382-389.
6. Grose, J. H., Bergthorsson, U., Xu, Y. P., Sterneckert, J., Khodaverdian, B., and Roth, J. R. (2005) Assimilation of nicotinamide mononucleotide requires periplasmic AphA phosphatase in Salmonella enterica, *Journal of bacteriology 187*, 4521-4530.

7. Reshetnyak, E., d'Alencon, E., Kern, R., Taghbalout, A., Guillaud, P., and Kohiyama, M. (1999) Hemi-methylated oriC DMA binding activity found in non-specific acid phosphatase, *Mol Microbiol 31*, 167-175.

8. Koenigsknecht, M. J., Ramos, I., and Downs, D. M. (2007) Glutamine phosphoribosylpyrophosphate amidotransferase-independent phosphoribosyl amine synthesis from ribose 5-phosphate and glutamine or asparagine, *The Journal of biological chemistry 282*, 28379-28384.

9. Leone, R., Cappelletti, E., Benvenuti, M., Lentini, G., Thaller, M. C., and Mangani, S. (2008) Structural Insights into the Catalytic Mechanism of the Bacterial Class B Phosphatase AphA Belonging to the DDDD Superfamily of Phosphohydrolases, *J Mol Biol 384*, 478-488.

10. C. A. Reddy, T. J. B., John A. Breznak, George Marzluf, Thomas M. Schmidt, Loren R. Snyder (1994) *Methods for General and Molecular Bacteriology*, American Society for Microbiology Washington, DC.

11. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Analytical biochemistry 72*, 248-254.

12. Ames, B. N. (1966) Assay of inorganic phosphate, total phosphate and phosphatases, *Methods Enzymol. 8*, 115-118.

13. Johnson, K. (2003) *Kinetic analysis of macromolecules : a practical approach*, Oxford University Press, Oxford.

14. Schowen, K. B., and Schowen, R. L. (1982) Solvent Isotope Effects on Enzyme-Systems, *Method Enzymol 87*, 551-606.

15. Zhao, Y., and Zhang, Z. Y. (1996) Reactivity of alcohols toward the phosphoenzyme intermediate in the protein-tyrosine phosphatase-catalyzed reaction: Probing the transition state of the dephosphorylation step, *Biochemistry-Us 35*, 11797-11804.

16. Forleo, C., Benvenuti, M., Calderone, V., Schippa, S., Docquier, J. D., Thaller, M. C., Rossolini, G. M., and Mangani, S. (2003) Expression, purification, crystallization and preliminary X-ray characterization of the class B acid phosphatase (AphA) from Escherichia coli, *Acta Crystallogr D 59*, 1058-1060.

17. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D., and Aravind, L. (2006) Evolutionary genomics of the HAD superfamily: Understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes, *J Mol Biol 361*, 1003-1034.

18. Williams, A. (1992) Effective Charge and Transition-State Structure in Solution, *Adv Phys Org Chem 27*, 1-55.

19. Williams, N. H., Cheung, W., and Chin, J. (1998) Reactivity of phosphate diesters doubly coordinated to a dinuclear cobalt(III) complex: Dependence of the reactivity on the basicity of the leaving group, *J Am Chem Soc 120*, 8079-8087.

20. Liao, X., Anjaneyulu, P. S., Curley, J. F., Hsu, M., Boehringer, M., Caruthers, M. H., and Piccirilli, J. A. (2001) The tetrahymena ribozyme cleaves a 5'-methylene phosphonate monoester approximately 10(2)-fold faster than a normal phosphate diester: implications for enzyme catalysis of phosphoryl transfer reactions, *Biochemistry-Us 40*, 10911-10926.

21. Kirby, A. J., and Jencks, W. P. (1965) Reactivity of Nucleophilic Reagents toward P-Nitrophenyl Phosphate Dianion, *Journal of the American Chemical Society 87*, 3209-&.

22. Thatcher, G. R. J., and Kluger, R. (1989) Mechanism and Catalysis of Nucleophilic-Substitution in Phosphate-Esters, *Advances in Physical Organic Chemistry 25*, 99-265.

23. Hollfelder, F., and Herschlag, D. (1995) The nature of the transition state for enzyme-catalyzed phosphoryl transfer. Hydrolysis of O-aryl phosphorothioates by alkaline phosphatase, *Biochemistry-Us 34*, 12255-12264.

24. Zhang, Z. Y., and Vanetten, R. L. (1991) Pre-Steady-State and Steady-State Kinetic-Analysis of the Low-Molecular-Weight Phosphotyrosyl Protein Phosphatase from Bovine Heart, *Journal of Biological Chemistry 266*, 1516-1525.

25. Allen, K. N., and Dunaway-Mariano, D. (2009) Markers of fitness in a successful enzyme superfamily, *Curr Opin Struc Biol 19*, 658-665.

26. Allen, K. N., and Dunaway-Mariano, D. (2004) Phosphoryl group transfer: evolution of a catalytic scaffold, *Trends Biochem Sci 29*, 495-503.

CHAPTER THREE

3. Preliminary Mechanistic Studies of Human Mitochondrial Deoxyribonucleotidase

3.1 Introduction

Cells need four deoxyribonucleotide triphosphates (dNTPs) to replicate and repair DNA. Some specific diseases such as severe immune deficiency[(1)], apoptotic destruction of B and/or T cells[(2)] and neurogastrointestinal encephalomyopathy[(3)], are caused by imbalances in the dNTP pool. Therefore, regulation of the supply of these four deoxyribonucleotide triphosphates is required in order to control size of the dNTP pool. Several enzymes are involved in the synthesis of deoxyribonucleotides. For example, dATP, dCTP and dGTP are directly formed by reduction of ribonucleotides, followed by phosphorylation of the dADP, dCDP and dGDP intermediates. dTTP is also formed by a reductase promoted process, but needs more extensive remodeling[(4)]. Along with the main de novo DNA synthesis pathway, two side routes exist to regulate the material entering and leaving dNTP pool (Figure 3.1). When the supply of deoxyribonucleotides exceeds the requirements for the DNA replication, deoxyrobinucleotides are converted to deoxyribonucleosides by 5'-nucleotidases that leave the pool. While when dNTPs are in short supply, kinases phosphorylate deoxyribonucleosides to increase the DNA pool size. In this regulation loop, dNTP5'-nucleotidases serve as types of 5'-nucleotidases that dephosphorylate deoxynucleoside monophosphates to produce deoxynucleoside and inorganic phosphate. Seven mammalian 5'-nucleotidases with different subcellular locations are known to exist. The effort described below focuses on the mitochondrial

deoxyribonucleotidase, mdN (previously named dNT2)[5], the only 5'-nucleotidase found in mitochondria.

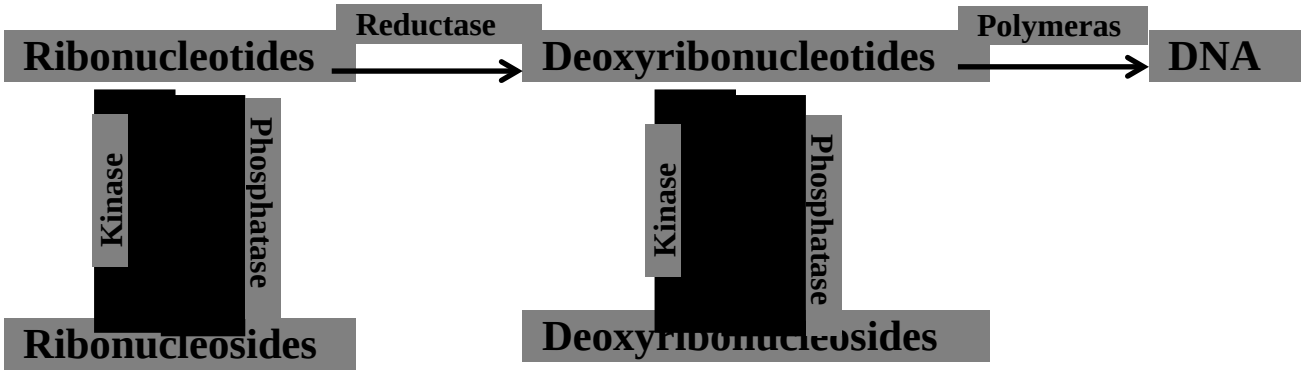
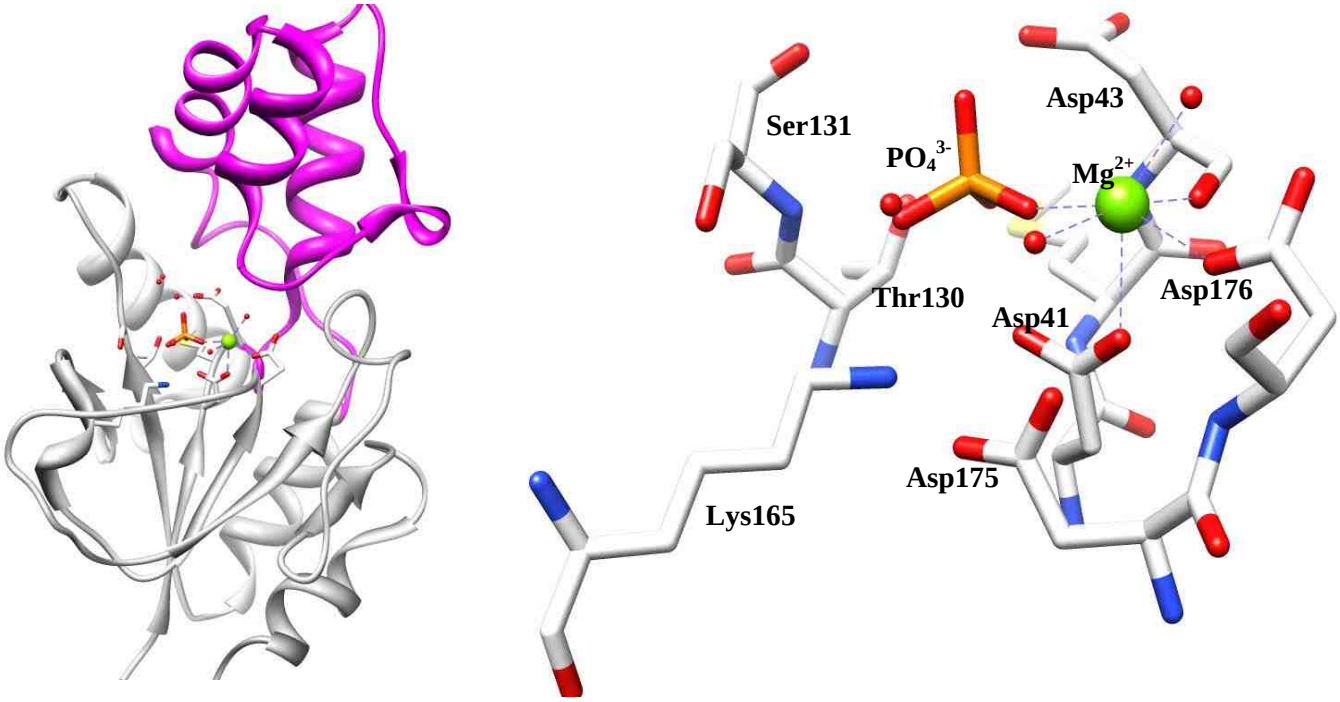


Figure 3.1 *de novo* synthesis of DNA.

3.1.1 Crystal Structure of Mitochondrial Deoxyribonucleotidase (mdN)

In 2002, the crystal structure of the first mammalian nucleotidase, mdN complexed with phosphate was solved [6]. The crystal structure, a dimer formed by 193 amino acids (amino acids 34 to 227), is similar to those of several proteins belonging to the HADSF. Sequence alignment analysis shows that mdN has the four highly conserved signature loops common to members of the HADSF and that each monomer of mdN is composed of two domains, a large core domain and a small cap domain (Figure 3.2 A). The core domain forms α/β Rossmann-like fold that contains all of the catalytic residues. The active site, located in a cleft between the two domains, is solvent accessible, and contains a Mg^{2+} ion and an inorganic phosphate moiety (Figure 3.2 B). Mg^{2+} ion is coordinated in an octahedral geometry by two water molecules, one phosphate group, side chains of Asp 41 and Asp 146 and backbone of Asp 43. Asp41 and Asp43 constitute motif 1, Try130

and Ser131 constitute motif 2, Arg163 and Lys165 constitute motif 3, while Asp175 and Asp176 constitute motif 4 of HADSF [7].



(A) (B)

Figure 3.2 (A) The structure of mdN phosphate complex. The cap domain is in magenta and the core domain is in grey. **(B)** Active site of mdN. Magnesium is shown in green sphere.

3.1.2 General Catalytic Mechanism of mdN

In 2002, a general catalytic mechanism of mdN catalyzed hydrolysis of nucleotide monophosphate was proposed by Nordlund and his coworkers [6], based on the crystal structure of mdN in complex with bound phosphate, thymidine and the intermediate mimic, beryllium trifluoride. Like all other HADSF members[8], the reaction was

suggested to proceed via a “in line” nucleophilic substitution pathway involving two nucleophiles. Juxtaposed to the phosphate moiety through bridging by the Mg^{2+} ion and Lys 165, Asp41 serves as the first nucleophile (Figure 3.2 B). Binding of Mg^{2+} ion to the

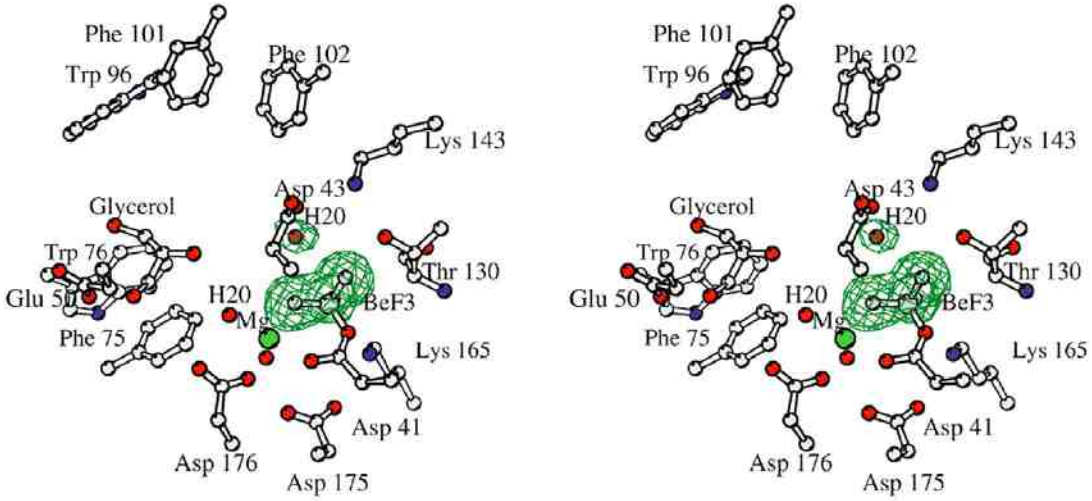


Figure 3.3 Stereo view of the active site of mdN in complex with BeF_3. *from ref. 6.*

nucleophile and the phosphorylated nucleotide substrate provides the required orientation and charge shielding for nucleophilic attack by ASP41[(9)]. The crystal structure of the BeF_3 complex of mdN provides evidence that suggest the formation of an aspartyl-phosphate intermediate in the reaction process (Figure 3.3). In the structure of the complex the conserved Lys165 side chain is observed to be in close proximity to the BeF_3 group and to the nucleophile Asp41 side chain. This finding indicates that Lys165 might play a key role in transition state stabilization. The negatively charged side chain of Lys165 residue may be to position the Asp nucleophile and as well to coordinate the juxtapositioned anionic phosphate group of the substrate. A water molecule acts as the second nucleophile and is responsible for the hydrolysis of the phospho-Asp41

intermediate. Other two highly conserved Asp residues, Asp175 and Asp176, which are located on loop 4, contribute to the Mg^{2+} ion binding via H-bond interaction with the water ligands, forming an inner sphere stabilization of Mg^{2+} coordination. Asp43, like Asp41, is also locates on loop one, where it serves as general acid/base, by both protonating the leaving group and activating the nucleophilic water molecule that dephosphorylates the phosphoenzyme intermediate. The bi-functional role of this "Asp+2" residue is conserved in most HADSF members.

3.1.3 Location and Physiological Role of mdN

The first purified and identified mdN derived from human placenta [(10)]. Its activity towards 5'-deoxyribonucleotides was found to be higher than that towards 5'-ribonucleotides. In addition, by observation of fluorescence signal of expressing mdN-green fluorescence protein fusion (GFP) protein in HEK 293 cells and inspecting the nature of the leader sequence, mdN was demonstrated to be located in mitochondria [(11)]. Full-length mdN and its truncated analog (without the first 31 NH_2-terminal leader peptides) expressed in *E. coli.* cells both have high activities towards deoxyribonucleotides.

In cytosol, a deoxynucleoside kinase together with a nucleotidases that catalyze opposing opposite irreversible phosphoryl transfer reactions act to regulate dNTP pools[(12)]. The mitochondrial nucleotidase, mdN, might play the same role in balancing the dNTP pool inside mitochondria. In the enzyme network two pathways exist for dNTP regulation[(13)], including de novo synthesis and ribonucleoside salvage, where de novo synthesis is responsible for the dNTP production by ribonucleotide reductase, deaminase, kinase and pyrophosphatase catalyzed processes. In contrast, the main salvage enzymes consist of a

spectrum of kinases, which phosphorylate all four deoxyribonucleotides. That no mitochondrial ribonucleotide reductase has been found until now[11] suggests that *de novo* synthesis of dNTPs might not take place in mitochondria, and that these substances or the corresponding deoxyribonucleosides are imported from the cytosol into mitochondria. Therefore, as a counterpart of kinases, mdN functions in salvage pathways to regulate the size of dNTP pools. Importantly, the mdN gene localization, which is on chromosome 17p11.2 in the Smith-Magenis syndrome-critical region, raises the possibility that mdN is related to this genetic disease.

While the general mechanism for mdN catalysis has been proposed, no studies have been conducted thus far to elucidate the detailed features of each step of the reaction not has any evidence been accumulated that sheds light on the nature of the phosphor-enzyme intermediate. Moreover, questions about the energy landscape of the process and the nature of the rate-limiting step remain unanswered. In studies described in this chapter, we have obtained preliminary results that provide information about these questions.

3.2 Materials and Experimental

3.2.1 Materials

All chemicals were obtained from Sigma. Primers and T4 DNA ligase were purchased from Invitrogen. Restriction enzymes were purchased from Biolab. *Pfu* polymerase and the pET28a vector kit were purchased from Stratagene. Host cells were purchased from Stratagene. cDNA from human (SC122889) was purchased from OriGene. 10k centrifugal filters were purchased from VWR.

3.2.2 Cloning, Expression, and Purification of mdN

cDNA encoding the gene HAD protein mdN (NCBI accession NM-020201.3, Swiss Pro accession: Q9NPB1) from *Homo sapiens* was amplified by PCR using the cDNA from human and P*fu* DNA polymerase. The first 32 amino acids constituting the mitochondrial localization leader sequence were not part of expressed construct. The PCR product was cloned into the PET28a vector, which was used to transform competent *E. coli.* BL21 (DE3) cells. Oligonucleotide primers (5'-CGGGCTGGGCCATATGGGAGGCCGCGCCCTAC) and (5'-CTGCTGAGCTAAGCTTTGCTTCGGGCTCCT), containing restriction endo- nuclease cleavage sites *NdeI* and *HindIII*, were used in the PCR reactions. The resulting PCR product was digested with restriction enzymes *NdeI* and *HindIII*, and after purification, the desired DNA fragment was ligated to *NdeI*/*HindIII*-linearized pET28a vector. The resulting clone vector was used to transform *E. Coli. BL21 (DE3)* competent cells. The cells were grown at 37 °C overnight on a LB-Agar plate containing 50 mg/L kanamycin. The plasmid DNA of one well-isolated colony was extracted, purified by using a Qiaprep Spin Miniprep Kit and sequenced by MClab. The confirmed clone was then used as a source for expression of recombinant mdN.

Following growth in 2 L of LB media (1x) containing 50 μg/mL kanamycin at 37 °C for 6 h at 200 rpm, induction was initiated with the addition of 0.2 mM isopropyl α-D-thiogalactopyranoside (IPTG). After incubation for 12 h at 20 °C and 160 rpm, the cells grow to an OD_{600nm} ~1.5A and were harvested by centrifugation (6500 rpm for 15 min at 4 °C). Accordingly, the cell pellet was suspended in lysis buffer (50 mM K^+HEPES containing 5 mM $MgCl_2$, 500 mM NaCl and 40 mM imidazole, pH 7.5.) and then passed

through a French Pressure cell press at 1,200 psi. The supernatant, obtained by centrifugation of the cell-free mixture, was loaded onto a 5ml HisTrap FF crude column at a flow rate of 5ml/min equilibrated with lysis buffer. The column was washed with 5CV lysis buffer followed by 5CV of wash buffer (50 mM K^+HEPES, 500 mM NaCl, 500 mM imidazole and 5mM MgCl2 at pH 7.5). Fractions containing mdN, as judged by SDS-PAGE analysis, were combined, concentrated to ~2 ml with an Amicon protein concentrator, and then dialyzed against 50 mM Na^+MES (pH 6, 4 °C) buffer containing 5 mM $MgCl_2$. The dialysate was chromatographed at 4 °C on a HiLoad 16/60 Superdex 200 column using the same dialyzed buffer as eluant. The fractions containing protein were pooled and concentrated.

3.2.3 Preparation of the mdN Mutant D43N

The mdN mutant D43N was prepared by using a PCR based strategy with the pET28a-mdN clone serving as template and commercial oligonucleotides as primers (5'-TGGTGGACATGAACGGCGTGCTGGCTGACTT). The purified clone was used to transform competent *E. coli* BL21 (DE3) cells after confirming the gene sequence by commercial DNA sequencing. The mutant protein was purified to homogeneity (based on SDS-PAGE analysis) from cultured cells using the method previously described for the preparation of the wild type mdN.

3.2.4 Determination of Steady State Kinetic Constants

Steady-state kinetic parameters (K_m and k_{cat}) of mdN and D43N were determined from initial velocities of reactions of nucleotides hydrolysis carried out using varying substrate concentrations in 5 mM $MgCl_2$ in 50 mM Na^+MES assay buffer (pH 6.0) at 25 °C. Protein concentrations were determined by using the Bradford assay[(14)] and inorganic

phosphate product was detected by using the acidified ammonium molybdate method[15]. Data were fit using the SigmaPlot program to equation 3.1,

$$V_0 = V_{max} [S] / (K_m + [S]) \qquad \text{(eq. 3.1)}$$

where V_0 = initial velocity, V_{max} = maximum velocity, [S] = substrate concentration and K_m = Michaelis-Menten constant for substrate. The k_{cat} value was calculated from V_{max} and [E] according to the equation $k_{cat} = V_{max} /[E]$, where [E] is the free enzyme concentration.

The rates of *p*-nitrophenyl phosphate (pNPP) hydrolysis reactions were determined by monitoring the increase in absorbance at 410 nm (ε = 1.86 mM^{-1} cm^{-1}) at 25 °C. The 0.5 mL assay mixtures contained 50 mM Na^+MES, pH 6.0, 5 mM $MgCl_2$, and various concentrations of pNPP.

3.2.5 Multi-turnover, mdN Catalyzed Reactions of [^{14}C(U)] dUMP

Multi-turnover mdN catalyzed reactions of [^{14}C(U)] dUMP reactions were performed at 25 °C using a rapid quench instrument from KinTek Instruments. Each reaction was performed by mixing 16 μL buffer A (50 mM Na^+MES containing 5 mM $MgCl_2$, pH 6.0) containing 600 μM mdN and 16 μL buffer A containing 3000 μM [^{14}C(U)] dUMP. Each reaction was quenched after a specified period of time by addition of 212 μl 0.2 N HCl. Each quenched reaction mixture was passed through a 10-kDa filter to remove the enzyme, and then analyzed by using a Shimadzu high-performance liquid chromatography (HPLC) system equipped with a C18 reverse (4 mm × 250 mm) column. The column was eluted at a flow rate of 1 mL/min by using the following program: mobile phase A was 0.1 mol/L KH_2PO4 (pH 5), and mobile phase B was 70% 0.01 mol/L KH_2PO4 (pH 3.5) and 30% methanol. 95% mobile phase A and 5% mobile phase

B was maintained for 3 min, a 2-min linear gradient to 50% mobile phase A was initiated, 50% mobile phase A was maintained for 18 min, a 2-min linear gradient to 95% mobile phase A was initiated, and 95% mobile phase A was maintained for at least 10 min before the next sample was injected. The radioactive content of the eluted product was measured by using a β-RAM model 4 radio flow-through detector (IN/US System, Inc.). The observed rate constants and k_{cat} values for the multi turnover reactions were obtained by fitting the time course data using the computer programs Kaleidagragh to the burst equations 3.2 and 3.3 [16],

$$[P] = A * (1 - e^{-k_{obs} t}) + k * t \qquad \text{(eq. 3.2)}$$

$$k = k_{cat} * [E]_0 \qquad \text{(eq. 3.3)}$$

where [P] is the observed concentration of the product adenosine, A is the product of the burst amplitude (A_0) and the active enzyme concentration ($[E]_0$), k_{obs} is the burst rate constant, k is the product of the steady-state turnover rate (approaching k_{cat} at saturating substrate) and the active enzyme concentration ($[E]_0$).

3.2.6 GMP/dUMP Binding Constant to Wild Type mdN and the D43N Mutant

The reaction was performed using an Applied Photophysics SX20 sequential stopped-flow spectrophotometer with a light path of 10 mm and a dead time of 2 ms. The light source was a 150-W xenon lamp and slit width was 1 mm. The drive syringes were driven by a pneumatic actuator operated by compressed nitrogen. The enzyme and substrates, GMP and dUMP, were prepared in 50 mM Na^+MES buffer containing 5 mM $MgCl_2$, pH 6. Protein (10 μM after mixing) and substrate (20 μM to 100 μM after mixing) were then mixed in the stopped flow apparatus. The excitation wavelength was

280 nm. The observed first-order rate constants k_{obs} was obtained by fitting the data using the software provided by Applied Photophysics to equation 3.4,

$$F_t = F_\infty + P \exp(-k_{obs}t) \qquad \text{(eq. 3.4)}$$

where F_t = the fluorescence at time t, F_∞ = fluorescence at infinite time, P = amplitude, k_{obs} = the observed first-order rate constant.

Time courses for fluorescence quenching at different GMP and dUMP concentrations and fixed enzyme concentrations were measured to obtain the association rate constant (k_{on}) and dissociation rate constant (k_{off}) for GMP and dUMP binding to mdN wild-type and the D34N mutant. The resulting k_{obs} values were plotted against the ligand concentrations ([L]) and the data were fitted to equation 3.5 in order to define the k_{on} and k_{off} values.

$$k_{obs} = k_{on}[L] + k_{off} \qquad \text{(eq. 3.5)}$$

3.3 Results and Discussion

3.3.1 Purification of Wild Type mdN

Recombinant mdN was prepared using transformed *E. coli* BL21 (DE3) cells and a column-based chromatography protocol for protein purification. Approximately 5 mg of homogeneous protein (determined by using SDS-PAGE analysis as shown in Figure 3.4) was obtained per gram of wet cell pellet. The protein concentration was determined by using the Bradford method and by measuring the absorption at 280 nm. The construct mdN does not include the first 32-NH_2 terminal amino acids, which are believed to constitute the mitocondrial localization peptide sequence that is cleaved after mdN enters mitochondrial space, but it does include the NH_2-terminal methionine. Full-length and

truncated mdN, which are expressed in *E coli.* Cells, have comparable high activities towards deoxyribonucleotides[(11)].

The molecular weight of mdN was determined as 24974.8 Da by using mass spectrometric analysis, a value that is in close agreement with the theoretical molecular weight of 24974.4 calculated by using ExPASy.

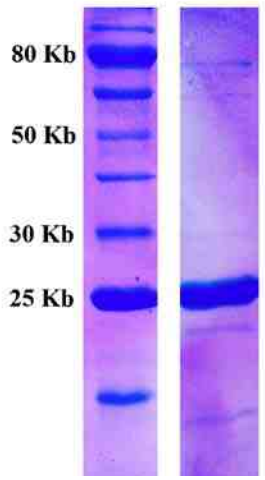


Figure 3.4 SDS-PAGE of the purified mdN construct. (lane 1, protein mass marker; lane 2, mdN construct)

3.3.2 Substrate Specificity of mdN

The steady state kinetic constants k_{cat} and K_m determined for mdN hydrolysis of several organic phosphate esters are listed in Table 3.1, with the value k_{cat}/K_m being a measure of substrate activity. The large values for dUMP (k_{cat}/K_m = 1.3 x 10^5 $M^{-1}s^{-1}$) and TMP (k_{cat}/K_m = 1.1 x 10^5 $M^{-1}s^{-1}$) indicate that these substances could be the natural substrates of mdN. UMP is also a good substrate, having a two fold lower k_{cat}/K_m than dUMP. In

Table 3.1 The steady-state kinetic constants k_{cat} and K_m values for mdN hydrolyzing various substrates are measured in 50 mM Na^+MES buffer containing 5mM $MgCl_2$, pH 6.0, 25 ^oC.

Substrate	K_m (uM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
dUMP	$(2.2 \pm 0.1) \times 10^2$	27.0 ± 0.4	1.3×10^5
TMP	$(9.0 \pm 0.5) \times 10^1$	10.3 ± 0.2	1.1×10^5
UMP	$(1.9 \pm 0.1) \times 10^2$	12.8 ± 0.2	6.7×10^4
GMP	$(1.2 \pm 0.1) \times 10^2$	$(3.3 \pm 0.1) \times 10^{-1}$	2.8×10^3
IMP	$(1.9 \pm 0.1) \times 10^2$	$(6.4 \pm 0.1) \times 10^{-1}$	3.3×10^3
dAMP	$(4.0 \pm 0.2) \times 10^2$	1.6 ± 0.1	3.8×10^3
dCMP	$(5.8 \pm 0.4) \times 10^2$	$(9.0 \pm 0.2) \times 10^{-2}$	1.6×10^2
3-AMP	$(1.5 \pm 0.1) \times 10^3$	$(2.1 \pm 0.1) \times 10^{-1}$	1.4×10^2
AMP	$(1.2 \pm 0.1) \times 10^3$	$(3.2 \pm 0.1) \times 10^{-1}$	2.7×10^2
CMP	$(3.0 \pm 0.1) \times 10^3$	$(3.3 \pm 0.1) \times 10^{-1}$	1.1×10^2
ADP	$(4.4 \pm 0.2) \times 10^3$	$(6.4 \pm 0.2) \times 10^{-1}$	1.4×10^2
3,5-ADP	$(7.6 \pm 0.5) \times 10^3$	$(7.0 \pm 0.3) \times 10^{-1}$	9.4×10^1
UTP	$(2.8 \pm 0.3) \times 10^3$	1.7 ± 0.1	6.3×10^2
pNPP	$(5.0 \pm 0.3) \times 10^3$	16.4 ± 0.5	3.3×10^3
Glucose-6-phosphate	ND	NH	ND
Glycerol-2-phosphate	ND	NH	ND
D-glucosamine-6-phosphate	ND	NH	ND
Ribose-5-phosphate	ND	NH	ND

ND, not detected. NH, no hydrolysis was detected with an enzyme concentration of 0.5 μM and substrate concentrations up to 0.5 mM in a final volume of 0.3 mL, for 5 min, at 25 °C.

addition, purine mononucleotides are also active substrates even though their k_{cat}/K_m values are 10-fold lower than the corresponding pyrimidine mononucleotides.

In comparison, nucleotide diphosphates and triphosphates are poor substrates with k_{cat}/K_m values of about 100 $M^{-1}s^{-1}$. The aryl phosphate pNPP is 100-fold less active than dUMP, however, once it is bound to enzyme its hydrolysis rate is very large (k_{cat} = 16.4 s^{-1}). Finally, the sugar phosphates, glucose-6-phosphate and ribose-5-phosphate, and glycerol-2-phosphate are not substrates for mdN.

Based on the results presented above (Table 3.1), it is likely that the natural substrates of mdN are those that contain aromatic rings and not aliphatic phosphate esters. Analysis of the crystal structures of native mdN complexed with thymidine[(6)] and with the nucleotide analogs AZTMP and BVdUMP[(17)] show that a substrate binding pocket formed by all the aromatic residues Phe75, Phe49, Trp76 and Trp96, exists and that the two Phe side chains in the enzyme stack parallel to the aromatic ring of the substrate. Thus, the binding pocket governs the substrate preference of mdN. The lower activity of mdN with purine mononucleotides compared to pyrimidine analogs can be explained by information gained from inspection of the crystal structures of complexes of mdN mutant D41N with bound 5'-dGMP and 5'-dUMP[(18)]. In the structure of the D41N mutant, the nucleophile Asp-41 is replaced by the poorer nucleophile asparagine, which forms hydrogen bonds to the substrate but it does not promote dephosphorylation. The structure of D41N with the bound purine derivative 5'-dGMP reveals that the distance between O5' and the general acid/base Asp43 is larger than when the pyrimidine 5'-monophosphate is bound, making proton transfer from Asp43 to the leaving group R less efficient. This phenomenon results in a lower activity of mdN with purine mononucleotides. mdN prefers the 2-

deoxyribo form of 5'- nucleoside monophosphate rather than ribo form. The crystal structures of D41N bound with 5'-UMP and 5'-dUMP show that the 2'-hydroxyl group of 5'-UMP is positioned in a hydrophobic pocket which is comprised of Phe-49, Phe-102 and Ile-133. Therefore, all the interactions between 2'-hydroxyl group and the surrounding residues are energetically unfavorable, which is the likely cause for the lower activity of ribo 5'-nucleoside monophosphate.

3.3.3 Kinetic Analysis of Wild Type mdN Catalyzed Hydrolysis of [^{14}C/U]dUMP under Multi-turnover Reaction Conditions

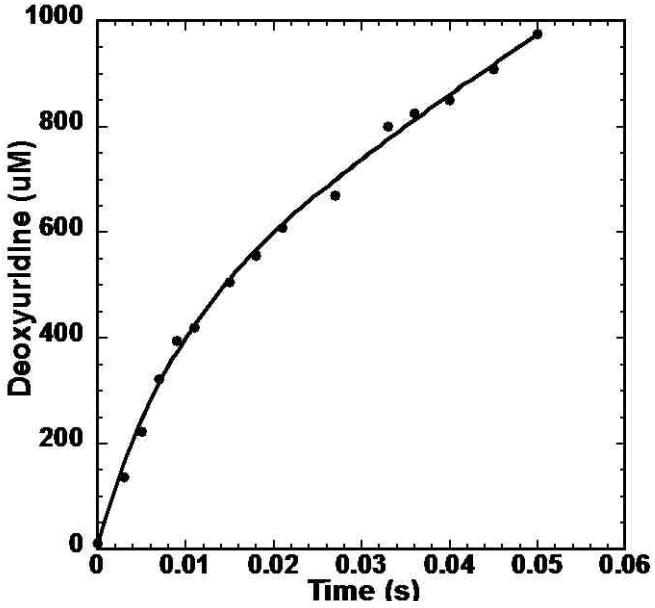


Figure 3.5 The time course of the multi turnover reaction of 300 μM mdN with 1500 μM dUMP in 50 mM Na^+MES buffer containing 5 mM $MgCl_2$, pH 6.0, and at 25°C. (□) experimental data; the curve shows the fit to the burst equation.

Pre-steady-state kinetic analysis has been used to detect intermediates in reaction pathways and to measure the rates of formation and decay of the intermediate. mdN-catalyzed dUMP hydrolysis under multi-turnover reaction conditions was carried out in 50 mM Na^+MES buffer, containing 5 mM $MgCl_2$, pH 6.0, 25 °C with 300 μM enzyme and excess substrate (1500 μM). At fixed time intervals (3 – 50 ms), the reactions were quenched and the amount of deoxyuridine produced was determined. A pre-steady-state burst was observed when the concentration of deoxyuridine was plotted versus reaction time (Figure 3.5). The formation of the burst occurred with a rate constant of 120 ± 18 s^{-1} and was followed by a slow steady-state breakdown of the phosphoenzyme intermediate, with a rate constant k_{cat} of 36.7 ± 3.3 s^{-1}. The rate constants were calculated uisng the burst phase equation and kcat was obtained by dividing k_{obs} by the enzyme concentration (300 uM).

3.3.4 Binding Constant of GMP and dUMP to Wild Type mdN and the D43N Mutant Measured by Using a Stopped-Flow Fluorescence Technique

It is likely that the observation of burst kinetics for mdN-catalyzed hydrolysis of dUMP described above shows that formation of phosphoenzyme intermediate is faster than the ensuing hydrolysis step. However, the “burst” could also be caused by a change in the conformation of the enzyme upon substrate binding.

Analysis of the crystal structure of wild type mdN shows that two tryptophans, Trp 76 and Trp 96, are present in the substrate binding pocket[(6)]. This feature provides an opportunity to monitor conformation changes of wild type mdN upon the substrate binding. A study of time-dependent fluorescence quenching of wild-type mdN (10 μM) promoted by binding of dUMP (80 μM) was conducted under pseudo-first-order

conditions, defined by a ligand to enzyme concentration ratio of 8:1 (Figure 3.6). The observed first-order rate constant, $k_{obs,}$ of 178 ± 110 s^{-1} was obtained by fitting the data to equation 3.4, which describes a single-exponential decay. Various concentrations (20, 30, 40, 50, 60 and 80 μM) of dGMP were used to measure the association rate constant (k_{on}) and dissociation rate constant (k_{off}) associated with binding to wild-type mdN. This treatment gave k_{on} = 11 ± 0.4 $\mu M^{-1}s^{-1}$ and k_{off} = 260 ± 21 s^{-1}.

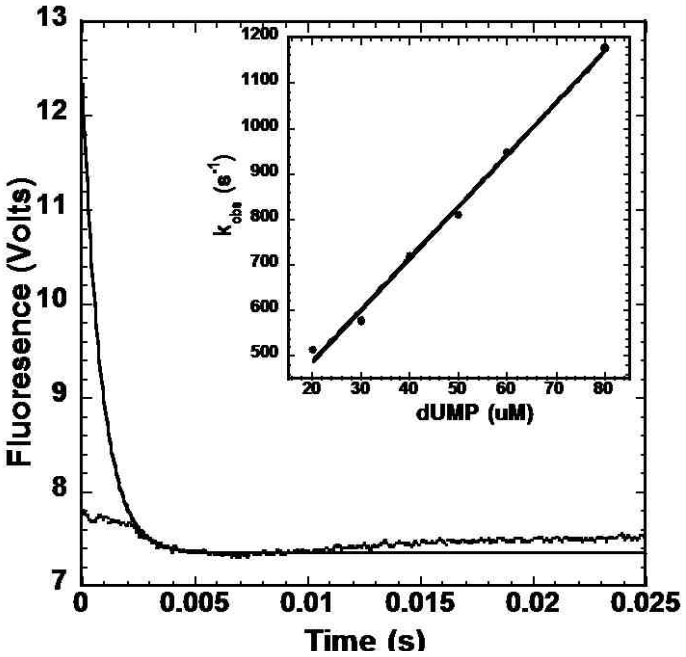
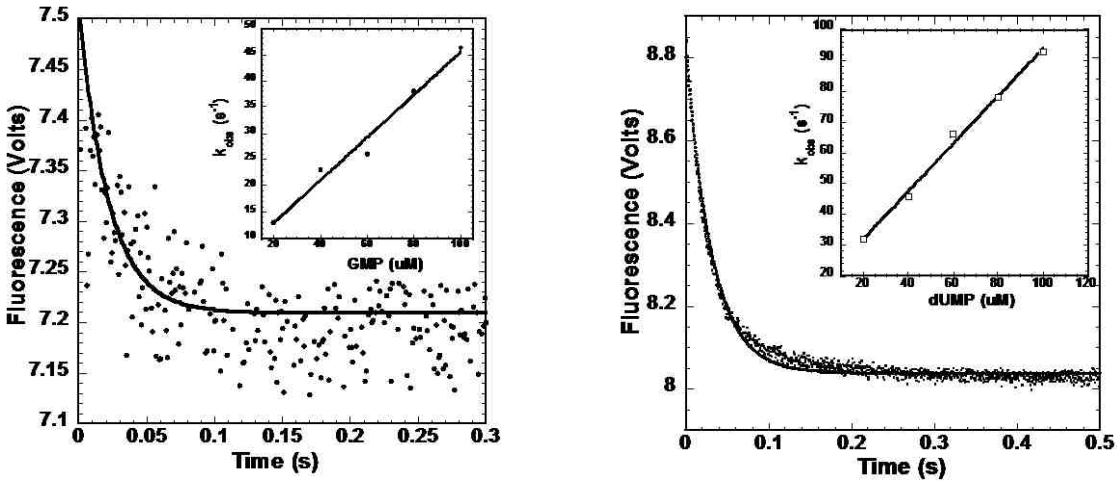


Figure 3.6 Time-dependent fluorescence change associated with the binding between 80 μM dUMP and 10 μM mdN wild type in 50 mM Na^{+}MES containing 5mM $MgCl_2$ (pH 6.0, 25 ^{o}C). The data were fitted to a single-exponential equation using the software provided by Applied Photophysics to obtain an observed rate constant (k_{obs}) of 1178 ± 110 s^{-1}. The inset shows the dependence of k_{obs} on the concentration of dUMP (20 - 80 μM).



(A) (B)

Figure 3.7 Time-dependent fluorescence change associated with the binding reaction **(A)** between 100 μM GMP and 10 μM mdN wild type in 50 mM Na^+MES containing 5 mM $MgCl_2$ (pH 6.0, 25 ^oC). **(B)** between 20 μM dUMP and 10 μM D43N. The data were fitted to a single-exponential equation using the software provided by Applied Photophysics to obtain an observed rate constant (k_{obs}) of (A) is 46.2 ± 4.2 s^{-1} and of (B) is 32.2 ± 0.3 s^{-1}. The inset shows the dependence of k_{obs} on the concentration of (A) GMP (20 – 100 μM) and (B) dUMP (20 - 100 μM).

Because the turnover rate of mdN with dUMP is fast (27 ± 0.4 s^{-1}), it is difficult to analyze the conformation change associated with binding of this substrate. The time course of the first 25 ms after mixing of mdN with dUMP is shown in Figure 3.6. The fluorescence signal at first decreases and then it increases, phenomena which might be caused by either the involvement of more than one step in binding of dUMP binding or by the possibility that the changes are associated with combined substrate binding and hydrolysis.

To determine which of these possibilities is operating, the relatively poor substrate GMP (k_{cat} = 0.33 ± 0.005 s^{-1}) along with the D43N mutant, in which the general acid/base Asp+2 aspartic acid residue is replaced by asparagine, were used in binding experiments. As the results shown in Table 3.2 demonstrate, the turnover rate of dUMP promoted by D43N is about 2000 fold lower than the wild type with a similar K_m value.

Table 3.2 The steady-state kinetic constants k_{cat} and K_m values for mdN D43N mutant catalyzed hydrolysis of dUMP measured in 50 mM Na^+MES buffer containing 5mM $MgCl_2$, pH 6.0, 25 ^oC.

Protein	K_m (uM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
Wild Type mdN	(2.2 ± 0.1) x 10^2	27.0 ± 0.4	1.3 x 10^5
D43N	(3.9 ± 0.1) x 10^2	(1.4 ± 0.1) x 10^{-2}	3.6 x 10^1

Time-dependent fluorescence quenching of wild-type mdN (10 μM) caused by binding of GMP (100 μM) and of the D43N mutant (10 μM) caused by binding of dUMP (20 μM) were monitored under pseudo-first-order conditions (Figure 3.7). In both cases, the stopped-flow determined, fluorescence time courses show continuous decreases in fluorescence intensities. These findings suggest that only one step is involved in both processes and that the step corresponds to substrate binding (E + S = ES). Fitting the fluorescence data to equation 3.4, gave the observed first-order rate constants for wild type binding of GMP (46.2 ± 4.2 s^{-1}) and for D43N binding dUMP of (32.2 ± 0.3 s^{-1}). The association rate constants (k_{on}) and dissociation rate constants (k_{off}) of substrate binding to the enzymes were obtained from experiments using various substrate

concentrations. The rate constants obtained were, for wild type mdN binding of GMP, k_{on} = 0.4 ± 0.03 $\mu M^{-1}s^{-1}$, k_{off} = 4.8 ± 2.3 s^{-1}, and for D43N binding of dUMP, k_{on} = 0.8 ± 0.03 $\mu M^{-1}s^{-1}$, k_{off} =17 ± 2.1 s^{-1}.

3.4 Conclusion

The x-ray crystal structures of wild type mdN and its complexes with product and substrate analogs and an intermediate mimic, along with those of the D41N mutant bound with substrates[6, 17, 18] provide direct insight into the nature of the phosphate ester hydrolysis reaction pathway. In addition, the substrate screening results (Table 3.1) agree with those coming from analysis of the crystal structures. Specifically, the observation that the substrate binding pocket in the enzyme contains several aromatic residues, including Phe49, Phe75, Trp76 and Trp96, is in full accord with the finding that that the catalytically most efficient substrates of mdN are those that contain aromatic rings and not aliphatic phosphate esters (Table 3.1).

In 2005, Nordlund and his coworkers[18] described seven crystal structures of D41N variant in complex with different substrates binding. Inspection of these structures led to the suggestion that two different phosphate binding modes, mode A and mode B, are adopted depending on the nature of the substrate. For example, they suggested that dUMP binds with mode B of the enzyme while binding of dGMP, and likewise GMP, occurs at a different location. However, they proposed that binding with mode B might be an artifact associated with the mutants used.

In our studies, we monitored binding of dUMP and GMP to mdN using a stopped-flow, fluorescence assay procedure. The observation that both binding processes follow

continuous decay modes demonstrates that only one binding mode/site exists in the enzyme. Therefore, Nordlund's finding is likely a consequence of a charge distribution change caused by substitution of a carboxyl group with an amide group in the D41N mutant.

The results of studies of the rapid-quench reaction of mdN-catalyzed dUMP under multi-turnover conditions revealed that a pre-steady-state burst occurs initially followed by slow steady-state turnover. Pre-steady-state kinetic analysis has been used to detect intermediates on reaction pathway. Therefore, the "burst" shows that formation of the intermediate in the catalytic pathway is much faster than the ensuing steps. Also, the possibility that the "burst" is caused by an enzyme conformation change upon substrate binding was ruled out by the results coming from the binding experiments (Figure 3.6).

Reference

1. Giblett, E. R., Ammann, A. J., Wara, D. W., Sandman, R., and Diamond, L. K. (1975) Nucleoside-phosphorylase deficiency in a child with severely defective T-cell immunity and normal B-cell immunity, *Lancet 1*, 1010-1013.
2. Cohen, A., Hirschhorn, R., Horowitz, S. D., Rubinstein, A., Polmar, S. H., Hong, R., and Martin, D. W. (1978) Deoxyadenosine Triphosphate as a Potentially Toxic Metabolite in Adenosine-Deaminase Deficiency, *P Natl Acad Sci USA 75*, 472-476.
3. Nishino, I., Spinazzola, A., and Hirano, M. (1999) Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder, *Science 283*, 689-692.
4. Reichard, P. (1988) Interactions between Deoxyribonucleotide and DNA-Synthesis, *Annu Rev Biochem 57*, 349-374.
5. Bianchi, V., and Spychala, J. (2003) Mammalian 5'-nucleotidases, *The Journal of biological chemistry 278*, 46195-46198.
6. Rinaldo-Matthis, A., Rampazzo, C., Reichard, P., Bianchi, V., and Nordlund, P. (2002) Crystal structure of a human mitochondrial deoxyribonucleotidase, *Nat Struct Biol 9*, 779-787.
7. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D., and Aravind, L. (2006) Evolutionary genomics of the HAD superfamily: Understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes, *J Mol Biol 361*, 1003-1034.
8. Allen, K. N., and Dunaway-Mariano, D. (2004) Phosphoryl group transfer: evolution of a catalytic scaffold, *Trends Biochem Sci 29*, 495-503.

9. Herschlag, D., and Jencks, W. P. (1989) Phosphoryl Transfer to Anionic Oxygen Nucleophiles - Nature of the Transition-State and Electrostatic Repulsion, *J Am Chem Soc 111*, 7587-7596.

10. Rampazzo, C., Johansson, M., Gallinaro, L., Ferraro, P., Hellman, U., Karlsson, A., Reichard, P., and Bianchi, V. (2000) Mammalian 5 '(3 ')-deoxyribonucleotidase, cDNA cloning, and overexpression of the enzyme in Escherichia coli and mammalian cells, *Journal of Biological Chemistry 275*, 5409-5415.

11. Rampazzo, C., Gallinaro, L., Milanesi, E., Frigimelica, E., Reichard, P., and Bianchi, V. (2000) A deoxyribonucleotidase in mitochondria: Involvement in regulation of dNTP pools and possible link to genetic disease, *P Natl Acad Sci USA 97*, 8239-8244.

12. Bianchi, V., Pontis, E., and Reichard, P. (1986) Interrelations between substrate cycles and de novo synthesis of pyrimidine deoxyribonucleoside triphosphates in 3T6 cells, *Proc Natl Acad Sci U S A 83*, 986-990.

13. Rampazzo, C., Miazzi, C., Franzolin, E., Pontarin, G., Ferraro, P., Frangini, M., Reichard, P., and Bianchi, V. (2010) Regulation by degradation, a cellular defense against deoxyribonucleotide pool imbalances, *Mutat Res-Gen Tox En 703*, 2-10.

14. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Analytical biochemistry 72*, 248-254.

15. Ames, B. N. (1966) Assay of inorganic phosphate, total phosphate and phosphatases, *Methods Enzymol. 8*, 115-118.

16. Johnson, K. (2003) *Kinetic analysis of macromolecules : a practical approach*, Oxford University Press, Oxford.

17. Wallden, K., Rinaldo-Matthis, A., Ruzzenente, B., Rampazzo, C., Bianchi, V., and Nordlund, P. (2007) Crystal structures of human and murine deoxyribonucleotidases: Insights into recognition of substrates and nucleotide analogues, *Biochemistry-Us 46*, 13809-13818.

18. Wallden, K., Ruzzenente, B., Rinaldo-Matthis, A., Bianchi, V., and Nordlund, P. (2005) Structural basis for substrate specificity of the human mitochondrial deoxyribonucleotidase, *Structure 13*, 1081-1088.

CHAPTER FOUR

4. Structure Function Analysis of a Haloacid Dehalogenase (HAD) Sugar Acid phosphatase YidA from *Escherichia coli*

4.1 Introduction

The haloacid dehalogenase superfamily (HADSF) is ubiquitous with members existing in all kingdoms of Life. These proteins, present in each organism, catalyze phosphate monoester hydrolysis reaction in the contexts of primary[(1)] and secondary metabolism [(2, 3)], regulation of metabolic pools [(4)], signal transduction[(5)], membrane remodeling [(6)], and active membrane transport [(7)]. To date, more than 4,000 gene sequences encoded HADSF proteins have been deposited in the gene bank, while only a fraction of these have defined structures and/or functions.

As mentioned in chapter one, domains containing structural elements for catalysis and substrate recognition are physically separated in HADSF members, a likely reason why the HADSF has been so highly evolvable. This feature facilitates new function acquisition by conserving the catalytic scaffold and using the cap domain to gain new substrate recognition. In this chapter, a study focusing on the HADSF phosphatase, YidA, from *E coli.*, which contains a C2 cap, is described. The goals of the effort were to (1) elucidate how the cap domains of HADS facilitate the evolution of new function, (2) identify the structural determinants that govern the expansion of the HADSF, and (3) formulate a unified model for the biological process that has formed this ubiquitous HADSF.

YidA is an effective catalyst for the hydrolysis of several known phosphate metabolites

such as arabinose-5-phosphate, erythrose-4-phosphate, α-mannnose-1-phosphate and α-glucose 1-phosphate as well as two oxidized derivatives (4-phospho-D-erythronate and 5-phospho-D-arabinonate). YidA is not located in an operon but it resides in the same "neighborhood" as the *dgo* operon (Figure 4.1). The dgo genes encode enzymes of the galactonate to pyruvate and glyceraldehyde 3-phosphate degradation pathway. The HADSF type C2 phosphatases are prevalent in secondary metabolic pathways. YdiA may play a role of a housekeeper to remove accumulated phosphorylated metabolites.

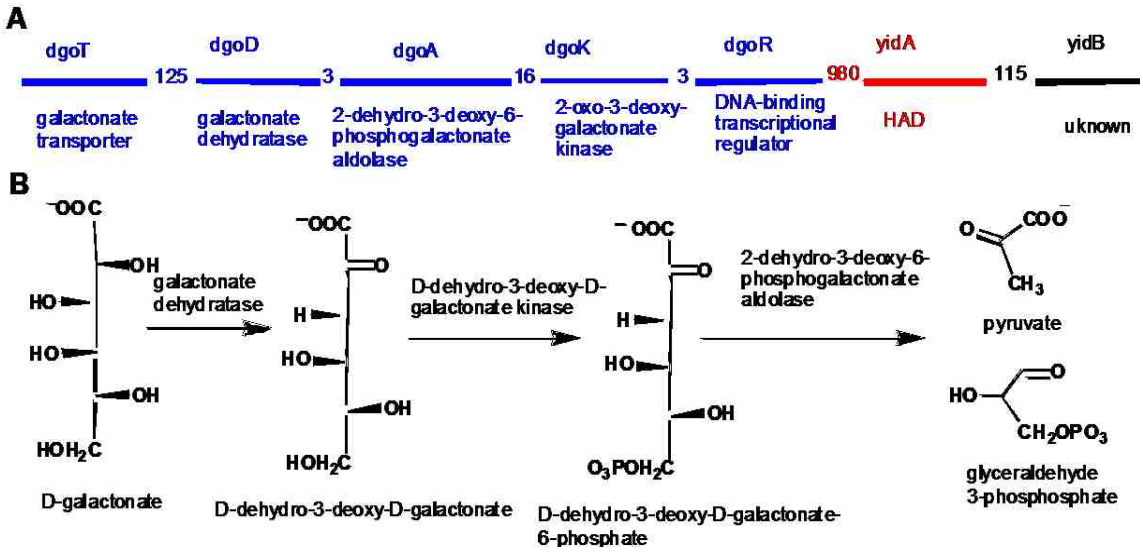


Figure 4.1 A. The *dgo* operon. The numbers represent the number of intervening nucleotides between genes. **B.** Schematic of the galactonate degradation pathway.

Analysis of the structure of YidA bound with a Mg^{2+} cofactor shows that it belongs to the C2b structural class, which has the αββ(αβαβ)αββ cap domain topology (Figure 4.2). The structure depicts the enzyme in the cap-open conformation. Catalytic turnover requires that the cap domain and catalytic domain associate so that the catalytic site becomes desolvated. The residues that bind the substrate-leaving group are located on the

cap domain. Thus, in order to determine the location of these residues a structure of the enzyme, in the cap closed conformation and complexed to substrate or to an inert substrate analog, is required.

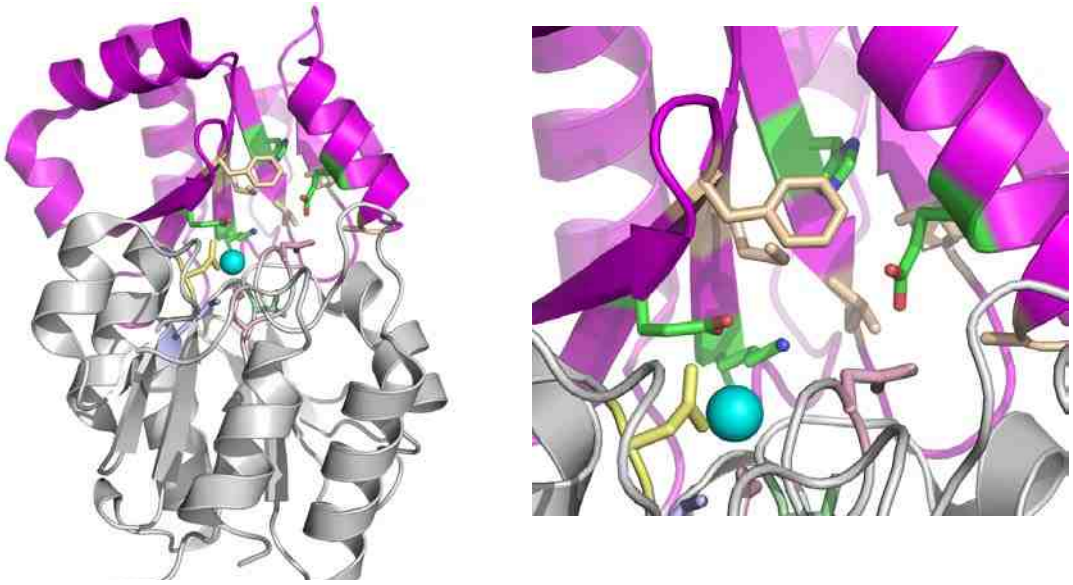


Figure 4.2 Crystal structure of *apo* YidA (1RKQ). The Mg^{2+} cofactor is shown as a blue sphere, the catalytic scaffold is colored in pastels, the core domain is gray, the cap domain magenta. The potential substrate binding groups of the cap domain are shown as sticks: polar are green whereas the nonpolar residues are tan.

4.2 Experimental Methods

4.2.1 Material

All chemicals were obtained from Sigma-Aldrich except where mentioned. The Enzcheck phosphate assay kit was purchased from Invitrogen Inc. Primers and T4 DNA ligase were from Invitrogen. Restriction enzymes were purchased from BioLab.

Deepvent polymerase, *Pfu* Turbo polymerase, pET28a and pET23a vectors were purchased from Stratagene. The Geneclean Spin Kit and the Qiaprep Spin Miniprep Kit were from Qiagen. Host cells were purchased from Novagen. Genomic DNA encoding YidA from *Escherichia coli* (ATCC 10798) were purchased from ATCC. The pET23a construct which contains the gene encoding YidA was from Dr. Liangbing Wang.

4.2.2 Purification of YidA

The *E. coli*. BL21 (DE3) cells transformed by pET23a construct were grown at 37 °C with agitation at 200 rpm in Luria broth (LB) containing 50 μg/mL ampicillin for 4-6 h to an OD_{600} = 0.6-0.8 and induced for 6 h at 20 °C at a final concentration of 0.4 mM isopropyl α-D-thiogalactopyranoside (IPTG). The cells were harvested by using centrifugation at 6500 rpm for 15 min at 4 °C to yield 3 g/L of culture medium. The cell pellet was suspended in 1 g wet cell/10 mL of ice-cold buffer consisting of 50 mM K^+-HEPES (pH 7.0 at 25 °C), 5 mM Mg^{2+} and 1 mM DTT. The cell suspension was passed through a French press at 1,200 PSIG before centrifugation at 20,000 rpm and 4 °C for 30 min. The supernatant was loaded onto a 5 mL HiTrap DEAE FF column (GE healthcare), which was eluted at 4 °C with a 0.2 L linear gradient of NaCl (0 to 0.5 M) in buffer A (50 mM K+HEPES, pH 7.0 containing 5 mM $MgCl_2$ and 1 mM DTT). The column fractions were analyzed by using SDS-PAGE, and the desired fractions were combined and concentrated at 4 °C using a 10K Amicon Ultra Centrifugal filter (Millipore). The concentrate was loaded onto a HiLoad 16/60 Superdex 200 column and the column was eluted at 4 °C with buffer A. The column fractions were analyzed by using SDS-PAGE, and the desired protein fractions were combined, concentrated at 4 °C and then stored at -80 °C. The yield is 8 mg protein/g wet cells.

The molecular weight of the protein was measured by using mass spectrometry (University of New Mexico Mass Spectrometry Facility). The theoretical molecular mass was calculated from the amino acid composition, derived from the gene sequence using the EXPASY Molecular Biology Server program [8]. The natural MW was estimated by FPLC size exclusion gel filtration chromatography (2.5 cm × 120 cm, GE Healthcare Hiprep 16/60 Sephacryl S-200 High Resolution column) eluted at 4 °C with 50 mM K^+ Hepes, pH 7.5 and 100 mM NaCl against protein standards (25-232 kDa, Amersham Pharmacia Biotech).

4.2.3 YidA Site-Directed Mutagenesis

YidA site-directed mutagenesis was prepared by a PCR based strategy with the pET28a-YidA clone serving as template and commercial oligonucleotides as primers (as listed in Table 4.1). The purified clone was used to transform competent *E. coli* BL21 (DE3) cells after confirming the gene sequence by commercial DNA sequencing at MC lab. The transformed cells were grown and harvested using the same procedure mentioned above. And then the cell pellet was suspended in 1g wet cell/10 mL of ice-cold lysis buffer (20 mM Tris•HCl, 40 mM imidazole, 1 mM DTT and 500 mM NaCl, pH 7.5). The cell suspension was passed through a French press at 1,200 PSIG before centrifugation at 20,000 rpm and 4°C for 30 min. The supernatant was loaded onto a 5mL HisTrap FF crude column (GE Healthcare) at a flow rate of 5 mL/min. The column was washed with 5 column volumns (CV) lysis buffer followed by 5 CV of wash buffer (20 mM Tris•HCl, 500 mM NaCl, 500 mM imidazole and 1mM DTT at pH 7.5). Fractions containing target protein, as judged by SDS-PAGE analysis, were combined, concentrated with a 10K Amicon Ultra Centrifugal filter (Millipore), and then stored at -80 °C.

Table 4.1 Primers used in YidA mutagenesis.

D11A	5’ - ATTGCTATCGATATGGCTGGCACCCTTCTGCTG
D11N	5’ - GCTATCGATATGGAATGGCACCCTTCTGCTGCCC
R46A	5’ - TAACGACGGGTGCCCCGTATGCAGGTGTGCA
H129A	5’ - GGCAACGAAGGATTCAGCCACCGTGTAGTAGC
S184A	5’ - ACCGTGCTGAAAGCTGCGCCGTACTTC
L111A	5’ - CATTTCCACGCCGCGGACCGCACCACG
Y187A	5’ - CTGAAAAGTGCGCCGGCCTTCCTCGAAATCCT
M156A	5’ - CATCAATCATCGCCACTTTCAGGAACTGGG
E130A	5’ - CTACACGGTGCATGCATCCTTCGTTGCCA
E190A	5’ - ACGCGTTTATCGAGGATTGCGAGGAAGTACG
F188A	5’ - AAGTGCGCCGTACGCCCTCGAAATCCTCGAT
K154A	5’ - CATCAATCATCATCACTGCCAGGAACTGGGT

4.2.4 Steady State Kinetic Constant Determination

The steady-state kinetic parameters K_m and k_{cat} of phosphorylated substrates were determined from initial reaction velocities measured at varying substrate concentrations (ranging from 0.5-5 K_m) for reactions containing 5 mM $MgCl_2$ in 50 mM K^+Hepes assay buffer (pH 7.5) at 25 °C. The total level of released inorganic phosphate was measured using Enzcheck Phosphate Assay Kit (Invitrogen). The reaction was initiated by adding 0.01-10 μM protein. Protein concentrations were determined by using the Bradford assay [(9)]. Data were fit using the SigmaPlot program to equation 4.1,

$$V_0 = V_{max} [S] / (K_m + [S]) \qquad \text{(eq. 4.1)}$$

, where V_0 = initial velocity, V_{max} = maximum velocity, [S] = substrate concentration and K_m = Michaelis-Menten constant for substrate. The k_{cat} value was calculated from V_{max} and [E] according to the equation $k_{cat} = V_{max} /[E]$, where [E] is the free enzyme concentration.

The rate of *p*-nitrophenyl phosphate (pNPP) hydrolysis was determined by monitoring the increase in absorbance at 410 nm ($\Delta\varepsilon$=18.4 mM^{-1} cm^{-1}) at 25°C. The 0.5 mL assay mixtures contained 50 mM K^+Hepes, pH 7.5, 5 mM $MgCl_2$, and various concentrations of pNPP.

4.2.5 pH Profile Analysis

Initial velocities of the *E.coli.* YidA catalyzed hydrolysis of ribose-5-phosphate were measured at 25 ^oC by monitoring the amount of released free phosphate using an Enzcheck phosphate assay kit (Invitrogen). The 0.5 mL assay mixture, containing 5 mM $MgCl_2$, 0.5-5 K_m substrate, was incubated at 25 °C for 10 min or longer until the absorbance of the solution at 360 nm reached a constant value to account for background hydrolysis, and then 0.025 μM enzyme was added. The enzyme concentration was determined using the Bradford assay. The buffer solutions used were 50 mM Na^+MES, for pH 5.5, 6.0 and 6.5, and 50 mM K^+Hepes for pH 7.0, 7.5 and 8.0. The K_m and v_{max} were acquired by fitting the rates into the SigmaPlot program. The k_{cat} value was calculated from the ratio of v_{max} and the enzyme concentration. The pH dependence of the steady-state kinetic constants were fit initial to equation 4.2, where

$$\log Y = \log [C / (1 + [H]/K_a + K_b/[H])] \qquad \text{(eq. 4.2)}$$

Y is the observed value of either k_{cat}/K_m or k_{cat}, C is the pH-independent plateau value of Y, [H] is the proton concentration, K_a and K_b are apparent acid/base dissociation constants for groups on the enzyme or substrate.

4.2.6 Preparation of " Metal-Free" Wild-Type YidA and Measurement of Catalytic Dependence on Mg^{2+}

Freshly prepared *YidA* was dialyzed first against four changes of 50 mM K^+HEPES (pH 7.5) containing 20 mM EDTA and 1 mM DTT and then against 50 mM K^+HEPES (pH 7.5)/1 mM DTT at 4 ^oC. The initial velocity of YidA hydrolyzed erythrose-4-phosphate reactions in solutions containing 0-0.1 mM $MgCl_2$ in 1 mL of 50 mM K^+HEPES (pH 7.5; 25 ^oC) was determined by using EnzChek assay (Invitrogen). The kinetic data were analyzed by using the computer program of Cleland [10] using equation 4.3,

$$V_o = V_m [A] / (K_M + [A]) \qquad \text{(eq. 4.3)}$$

where [A] is the $MgCl_2$ concentration, V_o is the initial velocity, V_m is the maximum velocity, K_M is the Michaelis constant for Mg^{2+} activation. The k_{cat} was calculated from the ratio of V_{max} and the enzyme concentration.

4.2.7 Determination of Inhibition Constants

Competitive inhibition constants K_i for different inhibitors were determined by measuring the initial velocities of YidA catalyzed substrate hydrolysis as a function of substrate concentrations (0.5 K_m to 5 K_m) and inhibitor concentrations (0 – 5 K_i). The initial velocity data were fitted using SigmaPlot Enzyme Kinetic Module (competitive inhibition) to equation 4.3,

$$V = V_{max} [A] / ([A] + K_m (1 + [I]/K_i)) \qquad \text{(eq. 4.4)}$$

where [A] is the substrate concentration, V is the initial velocity, V_{max} is the maximum

velocity, K_m is the Michaelis constant, K_i and K_{ii} are the inhibition constant and [I] is the inhibitor concentration.

The competitive inhibition by the phosphate analog tungstate was measured using ribose-5-phosphate as a substrate. Reactions were carried out in 50 mM Tris (pH 7.5) buffer containing 5mM $MgCl_2$, 0.02μM wild-type YidA, and varying concentrations of ribose-5-phosphate (0.3-10 mM) with different concentrations of tungstate (0 - 20 μM), at 25 °C. The inhibition by Ca^{2+} on activity of YidA was tested in the presence of Mg^{2+}. The initial velocities of YidA catalyzed erythrose-4-phosphate vs Mg^{2+} concentration were measured by varying the concentration of Ca^{2+} (0 - 20 μM). The reaction was initiated by adding 0.02 μM YidA in 50 mM Tris (pH 7.5) buffer, at 25 °C.

4.3 Results and Discussion

4.3.1 YidA Substrate Specificity

The YidA phosphatase activity toward hydrolysis of pNPP was determined. The moderate activity reflected in k_{cat}/K_m = 16.6 $M^{-1}s^{-1}$ (Table 4.2) shows that YidA is a phosphatase. To identify the substrate specificity of wild type YidA, several aliphatic phosphates were tested as substrates (Table 4.2). The genome context of the gene encoding YidA suggested that this enzyme might play the role of a housekeeper in galactonate metabolism. The activities toward 2-keto-3-deoxy 6-phosphogalactonate

Table[a] 4.2 Steady-state kinetic constants k_{cat} and K_m for wild type YidA catalyzed hydrolysis of phosphate esters in 50 mM K^+Hepes (pH 7.5, 25 ^oC) containing 5 mM $MgCl_2$.

substrate	k_{cat} (sec^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
p-nitrophenylphosphate	$(3.5 \pm 0.1) \times 10^{-2}$	$(2.2 \pm 0.1) \times 10^3$	1.7×10^1
5-phospho-D-arabinonate	9.5 ± 0.2	$(5.2 \pm 0.3) \times 10^2$	1.9×10^4
D-erythrose 4-phosphate	10.2 ± 0.5	$(7.5 \pm 0.6) \times 10^2$	1.4×10^4
4-phospho-D-erythronate	1.0 ± 0.1	$(1.0 \pm 0.1) \times 10^2$	1.0×10^4
D-arabinose 5-phosphate	14.0 ± 0.4	$(2.2 \pm 0.2) \times 10^3$	6.5×10^3
2-keto-3-deoxy-6-phosphogalactonate	1.6 ± 0.1	$(2.5 \pm 0.4) \times 10^2$	6.4×10^3
2-keto-3-deoxy-6-phosphogluconate	3.2 ± 0.3	$(0.8 \pm 0.1) \times 10^3$	4.0×10^3
α-D-mannose 1-phosphate	14.6 ± 0.3	$(5.2 \pm 0.3) \times 10^3$	2.8×10^3
α-D- glucose 1-phosphate	4.2 ± 0.1	$(2.7 \pm 0.2) \times 10^3$	1.5×10^3
D-fructose 6-phosphate	3.8 ± 0.2	$(6.6 \pm 0.7) \times 10^3$	5.8×10^2
D-ribose 5-phosphate	8.4 ± 0.5	$(11.0 \pm 0.1) \times 10^3$	8.1×10^2
D-glucose 6-phosphate	6.0 ± 0.6	$(2.1 \pm 0.3) \times 10^2$	2.9×10^2
α-DL-glycerophosphate	1.0 ± 0.1	$(3.7 \pm 0.5) \times 10^3$	2.7×10^2
DL-glyceraldehyde 3-phosphate	1.10	7.8×10^3	1.4×10^2

[a] from Dr. Liangbin Wang unpublished data

(k_{cat}/K_m = 6.4 × 10^3 $M^{-1}s^{-1}$) and its epimer 2-keto-3-deoxy-6-phosphogluconate (k_{cat}/K_m = 4.0 × 10^3 $M^{-1}s^{-1}$) supported this proposal. Other 6-carbon containing phosphates also have relatively high activities. Compared to 2-keto-3-deoxy 6-phosphogalactonate, the of α-D-mannose 1-phosphate displays a 2-fold lower (k_{cat}/K_m = 2.8 × 10^3 $M^{-1}s^{-1}$) that is similar to that of α-D-glucose 1-phosphate (k_{cat}/K_m = 1.5 × 10^3 $M^{-1}s^{-1}$), which is 4-fold less. The C-6 phospho-sugars, D-glucose 6-phosphate (k_{cat}/K_m = 2.9 × 10^2 $M^{-1}s^{-1}$) and D-fructose 6-phosphate (k_{cat}/K_m = 5.8 × 10^2 $M^{-1}s^{-1}$), are less active than the C (1) phospho-sugars.

Next, the C-5, C-4 and C-3 aldose/carboxylate phosphates were examined. The results show that these substrates have the highest activity (Table 4.2). 5-Phospho-D-arabinonate, D-erythrose 4-phosphate and 4-phospho-D-erythronate have k_{cat}/K_m values of about 10^4 $M^{-1}s^{-1}$. Arabinose 5-phosphate (k_{cat}/K_m = 6.5 × 10^3 $M^{-1}s^{-1}$) and its C(2)-OH epimer ribose 5-phosphate (k_{cat}/K_m = 8.1 × 10^2 $M^{-1}s^{-1}$) are less active. The activity of α-DL-glycerophosphate (k_{cat}/K_m = 2.7 × 10^2 $M^{-1}s^{-1}$) is 20-fold lower showing that it is not an active substrate.

The combined results suggest that the minimal unit for the leaving group of an active YidA substrate is a C-4 aldose or the corresponding carboxylate. Thus, the active site of this phosphatase is able to discriminate sugar phosphates based on interactions with the leaving group.

4.3.2 pH Rate Profile of Wild-Type YidA Catalyzed Hydrolysis of Ribose-5-Phosphate

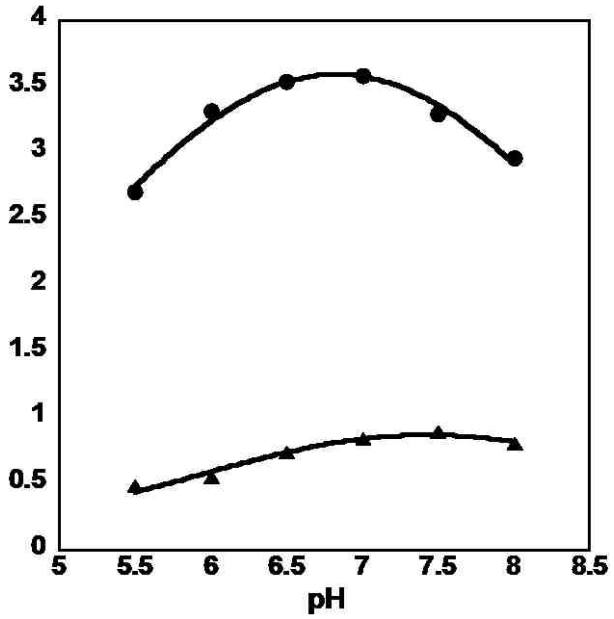


Figure 4.3 The log(k_{cat}/K_m) (●) or log(k_{cat}) (▲) pH profiles of wild-type YidA catalyzed hydrolysis of ribose-5-phosphate in H_2O buffer.

Generally, the variation of the rate of an enzyme catalyzed reaction with pH is caused by several factors. Protonation or deprotonation of enzyme side chains may induce a conformational changes that effects catalysis and/or the ionization state of a substrate or active site groups might influence the binding affinity or the rate at which the substrate is converted to products. Therefore, in order to characterize an enzyme, the pH profile needs to be determined. The pH profile of k_{cat} reflects how the ionization constants of enzyme-substrate complex ES (k_{cat} describes the first order rate constant for ES $\rightarrow$ E + P). In contrast, the pH dependence of k_{cat}/K_m monitors the ionization constants of the

uncomplexed enzyme and substrate (k_{cat}/K_m describes the second order rate constant for $E + S \rightarrow E + P$).

The pH dependence of wild-type YidA catalyzed hydrolysis of ribose-5-phosphate was determined by continuously monitoring the release of inorganic phosphate using the EnzChek assay. The pH profile measured for the k_{cat}/K_m of YidA with ribose-5-phosphate as a substrate has a bell-shape, which when fitted to equation 4.2 gives pK_a = 3.8 ± 0.3 and pK_b = 11.2 ± 1.6 (Figure 4.3). The k_{cat} pH profile of YidA hydrolysis of ribose-5-phosphate is also bell-shaped. The pK_a of 6.5 ± 0.1 and pK_b of 7.7 ± 0.4 were obtained by fitting the data to equation 4.2.

Like most enzymes, wild-type YidA functions optimally at physiological pH 7.4 (Figure 4.3) and it maintains activity over a pH range of 5.5 to 8.0, which is the range we used to measure the pH effects. Plots of log k_{cat} and log k_{cat}/K_m as a function of pH reveal plateaus with maximum values at neutral pH and that the values drop at acid and alkaline pH values. This observation suggests that one ionizable group loses its function upon protonation while another ionizable group loses its function upon deprotonation. In another words, one residue must be deprotonated for maximal enzymatic activity, while the other residue must be protonated for the maximal activity. A highly conserved Asp residue, which serves as a general acid/base for the catalytic reaction, in all phosphatase/phosphomutase branch of HADSF[(11-13)] except for ATPase, which has a Thr in the place of Asp[(14)]. In YidA, this highly conserved (D + 2) Asp (two residues downstream from the Asp nucleophile) is Asp11, which is deprotonated by the leaving group in the forming of the phosphoenzyme intermediate and protonated by a water molecule in the decomposition of aspartyl phosphate intermediate. The k_{cat}/K_m and k_{cat}

pH profiles indicate that an acid/base group must be ionized for catalytic turnover, which is consistent with the general mechanism of YidA.

4.3.3 Inhibition of Divalent Metal Ca^{2+} Ion

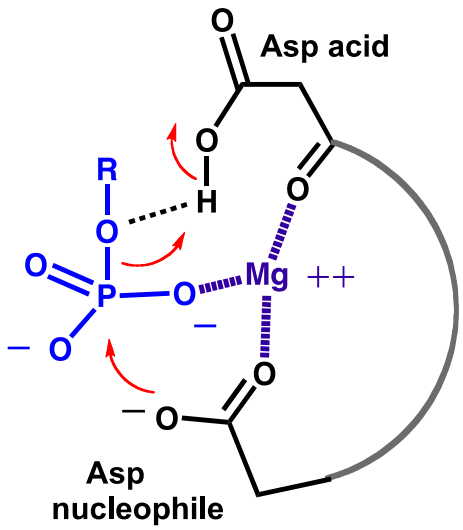


Figure 4.4 Schematic of the interactions between Mg^{2+} and active site residues of HADSF phosphatase.

Mg^{2+} is an important metal cofactor in many enzymes, including nucleases [15-18], synthetase [19], isomerase [20] and phosphatases[21, 22]. In these proteins, Mg^{2+} either binds to the substrate to form a magnesium-substrate complex or directly to the enzyme, resulting in alteration of its structure. In HADSF phosphatases, Mg^{2+} serves both of these roles (Figure 4.4). Like all phosphatases within the HAD superfamily, the presence of Mg^{2+} is essential for the YidA catalysis. The maximum activity of enzyme, measured by the release of Pi, was obtained when Mg^{2+} is present. In the absence of added divalent cation, the activity of YidA is only ~40% of the maximal activity observed in the

presence of Mg^{2+}(Table 4.3). When Mg^{2+} ion was removed by using dialysis against EDTA, the activity decreased to a half compared to the native enzyme (Table 4.3). While the addition of Mg^{2+} rescued partial activity, addition of Ca^{2+} caused a further activity loss (Table 4.3).

Table 4.3 The comparison of the apparent k_{cat} measured for wild type YidA and dialyzed YidA (D) in catalysis of 0.4 mM erythrose-4-phosphate hydrolysis in the presence and absence of added Mg^{2+} (1 mM) or Ca^{2+} (1 mM) in 50 mM Tris-HCl (pH 7.5) buffer containing 5 mM $MgCl_2$.

	YidA	YidA +Mg^{2+}	YidA(D)+Mg^{2+}	YidA (D)	Yida(D)+Ca^{2+}
k_{cat} (s^{-1})	3.1	7.8	2.1	1.5	0.6

To determine the amount of Mg^{2+} bound to the purified YidA before and after dialysis against EDTA, samples were analyzed by using Inductively Coupled Plasma (ICP) at the Earth & Planetary Science Department of UNM. The untreated YdiA was found to contain a ratio of Mg^{2+}: active site of 0.96:1.00, whereas the ratio of the EDTA treated enzyme was found to be 0.16:1.00. The majority fact that all Mg^{2+} was not removed might account for the residual activity observed for the enzyme after dialysis. The further loss of activity caused by the replacement of Mg^{2+} by Ca^{2+} raises the possibility that Ca^{2+} is an inhibitor, which disrupts the Mg^{2+}-YidA binding by having a stronger interaction with active site residues. The finding that Ca^{2+} inhibited the activity of YidA hydrolysis of erythrose-4-phosphate in the presence of Mg^{2+} enhances this possibility.

The YidA Ca^{2+} binding constant was determined by measuring the competitive inhibition

constant (K_i) vs Mg^{2+} using the steady-state initial velocity technique. The data (Figure 4.5) were fitted using equation 3.4 to define K_i = 3.1 ± 0.1 μM, which is much smaller than the Mg^{2+} binding constant (K_a = 16.6 ± 0.8 μM). Analysis of the structure of the active site of human phosphoserine phosphatase (HPSP) with Ca^{2+} bound, determined in 2004[23], provides a possible explanation of the tighter binding of and the inhibition by Ca^{2+} ion. The structure of HPSP showed that Ca^{2+} forms a sevenfold coordination complex with active site residues instead of the six fold coordination associated with Mg^{2+}-binding residues. Both side-chain oxygen atoms of the nucleophile Asp20 ligate with Ca^{2+} ion, while Mg^{2+} is only bound to one oxygen. The bidentate binding of Asp20 to Ca^{2+} hinders its key nucleophilic activity in catalysis.

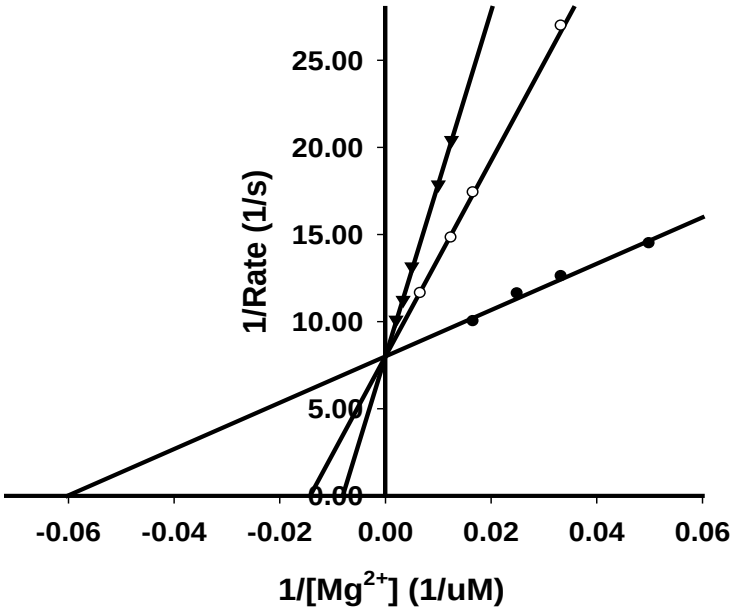


Figure 4.5 Lineweaver-Burk plot of the Ca^{2+} vs Mg^{2+} inhibition data measured for YidA catalyzed hydrolysis of erythrose-4-phosphate in 50 mM Hepes (pH 7.5, 25 ^{o}C). The concentration of enzyme was fixed at 0.02 μM, the Ca^{2+} concentration varied from 0 to

20 μM, and the Mg^{2+} concentration varied from 0.01 to 1 mM. (●) 0 μM Ca^{2+}; (○) 10 μM Ca^{2+}; (▼) 20 μM Ca^{2+}.

4.3.4 Inhibition by the Phosphate Analog WO_4^{3-}

As a transition state analog, tungstate (WO_4^{3-}) has been used to probe reaction mechanisms of a series of enzymes that catalyze cleavage of P-O bonds[2, 24-28]. The mechanism(s) by which phosphatases are inhibited by tungstate is not clear. One explanation is that tungstate is bound in the enzyme active site with a structure resembling the transition state for the phosphoryl transfer reaction[24] (Figure 4.6).

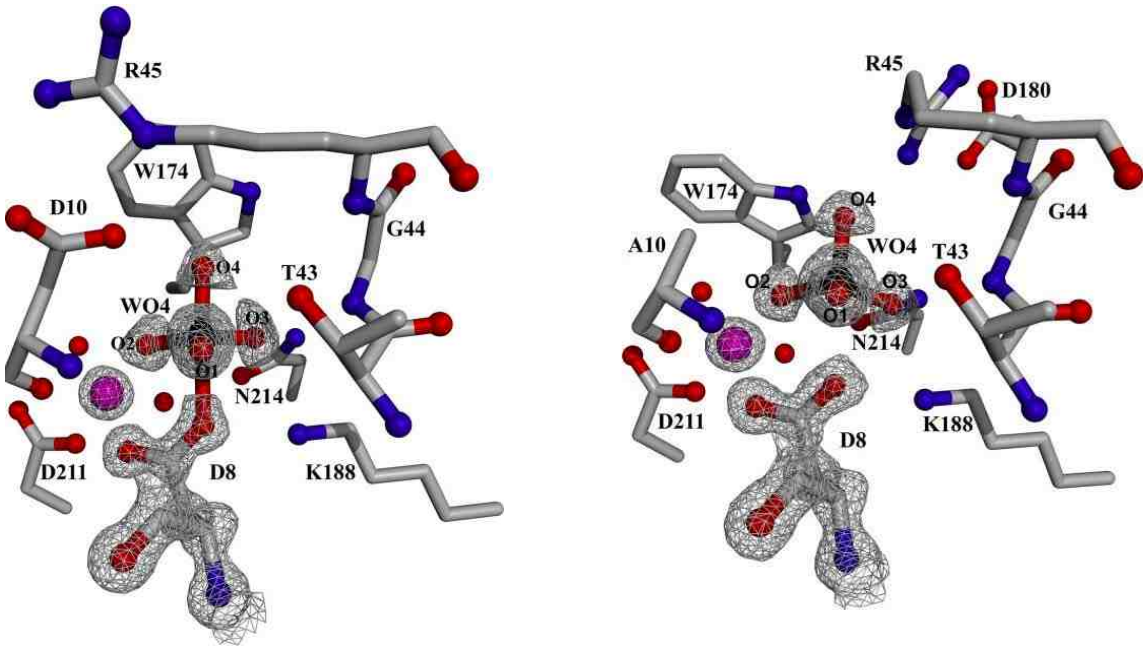


Figure 4.6 (A) The hexose phosphate phosphatase BT4131 from *Bacteroides thetaiotaomicron* wild type and **(B)** D10A variant in the presence of phosphate mimics WO_4^{3-} and the cofactor Mg^{2+} (magenta sphere). *from ref.* [24].

Tungstate was observed to be a competitive inhibitor for YidA hydrolysis of ribose-5-phosphate with a Ki = 3.7 ± 0.2 μM (Figure 4.7). The small K_i value means tungstate binds tightly to YidA and that it might, therefore, serve as a good ligand for co-crystalization. This would provide an ideal model to demonstrate interactions that take place between YidA active-site residues and the transition state formed along the reaction coordinate.

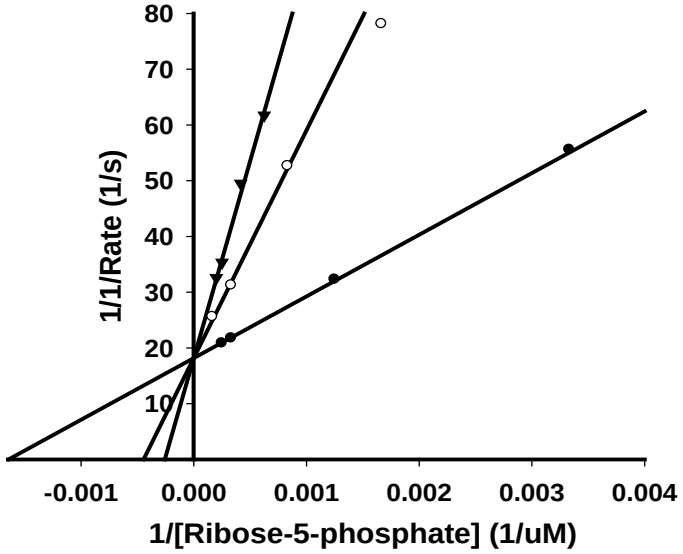


Figure 4.7 Lineweaver-Burk plot of the tungstate (0, 10 and 20 μM) inhibited YidA catalyzed ribose-5-phosphate (0.3 to 6 mM) hydrolysis in 50 mM Hepes (pH 7.5, 25 ^{o}C) containing 5 mM $MgCl_2$. (●) 0 μM WO_4^{3+}; (○) 10 μM WO_4^{3+}; (▼) 20 μM WO_4^{3+}.

4.3.5 YidA Site-Directed Mutagenesis

4.3.5.1 The Preserved Residue Asp12 Plays an Essential Role in Catalysis

As mentioned in Chapter one, all HADSF phosphatases and phosphate mutases have a

highly conserved core domain[29]. It is believed that the presence of this conserved core domain makes this enzyme family highly evolvable. The catalytic core domain, which has a robust Rossmann-fold, facilitates new function acquisition because it is able to enforce the conservation of the catalytic scaffold while accommodating the charged catalytic residues and the cap domain. The core domain binds the Mg^{2+} cofactor and the organophosphate reactant in a favorable orientation that allows attack by the Asp nucleophile and general acid/base catalysis by the (Asp+2) Asp positioned two residues upstream[30]. In YidA the conserved nucleophile is Asp10, and Asp12 plays the two-fold general acid/base functional role. Asp12 acts first as an acid to transfer a proton to the oxygen atom of the leaving group and then it acts as a base to abstract a proton from the water molecule (Figure 4.8).

Figure 4.8 The general catalytic mechanism of phosphatase members of the HADSF. The reaction proceeds through an aspartylphosphate intermediate.

The structure of wild type YidA docking with erythrose-4-phosphate (Figure 4.9) suggests that Asp12 protonates the C(4)O of the substrate during the enzyme phosphorylation step and forms a hydrogen bond to the leaving group oxygen or attacking nucleophile stabilizing the trigonal-bipyramidal intermediate structure during

Table 4.4 The steady-state kinetic constants k_{cat} and K_m for YidA mutations hydrolyzing erythrose-4-phosphate are measured by *EnzChek* Phosphate Assay Kit in 50 mM Hepes buffer containing 5mM $MgCl_2$, pH 7.5, 25 ^oC.

Protein	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
Wild Type YidA	$(7.5 \pm 0.6) \times 10^2$	10.2 ± 0.5	1.4×10^4
D12N	$(3.0 \pm 0.1) \times 10^2$	$(6.5 \pm 0.3) \times 10^{-3}$	2.0×10^1
D12A	NA[a]	------	------
Y187A	$(8.0 \pm 0.2) \times 10^1$	$(2.2 \pm 0.1) \times 10^{-1}$	2.8×10^3
E190A	$(1.4 \pm 0.1) \times 10^2$	$(4.6 \pm 0.1) \times 10^{-1}$	3.4×10^2
E130A	$(2.6 \pm 0.2) \times 10^2$	$(8.0 \pm 0.3) \times 10^{-2}$	3.1×10^2
S184A	$(3.6 \pm 0.2) \times 10^2$	1.6 ± 0.1	4.4×10^3
H129A	$(4.6 \pm 0.4) \times 10^2$	7.4 ± 0.2	1.6×10^4
M156A	$(1.5 \pm 0.2) \times 10^3$	12.5 ± 0.6	8.2×10^3
K154A	$(4.2 \pm 0.3) \times 10^2$	$(7.8 \pm 0.3) \times 10^{-2}$	1.9×10^2
F188A	$(1.1 \pm 0.1) \times 10^3$	5.0 ± 0.2	4.5×10^3
L111A	$(4.8 \pm 0.3) \times 10^2$	8.5 ± 0.3	1.8×10^4
R46A	$(4.0 \pm 0.4) \times 10^2$	$(1.6 \pm 0.1) \times 10^{-2}$	4.0×10^1

[a]No activity detected for a reaction solution containing 5 μM YidA D11A and 500 μM erythrose-4-phosphate and incubated for 25 min. The detection limit under these conditions is a turnover rate of 1.6×10^{-2} s^{-1}.

catalysis. In order to measure its contribution to the hydrolysis rate, Asp12 was replaced by Ala and Asn by using site-directed mutagenesis. The methyl side chain of the D12A mutant cannot interact with the substrate, thus eliminating the acid/base catalysis and

inactivating the enzyme (Table 4.4). In principle, the amide side chain of the D12N mutant could bind to the substrate and stabilize the intermediate via hydrogen bond formation.

As the data in Table 4.4 show, the apparent dissociation constant of D12N is comparable to the K_m of the wild type enzyme. In contrast, the turnover number (k_{cat}) of the D12N mutant decreased ~1500-fold.

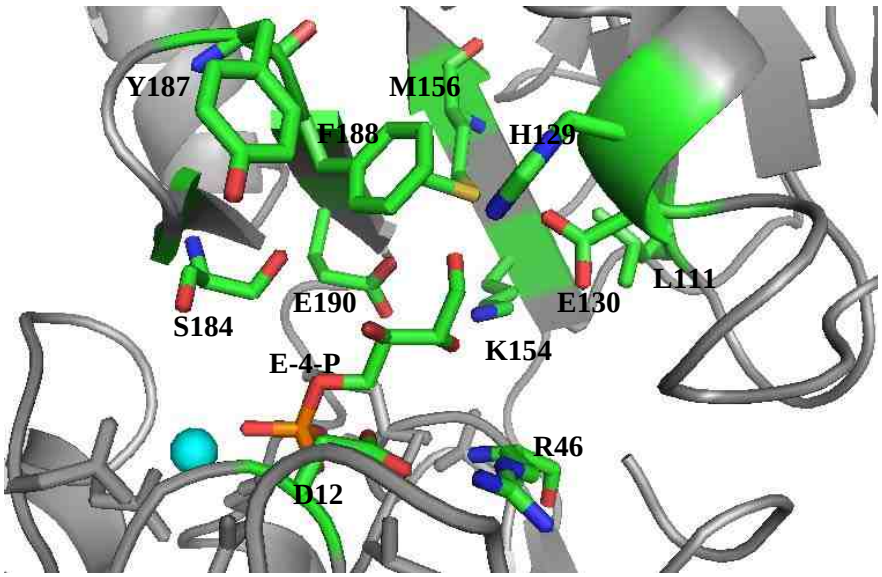


Figure 4.9 A Pymol generated structure of the hydrogen bonding patterns between the active sites of YidA and erythrose 4-phosphate. The Mg^{2+} is shown as teal sphere.

4.3.5.2 Determination of the Substrate Recognition Residues

Many enzymes, including acetyl cholinesterase, hydroxynitrile lyase, enolase and HADSF, have catalytic promiscuity[(31)]. These proteins show high activities towards their physiological substrates, as well as relatively low activities towards nonphysiological

substrates. The emergence of these enzymes is proposed to involve modification of existing genes to generate new protein structures and functions that are related to those of their ancestors. Therefore, all enzymes from a common ancestor, which belong to a same superfamily, share the active site catalytic residues and the catalytic mechanism [(32)]. Their promiscuity is linked to conformational diversity, i.e., conformational changes enable the same enzyme to accommodate different substrates[(33)].

HADSF phosphatases are typical of promiscuous, with the vast majority having a conserved catalytic core domain and a tethered cap domain. The catalytic scaffold binds the transferring phosphoryl group, whereas the cap domain binds the leaving group. Thus, HADSF phosphatase catalytic residues positioned are physically separated from the substrate recognition residues positioned on the cap domain. As a result, replacement of substrate recognition residues should switch specificity from one substrate to another, but it should not greatly perturb the environment of the catalytic residues. The cap domain with a large surface can accommodate numerous substrate recognition residues. However, only a few residues participate in binding the physiological substrate. In order to explore this issue, a series of site-directed mutants of YidA were prepared to explore the location of critical substrate-binding residues in its cap domain.

According to the substrate specificity of YidA (Table 4.2) and the location of its gene, two substrates, erythrose-4-phosphate and 2-keto-3-deoxy-6-phosphogalactonate, were chose to dock into the apo YidA crystal structure 1RKQ (Figure 4.9 and 4.10). Analysis of the structures show that 10 residues, Arg46, Leu111, His129, Glu130, Lys154, Met156, Ser184, Tyr187, Phe188 and Glu190, are located near the substrate-binding site and may be involved in interactions with substrate. The individual contribution of each of these

Table 4.5 The steady-state kinetic constants k_{cat} and K_m for YidA mutants promoted hydrolysis of 2-keto-3-deoxy-6-phosphogluconate (KDPG), measured by using *EnzChek* Phosphate Assay Kit in 50 mM Hepes buffer containing 5mM $MgCl_2$, pH 7.5, 25 ^oC.

Protein	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
Wild Type YidA	$(0.8 \pm 0.1) \times 10^3$	3.2 ± 0.3	4.0×10^3
Y187A	$(1.9 \pm 0.1) \times 10^3$	5.8 ± 0.1	3.0×10^3
E190A	$(2.8 \pm 0.2) \times 10^3$	1.7 ± 0.1	6.2×10^2
E130A	$(1.4 \pm 0.1) \times 10^3$	$(9.0 \pm 0.4) \times 10^{-1}$	7.2×10^2
S184A	$(3.1 \pm 0.2) \times 10^3$	1.3 ± 0.1	4.2×10^2
H129A	$(3.0 \pm 0.3) \times 10^3$	5.7 ± 0.4	1.9×10^2
M156A	$(1.6 \pm 0.1) \times 10^3$	$(9.0 \pm 0.4) \times 10^{-1}$	5.6×10^2
K154A	$(5.2 \pm 0.3) \times 10^2$	$(4.4 \pm 0.1) \times 10^{-2}$	8.5×10^1
F188A	NA[a]	---	---
L111A	$(5.5 \pm 0.3) \times 10^3$	1.4 ± 0.1	2.5×10^2
R46A	$(9.5 \pm 0.5) \times 10^2$	$(3.9 \pm 0.1) \times 10^{-3}$	4

[a]No activity detected for a reaction solution containing 2.32 μM YidA F188A and 800 μM KDPG and incubated for 25 min. The detection limit under these conditions is a turnover rate of 1.6×10^{-2} s^{-1}.

residues to substrate binding was evaluated by using appropriate site directed mutants and steady-state kinetic analysis of the mutant enzymes for catalysis of erythrose-4-phosphate and 2-keto-3-deoxy-6-phosphogluconate (KDPG) hydrolysis (Table 4.4 and 4.5).

Arg46 is conserved among HADSF members, acting in catalysis by forming a hydrogen bond to the general acid/base Asp12 that stabilizes the aspartylphosphate intermediate [(24)]. Therefore, it is not surprising that replacement of Arg46 by Ala results in a dramatic activity reduction for both erythrose-4-phosphate (~300 fold decrease) and for KDPG (~1000 fold decrease). The results show that Tyr187 does not play a significant role in catalysis activity, based on the observation that Y187A retains the activity as the wild type YidA. In accord with this observation is the analysis of the model (Figure 4.9) that shows that Try187 is located a little far away from the binding substrate (~7.6 Å between its side chain oxygen and erythrose-4-phosphate hydroxyl group).

The substrate specificity screen results (Table 4.2) shows that an active YidA substrate should process a 4,5 or 6-carbon aldose or corresponding carboxylate leaving group. These leaving groups have several polar functional groups and, thus, it is not surprising that enzyme cap residues that have charged polar side chains, eg., Glu130, Glu190 and Lys154, are very important in positioning the substrate. Consequently, Ala substitutions for these residues cause a ~100 fold activity reduction with erythrose-4-phosphate and ~10 fold decrease with KDPG.

Ala replacement of His129, which also has a charged polar side chain, causes a large effect on the activity towards KDPG (20-fold decrease) and a subtle effect on hydrolysis of erythrose-4-phosphate. The model of 2-keto-3-deoxygalactonate 6-phosphate docked in the active-site cage generated for YidA (PDB: 1RKQ) (Figure 4.10) shows that the C6-

phosphate moiety is positioned proximal to the aspartate nucleophile, Asp10, bringing the carboxylate moiety of the docked ligand within hydrogen-bond distance of substrate-specificity determinant His129.

Replacements of other residues, including Phe188, Met156, Leu111 and Ser184, have an effect on the rate of six carbon KDPG hydrolysis and but do not greatly change the catalytic activity of four carbon substrates. The hydrophobic interactions between Phe188 benzyl ring and KDPG carbon chain seems to be a key factor in determining how the C-6 carboxylate unit is accommodated by the binding site.

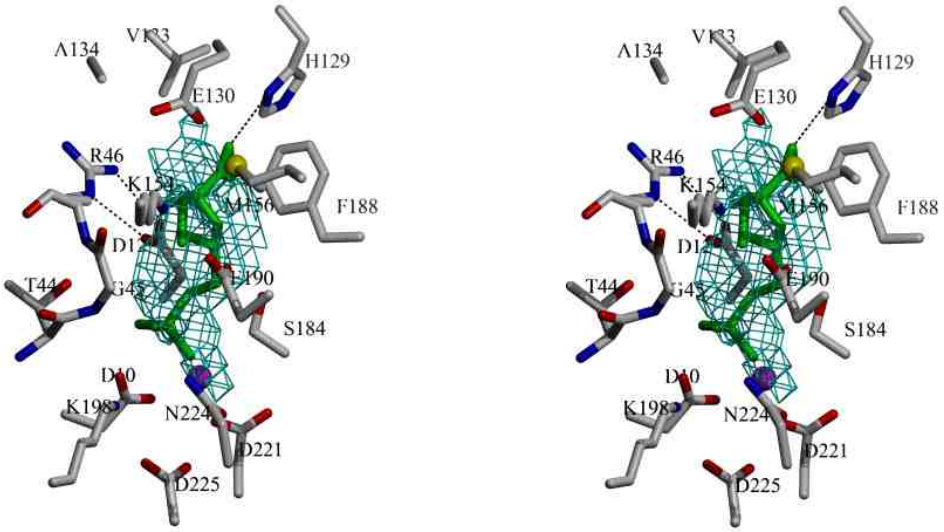


Figure 4.10 Stereo view of the model 2-keto-3-deoxygalactonate 6-phosphate (green) docked in the solvent accessible cage (cyan) calculated for the structure of YidA.

4.4 Conclusion

HADSF is an ideal enzyme superfamily to formulate and test theories of enzyme superfamily evolution because of the structural and functional diversity that are inherent

in its enormous size. In this chapter, a C2 type phosphatase YidA was chosen to examine the cap domain substrate recognition mechanism and from this knowledge to identify the structural determinants that govern the expansion of the HADSF.

The substrate screening results show that YidA is an effective catalyst for the hydrolysis of several known sugar phosphate metabolites (arabinose-5-phosphate, erythrose-4-phosphate, mannnose-1-phosphate and glucose 1-phosphate). The minimal unit for the leaving group of an active YidA substrate is a C-4 aldose or its corresponding carboxylate derivative. The cap residues Glu130, Glu190 and Lys154 play important roles in substrates recognition and binding. His129 and F188 are critical in binding of the larger (C-6) sugar phosphates.

The YidA gene is located near the *dgo* gene cluster, but its low activity (k_{cat}/K_m = 1.4 x 10^2 $M^{-1}s^{-1}$) towards the pathway product, glycelaldehyde-3-phosphate, excludes the possibility that YidA functions to hydrolize this phosphate. The observation that the phosphorylated intermediate (k_{cat}/K_m = 6.4 x 10^3 $M^{-1}s^{-1}$) 2-keto-3-deoxy 6-phosphogalactonate has moderate activity suggests that YidA might play the role of a housekeeper to remove accumulated phosphorylated metabolites, which can cause in catabolite inhibition.

Reference

1. Collet J-F, G. I., Rider MH, Veiga-da-Cunha M, Van Schaftingen E. (1997) Human-3-phosphoserine phosphatase: sequence, expression and evidence for a phosphoenzyme intermediate, *FEBS Lett 408*.
2. Lu, Z., Wang, L., Dunaway-Mariano, D., and Allen, K. N. (2009) Structure-function analysis of 2-keto-3-deoxy-D-glycero-D-galactononononate-9-phosphate phosphatase defines specificity elements in type C0 haloalkanoate dehalogenase family members, *The Journal of biological chemistry 284*, 1224-1233.
3. Biswas, T., Yi, L., Aggarwal, P., Wu, J., Rubin, J. R., Stuckey, J. A., Woodard, R. W., and Tsodikov, O. V. (2009) The tail of KdsC: conformational changes control the activity of a haloacid dehalogenase superfamily phosphatase, *The Journal of biological chemistry 284*, 30594-30603.
4. Gallinaro, L., Crovatto, K., Rampazzo, C., Pontarin, G., Ferraro, P., Milanesi, E., Reichard, P., and Bianchi, V. (2002) Human mitochondrial 5'-deoxyribonucleotidase. Overproduction in cultured cells and functional aspects, *The Journal of biological chemistry 277*, 35080-35087.
5. Neel, B. G., and Tonks, N. K. (1997) Protein tyrosine phosphatases in signal transduction, *Current opinion in cell biology 9*, 193-204.
6. O'Hara, L., Han, G. S., Peak-Chew, S., Grimsey, N., Carman, G. M., and Siniossoglou, S. (2006) Control of phospholipid synthesis by phosphorylation of the yeast lipin Pah1p/Smp2p Mg2+-dependent phosphatidate phosphatase, *The Journal of biological chemistry 281*, 34537-34548.

7. Passariello, C., Forleo, C., Micheli, V., Schippa, S., Leone, R., Mangani, S., Thaller, M. C., and Rossolini, G. M. (2006) Biochemical characterization of the class B acid phosphatase (AphA) of Escherichia coli MG1655, *Biochimica et biophysica acta 1764*, 13-19.

8. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J.-C., Williams, K. L., Appel, R. D., and Hochstrasser, D. F. . (1998) Protein Identification and Analysis Tools in the ExPASy Server, in 2-D Proteome Analysis Protocols (Link, A. J., Ed.). , *Humana Press, Totowa, NJ*.

9. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Analytical biochemistry 72*, 248-254.

10. Cleland, W. W. (1979) Statistical analysis of enzyme kinetic data, *Methods Enzymol 63*, 103-138.

11. Argiriadi, M. A., Morisseau, C., Hammock, B. D., and Christianson, D. W. (1999) Detoxification of environmental mutagens and carcinogens: Structure, mechanism, and evolution of liver epoxide hydrolase, *P Natl Acad Sci USA 96*, 10637-10642.

12. Parsons, J. F., Lim, K., Tempczyk, A., Krajewski, W., Eisenstein, E., and Herzberg, O. (2002) From structure to function: YrbI from Haemophilus influenzae (HI1679) is a phosphatase, *Proteins 46*, 393-404.

13. Dai, J. Y., Finci, L., Zhang, C. C., Lahiri, S., Zhang, G. F., Peisach, E., Allen, K. N., and Dunaway-Mariano, D. (2009) Analysis of the Structural Determinants Underlying Discrimination between Substrate and Solvent in beta-Phosphoglucomutase Catalysis, *Biochemistry-Us 48*, 1984-1995.

14. Toyoshima, C., Nakasako, M., Nomura, H., and Ogawa, H. (2000) Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 angstrom resolution, *Nature 405*, 647-655.

15. Jeltsch, A., Alves, J., Wolfes, H., Maass, G., and Pingoud, A. (1993) Substrate-Assisted Catalysis in the Cleavage of DNA by the Ecori and Ecorv Restriction Enzymes, *P Natl Acad Sci USA 90*, 8499-8503.

16. Baldwin, G. S., Vipond, I. B., and Halford, S. E. (1995) Rapid reaction analysis of the catalytic cycle of the EcoRV restriction endonuclease, *Biochemistry-Us 34*, 705-714.

17. Black, C. B., and Cowan, J. A. (1998) A critical evaluation of metal-promoted Klenow 3 '-5 ' exonuclease activity: calorimetric and kinetic analyses support a one-metal-ion mechanism, *J Biol Inorg Chem 3*, 292-299.

18. Vipond, I. B., Baldwin, G. S., and Halford, S. E. (1995) Divalent metal ions at the active sites of the EcoRV and EcoRI restriction endonucleases, *Biochemistry-Us 34*, 697-704.

19. Liaw, S. H., and Eisenberg, D. (1994) Structural model for the reaction mechanism of glutamine synthetase, based on five crystal structures of enzyme-substrate complexes, *Biochemistry-Us 33*, 675-681.

20. Vanbastelaere, P. B. M., Kerstershilderson, H. L. M., and Lambeir, A. M. (1995) Wild-Type and Mutant D-Xylose Isomerase from Actinoplanes-Missouriensis - Metal-Ion Dissociation-Constants, Kinetic-Parameters of Deuterated and Non-Deuterated Substrates and Solvent-Isotope Effects, *Biochem J 307*, 135-142.

21. Fortpied, J., Maliekal, P., Vertommen, D., and Van Schaftingen, E. (2006) Magnesium-dependent phosphatase-1 is a protein-fructosamine-6-phosphatase potentially involved in glycation repair, *The Journal of biological chemistry 281*, 18378-18385.

22. Anderson, R. A., Bosron, W. F., Kennedy, F. S., and Vallee, B. L. (1975) Role of Magnesium in Escherichia-Coli Alkaline-Phosphatase, *P Natl Acad Sci USA 72*, 2989-2993.

23. Peeraer, Y., Rabijns, A., Collet, J. F., Van Schaftingen, E., and De Ranter, C. (2004) How calcium inhibits the magnesium-dependent enzyme human phosphoserine phosphatase, *Eur J Biochem 271*, 3421-3427.

24. Lu, Z., Dunaway-Mariano, D., and Allen, K. N. (2008) The catalytic scaffold of the haloalkanoic acid dehalogenase enzyme superfamily acts as a mold for the trigonal bipyramidal transition state, *P Natl Acad Sci USA 105*, 5687-5692.

25. Stankiewicz, P. J., and Gresser, M. J. (1988) Inhibition of Phosphatase and Sulfatase by Transition-State Analogs, *Biochemistry-Us 27*, 206-212.

26. Foster, J. D., Young, S. E., Brandt, T. D., and Nordlie, R. C. (1998) Tungstate: A potent inhibitor of multifunctional glucose-6-phosphatase, *Arch Biochem Biophys 354*, 125-132.

27. Davies, D. R., Interthal, H., Champoux, J. J., and Hol, W. G. J. (2002) Insights into substrate binding and catalytic mechanism of human tyrosyl-DNA phosphodiesterase (Tdp1) from vanadate and tungstate-inhibited structures, *J Mol Biol 324*, 917-932.

28. Wagner, U. G., Stupperich, E., and Kratky, C. (2000) Structure of the molybdate/tungstate binding protein mop from Sporomusa ovata, *Structure 8*, 1127-1136.

29. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D., and Aravind, L. (2006) Evolutionary genomics of the HAD superfamily: Understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes, *J Mol Biol 361*, 1003-1034.

30. Allen, K. N., and Dunaway-Mariano, D. (2004) Phosphoryl group transfer: evolution of a catalytic scaffold, *Trends in biochemical sciences 29*, 495-503.

31. O'Brien, P. J., and Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities, *Chem Biol 6*, R91-R105.

32. Khersonsky, O., Roodveldt, C., and Tawfik, D. S. (2006) Enzyme promiscuity: evolutionary and mechanistic aspects, *Curr Opin Chem Biol 10*, 498-508.

33. James, L. C., and Tawfik, D. S. (2003) Conformational diversity and protein evolution--a 60-year-old hypothesis revisited, *Trends in biochemical sciences 28*, 361-368.

CHAPTER FIVE

5. Biological Function Assignment of HADSF Phosphatases

5.1 Introduction

As the genome sequencing technology has become more sophisticated, more and more protein sequences have been deposited in the database. Approximately 200,000,000 sequence records exist in the GenBank as of April 2011. The abundance of protein sequences help to define the metabolism and physiology of an organism and to understand the functions of proteins. However, experimental characterization has been performed on only a small fraction of the proteins. To cope with the dramatic increase in sequences, computational predictions are required to auto-annotate the individual functions of proteins. The computational method is based on the assumption that proteins with similar sequences will have similar functions. However, the reliability of this functional assignment technique has become questionable. Brenner[(1)] compared the functional annotation of three different groups for the *Mycoplasma genitalium* genome and found at least 8% for the 340 annotated genes were incorrect. Later, Devos and Valencia[(2)] examined the functional assignment in the first three published genomes, *H.influenzae, M. genitalium* and *Methanococcus jannaschii,* and suggested that the misannotation level for different types of function was as high as 40%. Recently, Babbitt[(3)] investigated the misannotation levels of 37 enzyme families in four public protein sequence databases (UniProtKB/Swiss-Prot, GenBank NR, UniProtKB/TrEMBL, and KEGG), for which extensive experimental information exists, Babbitt concluded that 40% of the sequences deposited in these four databases as recently as 2005 were

misannotated.

Therefore, a more reliable method for protein biofunction assignment needs to be created. The Enzyme function initiative (EFI)[4], which is a large scale collaborate project aimed at determining the functions of unknown proteins using a multidisciplinary strategy, was created recently. The EFI's strategy for enzyme functional assignment can be summarized in three steps. First, a computational prediction based high throughput screening is employed to narrow the substrate specificity of a protein. Subsequent experimental enzymological studies are used to test the predicted function of the protein. Finally, in vivo studies are conducted to evaluate the assigned function.

The EFI is composed of six scientific cores, including the superfamily/genome core, protein core, structure core, computation core, microbiology core and data/dissemination core. The initiative focuses on five superfamilies including the amidohydrolases (AH), enolases (EN), glutathione transferases (GST), haloalkanoic acid dehalogenases (HAD) and isoprenoid synthases (IS)]. Members of these five superfamilies are functionally diverse and have mechanistically conserved chemical reactions or chemical capabilities while having divergent overall functions.

In this chapter, I describe preliminary results arising from the EFI effort that have come from an approach that combines experimental structural biology and computation-based experimental enzymology to define the function of five members of the HAD superfamily.

As mentioned above, the HADSF was chosen because of its substrate diversity. The majority of the enzymes in this superfamily promote phosphoryl transfer reactions, such as the phosphate monoester hydrolases (phosphatases) or phosphoanhydride hydrolases

P-type (ATPases), phosphonoacetaldehyde hydrolase (phosphonatase) and phosphomutases (phosphoglucomutase and phosphomannomutase). Each of the HAD phosphotransferases requires a Mg^{2+} cofactor for catalysis[5, 6]. Although the sequence identity among HADSF members is very low (less than 15%)[7], they are characterized by containing four highly conserved catalytic sequence motifs[8]. The highly conserved Asp residue in the HADSF, located on loop one, acts as nucleophile catalytic group and contributes to the adaptability of the catalytic scaffold. Mg^{2+} binding to the nucleophilic Asp and phosphorylated substrate provides the required orientation and charge shielding for nucleophilic attack[9]. Two residues downstream from the Asp nucleophile is located another highly conserved amino acid (Figure 5.1). In phosphomonoesterases or phosphomutases, this position is occupied by a protonated/deprotonated Asp/Glu, acting as the acid/ base catalyst. In the ATPases, the occurrence of a Thr at this position allows for a reduced rate of aspartyl phosphate hydrolysis, which may result in a time lag necessary for a subsequent conformational change. In the phosphonatases, the space saving residue Ala instead of the second aspartate exists at this location, a phenomenon that is consistent with the unique role played by the Schiff-base intermediate (formed with the insert domain) as a general acid-base catalyst in aspartyl phosphate hydrolysis by these proteins[10].

The cap domain of HADs, which is the substrate recognizing part, is believed to be responsible for the diversity of the members of the superfamily. According to the location and the presence of the caps, HADSF phosphatases can be classified into three categories. The C1 types are those that contain inserts occurring in the middle of the β-hairpin of the flap motif, and fold into a structural unit distinct from the core domain. In

the C2 subclass, inserts occur in the linker immediately after strand S3. C0 class are structurally the simplest representatives of the HAD superfamily and only have small inserts in either of the two points of cap insertion. The cap domain, together with the squiggle and the flap in the core domain, is likely to be responsible for solvent exclusion and substrate access to the active site of the enzyme.

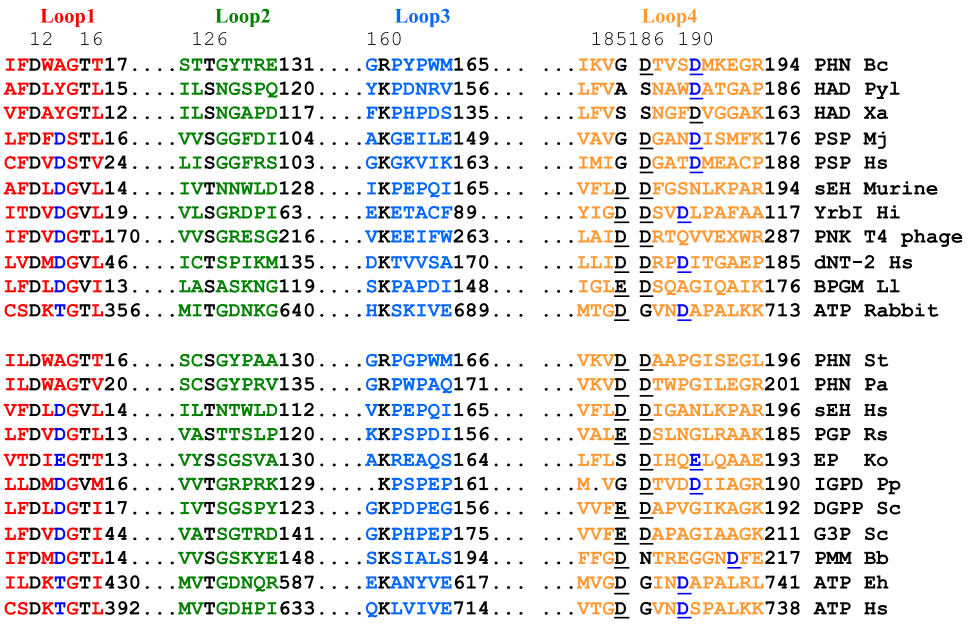


Figure 5.1 Sequence alignment of the core domain in different HAD proteins. In each loop, the residues identified by computer analysis of primary sequence are colored black and the residues identified by structure-function analysis are colored dark blue (loops 1 and 4, will be discussed next), in loop 4, the metal ion pair is underlined.

The issue of diversified functional assignment within the HAD superfamily can be best addressed through comparisons of probable orthologs. Four HADSF phosphatase orthologs from different species of bacteria were chosen based on high throughput

screening (HTPS) results, which show that all of these proteins have high a catalytical specificity towards riboflavin 5'-monophoaphate (FMN). The four proteins are YigB (EFI-501262) and YbjI(501335) from *Escherichia coli*, BT2542 (501088) from *Bacteroides thetaiotaomicron* and Q9RUP0 (501193) from *Deinococcus radiodurans*.

5.2 Experimental Methods

5.2.1 Material

All chemicals were obtained from Sigma-Aldrich except where mentioned. Primers and T4 DNA ligase were purchased from Invitrogen. Restriction enzymes were purchased from BioLab. Deepvent polymerase, *Pfu* Turbo polymerase, pET14b vector were purchased from Stratagene. The Geneclean Spin Kit and the Qiaprep Spin Miniprep Kit were purchased from Qiagen. Host cells were purchased from Novagen. Genomic DNA encoding YigB from *Escherichia coli* were purchased from ATCC (ATCC#: 10798).

5.2.2 Protein Cloning, Expression and Purification

5.2.2.1 Cloning, Expression, and Purification of YigB

The cDNA encoding the gene HAD protein YigB (NCBI accession U00096, Swiss Pro accession: P0ADP0) from *E coli* was amplified by PCR using the genomic DNA from *E. coli* K12 *(*ATCC 10798) and P*fu* DNA polymerase. The PCR product was cloned into the PET-14b vector, which was used to transform competent *E. coli.* BL21 (DE3) cells. Oligonucleotide primers (5'-TCCACGCGCCATATGGGGGAAATAATGCGTTTT) and (5'-GGGTATATATAAAGCTTTGCTGATTATATCAGCGAGG) containing restriction endo- nuclease cleavage sites *NdeI* and *HindIII* were used in the PCR reactions. Plasmid

DNA was purified using a Qiaprep Spin Miniprep Kit. The gene sequencing carried out at the Center for Genetics in Medicine, University of New Mexico School of Medicine.

The *E. coli.* BL21 (DE3) cells transformed by pET14b construct were grown at 37 °C with agitation at 200 rpm in Luria broth (LB) containing 50 μg/mL ampicillin for 4-6 h to an OD_{600} = 0.6-0.8 and induced for 12 h at 20 °C at a final concentration of 0.4 mM isopropyl α-D-thiogalactopyranoside (IPTG). The cells were harvested by centrifugation at 6500 rpm for 15 min at 4 °C to yield 3 g/L of culture medium. And then the cells were resuspended in lysis buffer (50 mM HEPES, 500 mM NaCl, 10 mM imidazole, 5mM $MgCl_2$, pH=7.5) at 1 g wet cells/10 ml buffer and lysed by 2 passages through a French Press cell (12,000 psi). The cell debris was removed by centrifugation at 20,000 rpm for 30 min at 4 °C, and the supernatant containing YigB loaded onto a Ni-NTA column at 4 °C for His_6-tagged protein binding. The column was washed with 200 ml of lysis buffer and 100 ml of wash buffer (50 mM HEPES, 500 mM NaCl, 50 mM imidazole, 5mM $MgCl_2$, pH=7.5) before YigB was eluted using 100 ml of elution buffer (50 mM HEPES, 500 mM NaCl, 250 mM imidazole, 5mM $MgCl_2$, pH 7.5). The protein solution was concentrated by using a 10 kDa Macrosep centricon (Pall Filtron) device. The imidazole was removed by dialysis in 6 L of 50 mM K^+HEPES containing 500 mM NaCl and 5mM $MgCl_2$ (pH 7.5, 4 °C). The protein purity was verified by SDS-PAGE analysis. The protein concentration was determined by the Bradford method. The yield of the protein YigB is 2 mg/ g wet cells.

The N-terminal His_6-tag was removed by incubating 50 mM YigB with 1 units/ml thrombin in 50 mM Tris-HCl (pH 7.5) containing 500 mM NaCl and 5mM $MgCl_2$ on ice

for 12 h. The resulting solution was passed through the Ni-NTA column (equilibrated with the same incubating buffer) and the flow-through fractions were collected.

5.2.2.2 Expression, and Purification of BT2542

The CHS30 construct containing the gene encoding BT2542, which was from EFI protein core, was used to transform the *E. coli.* BL21 (DE3) cells. The transformed cells were grown and harvested using the same procedure mentioned above. And then the cell pellet was suspended in 1g wet cell/10 mL of ice-cold lysis buffer (20 mM Tris•HCl, 40 mM imidazole, 1 mM DTT and 500 mM NaCl, pH 7.5). The cell suspension was passed through a French press at 1,200 PSIG before centrifugation at 20,000 rpm and 4°C for 30 min. The supernatant was loaded onto a 5ml HisTrap FF crude column (GE Healthcare) at a flow rate of 5 mL/min. The column was washed with 5 column volumn (CV) lysis buffer followed by 5CV of wash buffer (20 mM Tris•HCl, 500 mM NaCl, 500 mM imidazole and 1mM DTT at pH 7.5). Fractions containing target protein, as judged by SDS-PAGE analysis, were combined, concentrated with a 10K Amicon Ultra Centrifugal filter (Millipore), and then stored at -80 °C.

5.2.2.3 BT2542 Site-Directed Mutagenesis

YidA site-directed mutagenesis was prepared by a PCR based strategy with the CHS30-YidA clone serving as template and commercial oligonucleotides as primers (as listed in Table 5.1). The purified clone was used to transform competent *E. coli* BL21 (DE3) cells after confirming the gene sequence by using commercial DNA sequencing at MC lab. The transformed cells were grown and harvested using the same procedure as the wild type BT2542. And the variant proteins were purified to homogeneity (as judged by S.D.S gel) using the method mentioned in Section 5.2.2.2.

5.2.3 Size exclusion chromatography

SEC was performed on ÄKTA P-920 FPLC system and Hiprep 16/60 Sephacryl S-200 HR column (GE Healthcare). The size exclusion column was equilibrated with 50 mM HEPES (pH 7.5), 500 mM NaCl, and 1 mM DTT. The column was calibrated using the gel-filtration low molecular weight and high molecular weight gel filtration calibration kit from Amersham Biosciences. Purified protein (2 mL, ca. 200 mM) was applied to the column.

Table 5.1 Primers used in BT2542 mutagenesis.

C25A	5’ - CTCGACCGCGAACGTGCTATTGAGAACTTTAAA
F29A	5’ - GTTGTATTGAGAACGCTAAAAAGATCGGGT
E54A	5’ - ATATTTTTGCAGCAGGCGAAAGGACTGATC
G56A	5’ - TTGCAGCAGGAGAAAGCACTGATCACTCCT
I58A	5’ - CAGGAGAAAGGACTGGCCACTCCTGCCGAA
F63A	5’ - GATCACTCCTGCCGAAGCCCGTGACGGCAT
R64A	5’ - ACTCCTGCCGAATTCGCTGACGGCATTCGG
R68A	5’ - TTCCGTGACGGCATTGCGGAGATGATGGGA
D77A	5’ - AAGATGGTAAGCGCCAAGCAGATCGACGCA
I80A	5’ - TGGTAAGCGACAAGCAGGCCGACGCAGCCT
D81A	5’ - CAAGCAGATCGCCGCAGCCTGGAACAGCTT
W84A	5’ - ATCGACGCAGCCGCGAACAGCTTTCTGGTA
L88A	5’ -GCCTGGAACAGCTTTGCGGTAGACATCCCG

5.2.4 Steady State Kinetic Constant Determination

The steady-state kinetic parameters (K_m and k_{cat}) of enzymes were determined from initial reaction velocities measured at varying substrate concentrations for reactions containing 5 mM $MgCl_2$ in 50 mM Na^+MES (pH 6.0) or 50 mM Tris-HCl (pH 7.5) assay buffer at 25 °C. Protein concentrations were determined by using the Bradford assay[(11)] and the released inorganic phosphate was detected by using the EnzChek assay or the acidified ammonium molybdate method[(12)]. Data were fit with the SigmaPlot program to equation 1,

$$V_0 = V_{max} [S] / (K_m + [S]) \qquad \text{(eq. 5.1)}$$

where V_0 = initial velocity, V_{max} = maximum velocity, [S] = substrate concentration and K_m = Michaelis-Menten constant for substrate. The k_{cat} value was calculated from V_{max} and [E] according to the equation $k_{cat} = V_{max} /[E]$, where [E] is the free enzyme concentration.

The rate of *p*-nitrophenyl phosphate (pNPP) hydrolysis was determined by monitoring the increase in absorbance at 410 nm (ε=1.86 mM^{-1} cm^{-1}) at 25°C. The 0.5 mL assay mixtures contained 50 mM Na^+MES, pH 6.0, 5 mM $MgCl_2$, and various concentrations of pNPP.

5.2.5 Determination of Inhibition Constants

Noncompetitive inhibition constants, K_i, or products inhibitors were determined by measuring the initial velocity of YigB catalyzed substrate hydrolysis as a function of substrate concentration (K_m to $5K_m$) and inhibitor concentration (0, $1K_i$ and $2K_i$ μM). The initial velocity data were fitted using KinetAsyst (IntelliKinetics, PA) to equation 5.2,

$$V = V_{max} [S] / [K_m (1+[I]/K_i) + [S] (1+[I]/K_{ii})] \quad \text{(eq. 5.2)}$$

where [A] is the substrate concentration, V is the initial velocity, V_{max} is the maximum velocity, K_m is the Michaelis constant, K_i and K_{ii} are the inhibition constant and [I] is the inhibitor concentration.

5.2.6 pH Rate Profile Analysis

The initial velocities of the *E.coli.* BT2542 catalyzed hydrolysis of FMN were monitored at 25 ^{o}C by measuring the release of free phosphate by using the acidified ammonium molybdate method[(12)]. The buffer systems, 50 mM 2-(N-morpholino) ethanesulfonate (MES) at pH 5.0-6.0, 50 mM HEPES at pH 6.5-7.5 with 5mM $MgCl_2$, were used to measure the pH dependence of the catalyzed reactions. The K_m and v_{max} were acquired by fitting the rates using SigmaPlot program. The k_{cat} values were calculated from the ratio of v_{max} and the enzyme concentrations. The enzyme concentrations were determined using the Bradford assay [(11)]. The pH dependence of the steady-state kinetic constant data were fit initially to equation 5.3

$$\log Y = \log [C / (1 + [H]/K_a + K_b/[H])] \quad \text{(eq. 5.3)}$$

Equation 5.3 was then used to fit data from pH profiles that decrease with unit slopes at both high and low pH. Y is the observed value of either k_{cat}/K_m or k_{cat}, C is the pH-independent plateau value of Y, [H] is the proton concentration, K_a and K_b are apparent acid/base dissociation constants for groups on the enzyme or substrate.

5.2.7 Metal-Free Wild Type BT2542 Preparation and Test of the Catalytic Dependence on Divalent Metal Ions

Freshly prepared BT2542 was dialyzed first against three changes of 50 mM Na^{+}MES (pH 6.0) containing 20 mM EDTA and 1 mM DTT and then against three changes of 50

mM Na^+MES (pH 6.0)/1mM DTT at 4 ^oC. The initial velocities of FMN hydrolysis in reaction solutions containing 0.2 μM metal-free BT2542, 500 μM FMN and varying concentrations of metal ions in 300 μL of 50 mM Na^+MES (pH 6.0, 25 ^oC) were determined by using the acidified ammonium molybdate assay. The kinetic data were analyzed using the computer program of Cleland[13] and equation 5.4,

$$V_o = V_m [A] / (K_M + [A]) \qquad \text{(eq. 5.4)}$$

where [A] is the metal ion concentration, V_o is the initial velocity, V_m is the maximum velocity, K_M is the Michaelis constant for metal activation. The k_{cat} was calculated from the ratio of V_{max} and the enzyme concentration.

5.3 Results and Discussion

5.3.1 Purification of Wild Type YigB

The gene encoding YigB, a HAD phosphatase from *E. coli (K-12)*, was cloned for over expression in *E. coli* BL21 (DE3) cells transformed with the YigB/pET-14b plasmid clone. Cells were collected 12 h after induction and then lysed and the resulting YigB was purified. The purity of protein is ~90% as determined by using SDS-PAGE analysis. The protein concentration was determined by using the Bradford method and by measuring the absorption at 280 nm. The final yield of YigB is 1.6 mg protein/g wet cells.

The exact subunit mass of YigB was determined by using MS-ES mass spectrometry. Analysis of the mass spectrum showed that YigB has a molecular mass of 29583 Da, which agrees closely with the theoretical masses of 29584 Da calculated by using the amino acid composition, derived from the gene sequence, and the ExPASy Molecular

Biology Server program Compute pI/MW. The native molecular weight of YigB was estimated by using FPLC gel filtration column chromatography employing a 1.6 cm x 60 cm Sephacryl S-200HR column (GE Healthcare). The protein was eluted by using 50 mM HEPES buffer containing 100 mM NaCl, pH 7.5, 4 °C, at a flow rate of 1mL/min. The native molecular mass was derived from the measured elution volume by extrapolation of the plot of the elution volume of the protein standards (13.7-220 kDa from GE Healthcare) versus log of the molecular mass. The elution volume of size exclusion chromatography for the native protein YigB was found to be 75.37 mL, from which a molecular weight of 17,906 was estimated, which means that the native YigB is a monomer in solution.

5.3.2 YigB Substrate Specificity

A substrate screen for phosphatase activity (P_i release) was carried out using a set of potential substrates representing all of the main divisions of the *E. coli* phosphometabolome (nucleotides, phosphorylated carbohydrates, organic acids, and amino acids) in order to identify the structural class to which the YigB physiological substrate belongs. The results (Table 5.2) showed that YigB has a modest yet significant level of phosphatase activity (k_{cat}/K_m = 2.2 x 10^4 M^{-1} s^{-1}) towards riboflavin mononucleotide and that it has some activity towards a nucleotide diphosphate (k_{cat}/K_m = 6.0 x 10^2 M^{-1} s^{-1}) and a nucleotide triphosphate (= 1.3 x 10^3 M^{-1} s^{-1}). YigB has no activity towards all sugar phosphates tested. Finally, YigB catalyzes cleavage of the P-N bond in imidodiphosphate with a catalytic efficiency k_{cat}/K_m of 3.9 x 10^3 M^{-1} s^{-1}.

Table 5.2 Steady-state kinetic constants of YigB with different substrates measured by *EnzChek* phosphate assay in 50 mM Tris-HCl buffer containing 5 mM $MgCl_2$, pH 7.5, at 25 °C.

Substrate	K_m (uM)	$k_{cat}(s^{-1})$	$k_{cat}/K_m(M^{-1}s^{-1})$
Flavin mononucleotide (FMN)[a]	$(4.8 \pm 0.2) \times 10^2$	10.7 ± 0.1	2.2×10^4
Imidodiphosphate	$(2.0 \pm 0.2) \times 10^3$	7.5 ± 0.2	3.9×10^3
ATP	$(7.0 \pm 0.9) \times 10^1$	$(9.0 \pm 0.3) \times 10^{-2}$	1.3×10^3
5-Phospho-D-ribose 1-diphosphate	$(1.9 \pm 0.7) \times 10^2$	$(1.1 \pm 0.1) \times 10^{-1}$	6.9×10^2
ADP	$(1.3 \pm 0.1) \times 10^2$	$(8.0 \pm 0.2) \times 10^{-2}$	6.0×10^2
Ribitol-5-phosphate	N.D	N.D	N.D
D-Glucose-6-phosphate	N.D	N.D	N.D
Serine-3-phosphate	N.D	N.D	N.D
ß-Glucose-1, 6-biphosphate	N.D	N.D	N.D
3-Phosphoglyeric acid	N.D	N.D	N.D
AMP	N.D	N.D	N.D

[a] which was measured by the acidified ammonium molybdate method in 50mM Na^+MES buffer containing 5mM $MgCl_2$, pH 6.0.

5.3.3 Product Riboflavin is a Competitive Inhibitor of YigB

Product inhibition, a phenomenon that is related to enzyme mechanism[(14-16)], is caused by accumulation of the product. Product inhibition can not only occur by microscopic

reversal of the enzymatic reaction, but also by a specific interaction between the enzyme and the product. In general, there are three types of product inhibition, termed competitive, uncompetitive and noncompetitive.

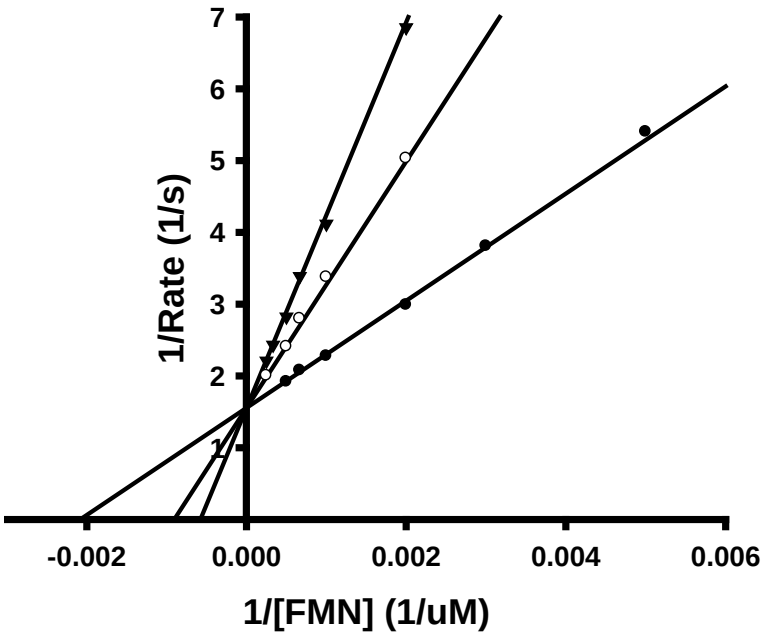


Figure 5.2 Lineweaver-Burk plots of riboflavin inhibited YigB catalyzed FMN hydrolysis in 50 mM MES (pH 6.0, 25 ^{o}C) containing 5 mM $MgCl_2$. (●) 0 mM riboflavin; (○) 1.65 mM riboflavin; (▼) 3.3 mM riboflavin.

The results of the product inhibition study described above show that riboflavin acts as a competitive inhibitor of YigB with a K_i = 1.27 ± 0.0045 mM (Figure 5.2) against FMN. This means that riboflavin binds directly to the free enzyme and, thus, that the order of release of the YigB hydrolysis products is inorganic phosphate first followed by riboflavin.

5.3.4 Purification of BT2542 Wild Type and Mutants

Recombinant wild type BT2542 and its mutants were purified from the *E. coli* BL21 (DE3) cells, transformed by using BT2542 (UniProtKB/Swiss-Prot: Q8A4Q5) or the corresponding mutagenesis. The purity of each protein is ~90% as determined by using SDS-PAGE analysis. After dialyzing against 50 mM MES buffer containing 5 mM $MgCl_2$, pH 6.0, the proteins were concentrated using a 10K Amicon Ultra Centrifugal filter (Millipore) and then stored at -80 °C.

Table 5.3 Steady-state kinetic constants of BT2542 hydrolyzed different substrates in 50 mM Na^+MES buffer, pH 6.0, 25°C.

Substrate	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
FMN	2.6 ± 0.1	(1.1 ± 0.1) x 10^2	2.4 x 10^4
α-D-Mannose-1-phosphate	4.3 ± 0.1	(2.3 ± 0.1) x 10^3	2.0 x 10^3
α-D-Glucose-1-phosphate	(8.8 ± 0.1) x 10^{-1}	(3.4 ± 0.1) x 10^3	2.6 x 10^2
Ribitol-5-phosphate	1.4 ± 0.1	(1.2 ± 0.1) x 10^4	1.2 x 10^2

5.3.5 BT2542 Substrate Specificity

According to the HTS results, BT2542 is a highly specific catalyst of FMN hydrolysis and it also has some activity towards other sugar phosphates (Table 5.3). The activity of this enzyme toward ribitol-5-phosphate was also tested. Part of this substrate mimics the part of the FMN molecule that does not contain the aromatic moiety. Compared to FMN, the catalytic efficiency of the reaction of ribitol-5-phosphate is 200-fold lower (from 2.4

x 10^4 $M^{-1}s^{-1}$to 1.2 x 10^2 $M^{-1}s^{-1}$). The decrease in reactivity is mainly caused by the K_m value (K_m of ribitol-5-phosphate is 100-fold larger than the K_m of FMN), which is the apparent dissociation constant that reflects the binding affinity of the substrate[17]. Therefore, the significant increase of K_m shows the BT2542 binding pocket prefers hydrophobic interactions with the aromatic ring. In a later section of this chapter, further studies designed to define substrate recognition residues in this enzyme will be described.

5.3.6 pH Rate Profile of Wild Type BT2542

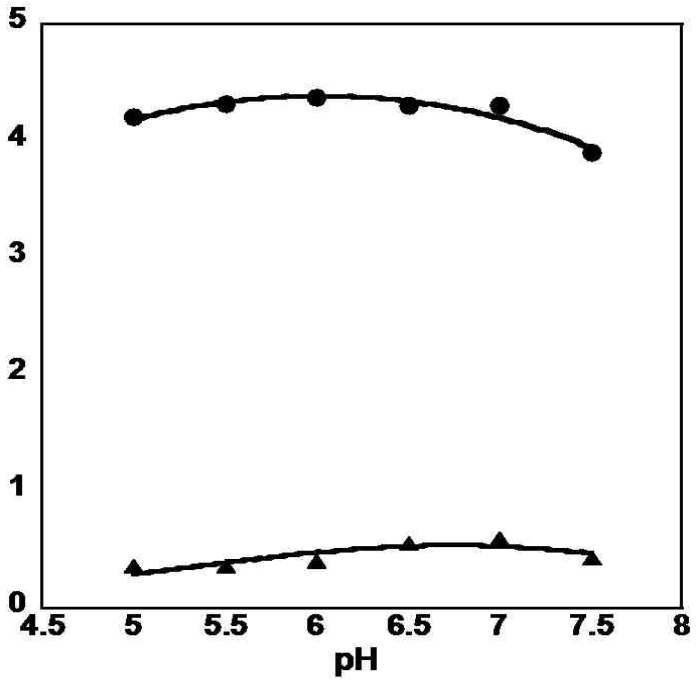


Figure 5.3 The overall pH kinetic pattern of BT2542 catalyzing FMN at 25°C. (●) $\log(k_{cat}/K_m)$, (▲) $\log k_{cat}$.

Many polar residues exist in enzyme active sites and some of them are capable of accepting or donating protons. Most residues participate in enzyme catalysis using only

one protonation form. Therefore, incorrect protonation or deprotonation of the residues could lead to reduced enzyme activity either by decreasing the effective concentration of the reactive species or jeopardizing the stability of an intermediate. Whether the side chain of a potentially acidic amino acid is protonated or not depends on its pKa value, which is influenced by the intrinsic property of the side chain and the microenvironment surrounding it within the enzyme active site[18]. A change in the pH of the reaction solution can affect the protonation state of a catalytically important residue changing it from an active to an inactive form, and thereby influence the kinetic patterns (pH rate profiles) of the enzyme catalysis. We have used pH rate profiles to provide information about the nature of the catalytic mechanism employed by the enzyme and also to identify these residues that play potential roles in acid/base catalysis.

The most commonly measured kinetic constants in pH profile studies are k_{cat} and k_{cat}/K_m[19]. Typically, the value of k_{cat} follows the pKa of the enzyme-substrate complex and the pH dependence of k_{cat} monitors the process ES = E + P. In contrast, the pH dependence of k_{cat}/K_m follows the ionization tendancies of the free enzyme and the free substrate. Thus, the pH profile of k_{cat}/K_m monitors the process E + S = E + P[17]. Usually, variations of kinetic constants with pH follow bell-shaped curves, with reduced values at both pH extremes. This pattern suggests that two ionizable groups are involved in the catalytic process, where one loses its function upon deprotonation and one loses function upon protonation.

The k_{cat} and k_{cat}/K_m pH profiles of wild type BT2542 are shown in Figure 5.3. The curves are well fit to eq. 5.3, giving a pKa = 4.97 ± 0.17 and a pKb = 7.6 ± 0.27 for k_{cat} pH profile, and a pKa = 4.69 ± 0.07 and a pKb = 10.25 ± 1.53 for k_{cat}/K_m. The drop in

k_{cat} with decreasing pH means that a group needs to remain unprotonated for maximum catalysis. The drop in k_{cat} with the increasing pH indicates the need for a group to remain protonated for maximum catalysis. No attempt has yet been made to assign these ionizable residues in active site of BT2542.

5.3.7 Metal Ion Dependence of Wild Type BT2542 Catalysis

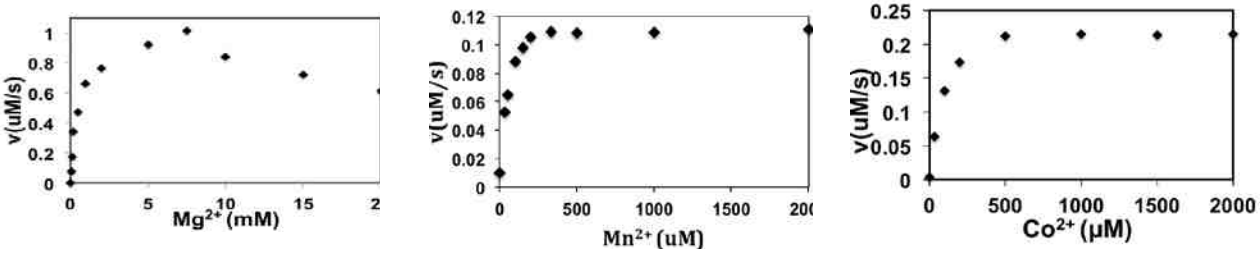


Figure 5.4 The effects of varies metal ions on the reaction of BT2542 hydrolyzed FMN. **(A)** Reaction of 0.2 μM metal-free BT2542 catalyzed hydrolysis of FMN vs Mg^{2+} concentration. **(B)** Reaction of 0.2 μM metal-free BT2542 and 0-2mM Mn^{2+}. **(C)** 0.2 μM metal-free BT2542 and 0-2mM Co^{2+}. All the reactions are performed in 50 mM Na^{+}MES buffer, pH 6.0, 25 °C.

Metal ions play important functional or structural roles roles in a broad variety of proteins. The HAD superfamily members, except for 2-haloacid dehalogenase[20], utilize Mg^{2+} as a cofactor for catalysis[21]. Some other divalent metal are also known to be activators for catalysis by HADSF phosphatases[22, 23]. The effects of various metal cations (Mg^{2+}, Mn^{2+} and Co^{2+}) on the activity of BT2542 were probed in this study (Figure 5.4) and the kinetic constants for each metal ion promoted reaction of FMN are listed in Table 5.4.

Table 5.4 The kinetic constants k_{cat}, K_m, and k_{cat}/K_m of several metal ions on BT2542 hydrolysis of FMN in 50 mM Na^+MES reaction buffer, pH 6.0, 25°C.

Metal	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
Mg^{2+}	2.0 ± 0.1	(8.5 ± 0.5) x 10^2	2.3 x 10^3
Mn^{2+}	(6.6 ± 0.1) x 10^{-1}	(5.0 ± 0.2) x 10^1	1.3 x 10^4
Co^{2+}	1.3 ± 0.1	(9.5 ± 0.4) x 10^1	1.4 x 10^4

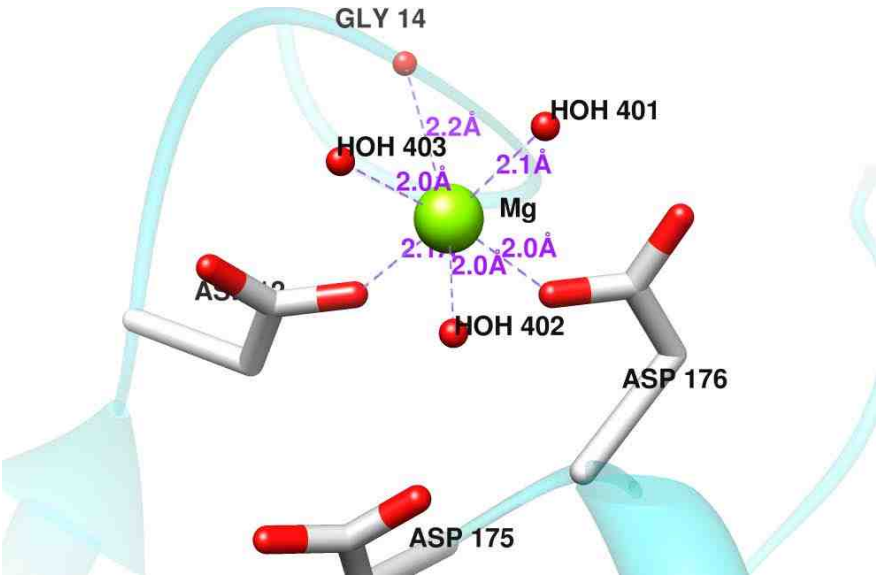


Figure 5.5 Structure of BT2542 with a Mg^{2+} in the active site (PDB codes 4DFD). Mg^{2+} (green) displays almost perfect octahedral coordination geometry with six ligands, including three water molecules (red), Gly14, Asp12 and Asn176 directly coordinate the Mg^{2+} ion.

In the absence of added divalent cations, BT2542 is not catalytically active. Thus, all of the metal ions tested here act as activators for BT2542 (Figure 5.4). Binding of Mn^{2+} and Co^{2+} to the protein is much tighter than is binding of Mg^{2+} (Table 5.4). The coordination geometry of Mg^{2+} in the BT2542 active site, shown in Figure 5.5, almost perfectly octahedral with six ligands, includes three water molecules, Asp12, Asp176 and Gly14.

5.3.8 BT2542 Substrate Recognition Residues

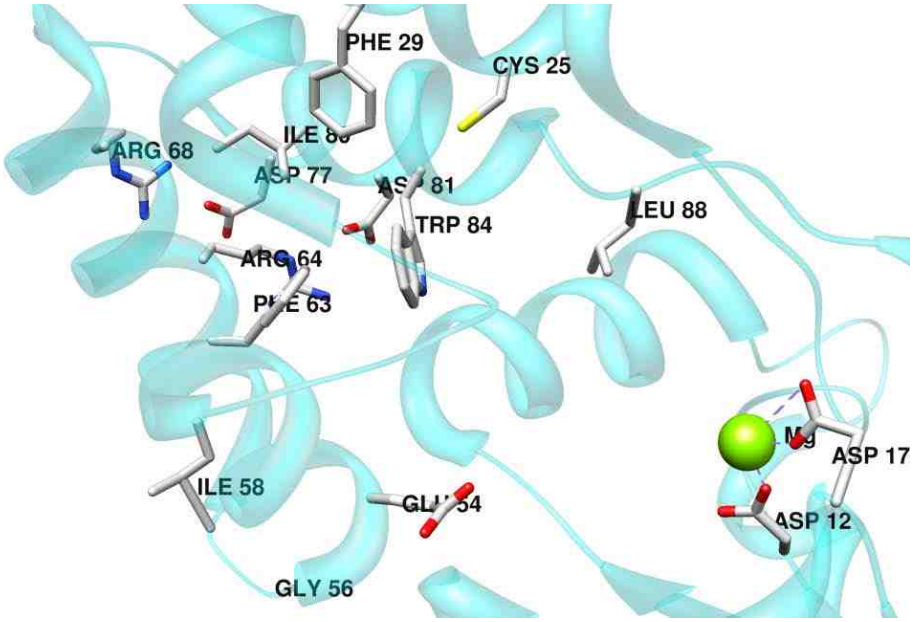


Figure 5.6 Location of conserved residues in the BT2542 cap domain. The magnesium ion is shown in green.

Although nucleophilic catalysis by HAD phosphatases is accomplished by using a conserved Asp residue in the core domain, residues located in the cap domain also play important roles in substrate recognition. In order to determine the residues in BT2542 that

are responsible for the substrate recognition, a series of site-directed mutagenesis studies were carried out (Table 5.5).

Table. 5.5 The steady-state kinetic constants k_{cat} and K_m for BT2542 wild and mutant hydrolysis of FMN are measured in 50 mM Na^+MES buffer containing 5 mM $MgCl_2$, pH 6.0, 25 ^oC.

Protein	k_{cat} (sec^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}sec^{-1}$)
BT2542 Wild Type	2.6 ± 0.1	(1.1± 0.1) x 10^2	2.4 x 10^4
C25A	5.5 ± 0.1	(1.2 ± 0.1) x 10^2	4.6 x 10^4
F29A		Insoluble	
E54A		No Activity	
G56A	(5.0 ± 0.1) x 10^{-2}	(3.5 ± 0.3) x 10^2	1.4 x 10^2
I58A	4.0 ± 0.1	(2.1 ± 0.1) x 10^2	1.9 x 10^4
F63A		Insoluble	
R64A		Insoluble	
R68A	1.1 ± 0.1	(4.3 ± 0.3) x 10^2	2.6 x 10^3
D77A	4.7 ± 0.2	(9.0 ± 0.9) x 10^2	5.2 x 10^3
I80A		Insoluble	
D81A		Insoluble	
W84A		No Overexpression	
W84F	3.3 ± 0.1	(3.1 ± 0.3) x 10^2	1.1 x 10^4
L88A		Insoluble	

Cap domain residues are highly conserved in orthologs which have >45% identity to BT2542 in *Bacteroides thetaiotaomicron*. Unfortunately, replacement of some of these cap domain residues, including Phe29, Phe63, Arg64, Ile80, Asp81 and Ley88, by Ala, results in formation of inclusion bodies (i. e., unstable proteins). In addition, replacement of the polar amino acid Glu54 with Ala in this domain results in loss of all detectable activity. Finally, a Trp84 to Ala, variant could be overexpressed in the host *E coli.* BL21 (DE3) cells.

However, replacement of Trp84 with the aromatic residue Phe did not have a significant effect on the catalytic efficiency (Table 5.5), and neither do replacements of Cys25 and Ile58 by Ala. For example the C25A and I58A mutants only have 2-fold lower K_m (K_m from 110 ± 7 μM for wild type to 120 ± 4 μM and 210 ± 10 μM, respectively). The crystal structure of BT2542 (Figure 5.6) indicates that these two stringently conserved residues are located on the surface of the protein, a possible reason for why neither contributes to substrate binding nor turnover.

The flavin moiety of substrate FMN contains several polar groups. Therefore, hydrogen bonds could be formed between FMN and cap residues resulting in positioning of the substrate. Not surprisingly, replacement of Arg68 and Asp77 by nonpolar Ala leads to reductions in the respective binding affinities by 4 and 8-fold (K_m from 110 ± 7 μM of wild type to 430 ± 30 μM and 900 ± 90 μM respectively). Replacement of Gly56 by Ala reduces the turnover rate 50-fold and the binding affinity 3-fold. Thus, because the space-saving Gly56 residue is located close to the substrate entrance site of the protein (Figure 5.6), an increase in the steric bulk at this position caused by the Ala mutation could hinder the binding of FMN and/or the release of the product riboflavin.

5.3.9 Mutagenesis of the BT2542 Active Site Residue Asp+2

It is well known that two important highly conserved Asp residues are present in the core domains of HADSF phosphatases[24]. One of these residues serves as the key catalytic nucleophile and the other, Asp+2, which is positioned two residues from the Asp nucleophile, functions as a general acid/base to first protonate the leaving group and then deprotonate the water nucleophile. While in BT2542, Gly14 is present in the Asp+2 position (Figure 5.7), alignment of BT2542 (data not shown) with its homologs in *Bacteroides* which have a 40% or higher identity indicates that the Gly residue is conserved in all these proteins.

The stringently conserved Gly14 residue was replaced by Ala to probe the effect of a modest increase in the bulk of the side chain, with Val to probe the impact of a large increase in side chain bulk, and by Asp to probe the tolerance of a polar large bulk increase side chain. The kinetic constants of the wild type BT2542 and its Gly14 mutants are shown in Table 5.6. In BT2542, Gly14 helps bind Mg^{2+} to form an octahedral conformation. While in other HADSF phosphatases, this position is occupied by an Asp residue (Figure 5.7). The replacement of Gly with Asp leads to a ~20-fold reduction in the catalytic efficiency (k_{cat}/K_m from 2.4×10^4 for wild type to 1.3×10^3 for G14D), which is mainly caused by a change in K_m. Both Ala and Val substitution for Gly14 are tolerated by the enzyme, causing only a subtle decrease in the catalytic efficiency (2-fold reduction for G14A and 3-fold reduction for G14V) (Table 5.6).

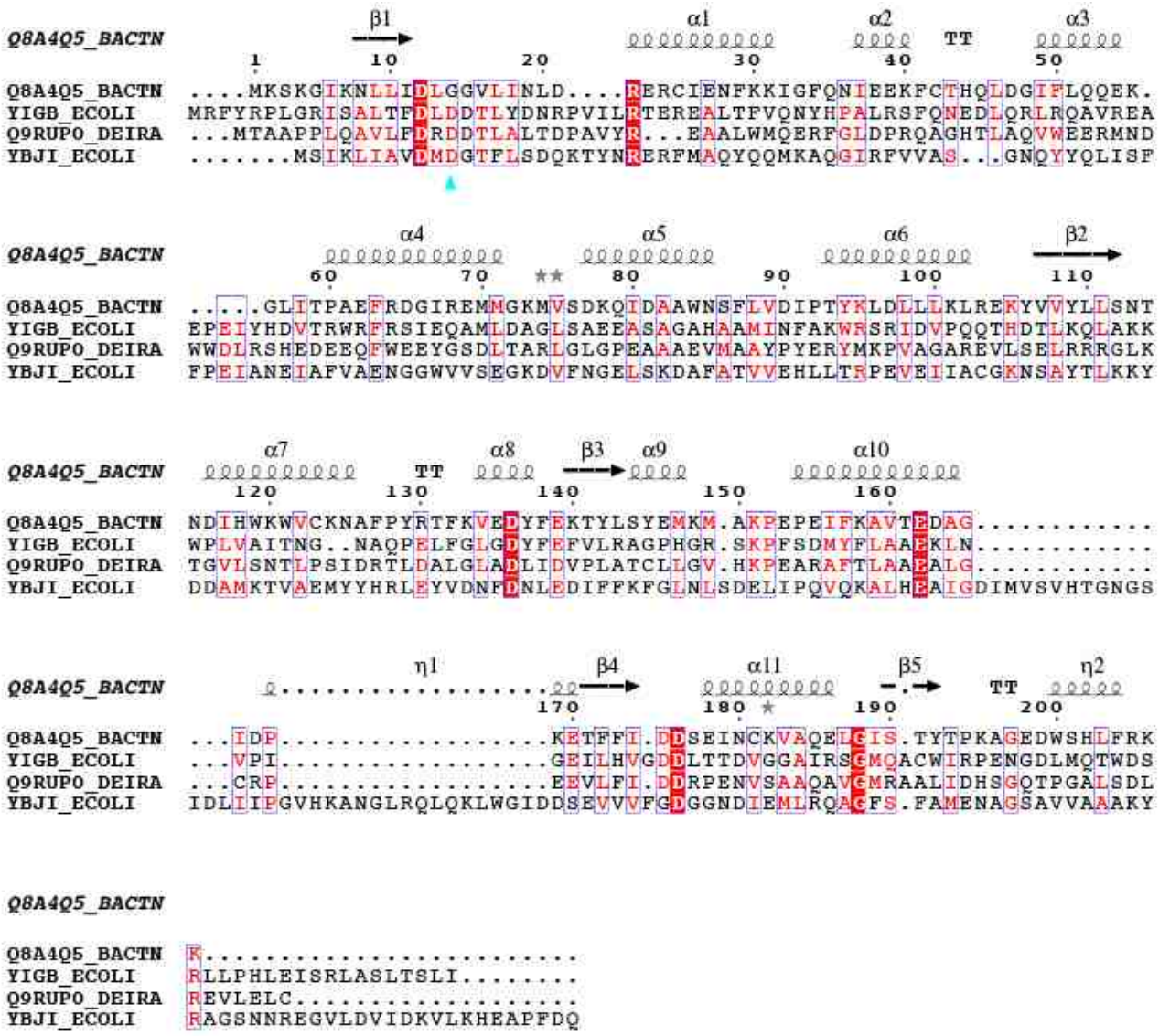


Figure 5.7 The sequence alignment of BT2542 (Q8A4Q5) with the other three putative FMN nucleotidases. Red background with white letter shows the stringently conserved residues, red and black letter in blue rectangle shows the partly conserved residues. The cyan triangle indicates the (Asp+2) position. The sequences were aligned using the Clustal Omega tool (http://www.ebi.ac.uk/Tools/msa/clustalo/) and displayed in ESPript (http://espript.ibcp.fr/ESPript/ESPript/).

Table 5.6 The steady-state kinetic constants k_{cat} and K_m for BT2542 wild and mutants hydrolyzing FMN are measured in 50 mM Na^+MES buffer containing 5mM $MgCl_2$, pH 6.0, 25 ^oC.

Protein	k_{cat} (sec^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}sec^{-1}$)
BT2542 wild type	2.6 ± 0.1	(1.1 ± 0.1) x 10^2	2.4 x 10^4
G14D	3.0 ± 0.1	(2.4 ± 0.2) x 10^3	1.3 x 10^3
G14A	2.4 ± 0.1	(2.2 ± 0.1) x 10^2	1.1 x 10^4
G14V	1.4 ± 0.1	(1.7 ± 0.1) x 10^2	8.2 x 10^3

5.3.10 Substrate Specificity of 501335 and 501193

Table 5.7 Steady-state kinetic constants of YbjL (EFI-501335) with different substrates measured by acidified ammonium molybdate method in 50 mM Na^+MES buffer containing 5mM $MgCl_2$, pH 6.0, at 25°C.

Substrate	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
FMN	2.8 ± 0.1	(4.6 ± 0.4) x 10^2	6.2 x 10^3
Ribitol-5-phosphate	18.3 ± 0.3	(6.7 ± 0.3) x 10^3	2.7 x 10^3
Mannose-6-phosphate	5.6 ± 0.2	(9.0 ± 0.6) x 10^3	6.2 x 10^2
Glucosamine-6-phosphate	1.6 ± 0.1	(8.5 ± 0.5) x 10^3	1.9 x 10^2

the K_m of glucose-6-phosphate is >10mM.

Information about the substrate specificity of 501335 and 501193 was obtained by using a high throughput screen (HTS) carried out by Allen and her coworkers at Boston University. Both proteins were found to be quite specific for several substrates, having highest activities towards FMN. The kinetic constants of 501335 and 501193 for FMN and several sugar phosphates are listed in Table 5.7 and Table 5.8.

501193 has a very low activity towards 2'-AMP and d, l-glyceraldehyde-3-phosphate. (Incubating 2μM 501193 with 500 μM substrate for 5min, the velocity is 6.8 x 10^{-3} μmol min^{-1} mg^{-1} protein).

Table 5.8 Steady-state kinetic constants of Q9RUP0 (EFI-501193) with FMN measured by acidified ammonium molybdate method in 50 mM Na^{+}MES buffer containing 5mM $MgCl_2$, pH 6.0, at 25°C.

Substrate	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
FMN	1.3 ± 0.1	(1.9 ± 0.2) x 10^3	6.9 x 10^2

5.4 Conclusion

Enzyme Function Initiative (EFI) is a large-scale collaborative project whose goal is to develop a strategy to determine enzyme function using an integrated sequence-structure based approach. In the studies described above, the catalytic properties of four putative HAD phosphatases, identified by using high throughput screening (HTPS) were determined and compared side by side in order to determine the functional diversity of HADs in different organisms. All four proteins showed significant phosphatase activity towards FMN at a level predicted to be of physiological relevance. HTPS also identified

the preferred substrates for all four proteins among *ca.* 160 phosphosugars, nucleotides and amino acids. 501335 has a broader substrate spectrum than the other three enzymes.

The results of a pH profile study of the BT2542 catalyzed hydrolysis reaction of FMN show that the protein has a slightly acidic pH efficiency optimum. Moreover, the investigation demonstrated that the phosphatase activity of BT2542 is strictly dependent on the presence of a divalent metal cation (Mg^{2+}, Mn^{2+} and Co^{2+}). Analysis of the crystal structure of BT2542, which was solved by the Allen group at Boston University, indicates that it is different from other HAD phosphatases in that the Asp+2 position in the core domain contains a Gly rather than Asp residue. When this Gly residue is replaced by Asp, the catalytic efficiency of the protein is 10-fold lowered compared to that of wild type BT2542, while Ala and Val substitution for Gly do not greatly alter the catalytic activity. The results of studies of the substrate recognition residues of BT2542 indicate that several residues, including Phe29, Phe63, Arg64, Ile80, Asp81 and Leu88, located in the cap domain are very important in governing the stability of the protein.

Analysis of the kinetic parameters (Table 5.2, 5.3, 5.7 and 5.8) of the four proteins revealed that all have a low affinity (K_m is high) for phosphorylated sugars. In addition, while they have high affinities to phosphate containing aromatic substances (Table 5.2), they only have a high catalytic turnover rate towards FMN. This finding indicates the substrate binding pocket of these proteins favors hydrophobic substrates.

In summary, this initial study has provided a brief view of the EFI's strategy for protein functional assignment. More detailed studies are required in this area to complete the task of functional assignments, especially in vivo testing and the crystal structure

determination of the complexes of the proteins with substrate analogs or of inactive mutants with the proposed physiological substrates.

Reference

1. Brenner, S. E. (1999) Errors in genome annotation, *Trends Genet 15*, 132-133.
2. Devos, D., and Valencia, A. (2001) Intrinsic errors in genome annotation, *Trends Genet 17*, 429-431.
3. Schnoes, A. M., Brown, S. D., Dodevski, I., and Babbitt, P. C. (2009) Annotation Error in Public Databases: Misannotation of Molecular Function in Enzyme Superfamilies, *Plos Comput Biol 5*.
4. Gerlt, J. A., Allen, K. N., Almo, S. C., Armstrong, R. N., Babbitt, P. C., Cronan, J. E., Dunaway-Mariano, D., Imker, H. J., Jacobson, M. P., Minor, W., Poulter, C. D., Raushel, F. M., Sali, A., Shoichet, B. K., and Sweedler, J. V. (2011) The Enzyme Function Initiative, *Biochemistry-Us 50*, 9950-9962.
5. Koonin, E. V., and Tatusov, R. L. (1994) Computer analysis of bacterial haloacid dehalogenases defines a large superfamily of hydrolases with diverse specificity. Application of an iterative approach to database search, *J Mol Biol 244*, 125-132.
6. Aravind, L., Galperin, M. Y., and Koonin, E. V. (1998) The catalytic domain of the P-type ATPase has the haloacid dehalogenase fold, *Trends Biochem Sci 23*, 127-129.
7. Morais, M. C., Zhang, W., Baker, A. S., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2000) The crystal structure of bacillus cereus phosphonoacetaldehyde hydrolase: insight into catalysis of phosphorus bond cleavage and catalytic diversification within the HAD enzyme superfamily, *Biochemistry 39*, 10385-10396.

8. Allen, K. N., and Dunaway-Mariano, D. (2004) Phosphoryl group transfer: evolution of a catalytic scaffold, *Trends in biochemical sciences 29*, 495-503.

9. Herschlag, D., Jencks, W.P. (1989) Phosphoryl transfer to oxyanions: the nature of the transition state and electrostatic repulsion, *J. Am. Chem. Soc. 111*.

10. Morais, M. C., Zhang, G. F., Zhang, W. H., Olsen, D. B., Dunaway-Mariano, D., and Allen, K. N. (2004) X-ray crystallographic and site-directed mutagenesis analysis of the mechanism of Schiff-base formation in phosphonoacetaldehyde hydrolase catalysis, *J Biol Chem 279*, 9353-9361.

11. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Analytical biochemistry 72*, 248-254.

12. Ames, B. N. (1966) Assay of inorganic phosphate, total phosphate and phosphatases, *Methods Enzymol. 8*, 115-118.

13. Cleland, W. W. (1979) Statistical analysis of enzyme kinetic data, *Methods Enzymol 63*, 103-138.

14. Miao, Y. L., Chen, J. Y., Jiang, X. J., and Huang, Z. (2012) Kinetic Studies on the Product Inhibition of Enzymatic Lignocellulose Hydrolysis, *Appl Biochem Biotech 167*, 358-366.

15. Zhang, Z. Y., and Vanetten, R. L. (1991) Pre-Steady-State and Steady-State Kinetic-Analysis of the Low-Molecular-Weight Phosphotyrosyl Protein Phosphatase from Bovine Heart, *J Biol Chem 266*, 1516-1525.

16. Robertson, B. A., Schroeder, G. K., Jin, Z. N., Johnson, K. A., and Whitman, C. P. (2009) Pre-Steady-State Kinetic Analysis of cis-3-Chloroacrylic Acid Dehalogenase: Analysis and Implications, *Biochemistry-Us 48*, 11737-11744.

17. Fersht, A. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman.

18. Highbarger, L. A., Gerlt, J. A., and Kenyon, G. L. (1996) Mechanism of the reaction catalyzed by acetoacetate decarboxylase. Importance of lysine 116 in determining the pK(a) of active-site lysine 115, *Biochemistry-Us 35*, 41-46.

19. Cleland, W. W. (1982) The Use of Ph Studies to Determine Chemical Mechanisms of Enzyme-Catalyzed Reactions, *Method Enzymol 87*, 390-405.

20. Ridder, I. S., and Dijkstra, B. W. (1999) Identification of the Mg2+-binding site in the P-type ATPase and phosphatase members of the HAD (haloacid dehalogenase) superfamily by structural similarity to the response regulator protein CheY, *Biochemical Journal 339*, 223-226.

21. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D., and Aravind, L. (2006) Evolutionary genomics of the HAD superfamily: Understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes, *J Mol Biol 361*, 1003-1034.

22. Kuznetsova, E., Proudfoot, M., Gonzales, C. F., Brown, G., Omelchenko, M. V., Borozan, I., Carmel, L., Wolf, Y. I., Mori, H., Savchenko, A. V., Arrowsmith, C. H., Koonin, E. V., Edwards, A. M., and Yakunin, A. F. (2006) Genome-wide analysis of substrate specificities of the Escherichia coli haloacid dehalogenase-like phosphatase family, *J Biol Chem 281*, 36149-36161.

23. Thaller, M. C., Schippa, S., Bonci, A., Cresti, S., and Rossolini, G. M. (1997) Identification of the gene (aphA) encoding the class B acid phosphatase/phosphotransferase of Escherichia coli MG1655 and characterization of its product, *Fems Microbiol Lett 146*, 191-198.

24. Allen, K. N., and Dunaway-Mariano, D. (2009) Markers of fitness in a successful enzyme superfamily, *Current opinion in structural biology 19*, 658-665.