5-2012

Structural and Mechanistic Investigations of Phosphothreonine Lyase Class of Enzymes

Alok Gopalkrishna Shenoy
Utah State University

MECHANISTIC AND STRUCTURAL INVESTIGATIONS OF THE PHOSPHOTHREONINE LYASE CLASS OF ENZYMES

by

Alok Gopalkrishna Shenoy

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Chemistry

Approved:

_____	_____
Dr. Alvan C. Hengge Major Professor	Dr. Sean J. Johnson Committee Member
_____	_____
Dr. Cheng Wei Tom Chang Committee Member	Dr. Mark R. McLellan Vice President for Research and Dean of School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2012

ABSTRACT

Structural and Mechanistic Investigations of Phosphothreonine Lyase Class of Enzymes

by

Alok Gopalkrishna Shenoy, Master of Science

Utah State University, 2012

Major Professor: Prof. Alvan C. Hengge
Department: Chemistry and Biochemistry

The phosphothreonine lyase class of enzymes represents a recently discovered set of enzymes that catalyze a dephosphorylation reaction. The catalysis is carried out using a unique elimination mechanism without any involvement of cofactors. Crystallographic studies of SpvC, a phosphothreonine lyase, and its mutant show that the mutation of the general catalytic acid does not result in any significant perturbations to the tertiary and the secondary structure of the protein. Using results from the structural studies and a deuterium isotope exchange experiment, we conclude that the reaction catalyzed by SpvC may not involve formation of a carbanion at the active site.

(56 Pages)

PUBLIC ABSTRACT

Structural and mechanistic investigations of phosphothreonine lyase class of enzymes

Phosphorylation and dephosphorylation are a highly pervasive mechanism in biology that is used by the cell to modulate enzymes and proteins. The presence of a phosphate group can activate or deactivate an enzyme. The phosphate group is linked to a protein by a phosphoester bond that is known to be highly stable in cytoplasmic pH range. Thus the breaking and formation of these bonds need to be effected by enzymes.

Recent discovery of the activity carried out by certain virulence related proteins (OspF released by Shigella and SpvC released by Salmonella) have resulted in a necessity to create a new class of dephosphorylating enzymes – phosphothreonine lyases.

This class of proteins forms an important drug target against salmonella and shigella. The objective of our work has been to decipher the mechanism by which these effector proteins catalyze the dephosphorylation. Elimination reactions can be carried out by several mechanisms, namely E1, E2, and E1cB. Identification of the mechanism will aid in drug design. This report presents the results of analytical and crystallographic techniques used to study the mechanism of the elimination catalyzed by SpvC.

Alok G. Shenoy

I dedicate the work presented herein to my Guru, my parents, in-laws, sunny and my wife Padmashri Suresh.

ACKNOWLEDGMENTS

I would like to thank Prof. Alvan C. Hengge for allowing me to pursue research in his esteemed lab. His patience and mentorship have been instrumental in achieving the body of work presented here. I would also like to thank my committee members, Prof. Sean J. Johnson, who lent me his valuable time and resources during my training in crystallography, and Prof. Tom Chang, who has been a fountain of knowledge for my synthetic chemistry work. Lastly, I would like to thank my colleagues in the Hengge Lab who have been a constant presence in the lab, late hours and early mornings – it has been a pleasure working with you.

I would like to thank my parents for being great role models and timely inspiration. I would like to give special and heartfelt thanks to my wife, Padmashri Suresh, who has been a constant companion through highs and lows, and a great source of inspiration and strength.

Alok G. Shenoy

CONTENTS

Page

ABSTRACT iii

PUBLIC ABSTRACT iv

ACKNOWLEDGMENTS vi

CONTENTS vii

LIST OF TABLES ix

LIST OF FIGURES x

CHAPTER

1. INTRODUCTION 1

Classification of Dephosphorylating Enzymes 1
Phosphothreonine lyases released by Salmonella & Shigella 2

2. LITERATURE REVIEW 4

OspF classified as a Dual specificity phosphatase 4
Structural and mutational studies show that OspF and SpvC dephosphorylate threonine residue 7
Structural data on phosphothreonine lyase class of enzymes is limited 14
The mechanism of the elimination reaction is yet to be determined. 16

3. MATERIALS AND METHODS 20

Genomics – cloning and site directed mutagenesis 20
Purification of the enzymes SpvC and SpvC H106N 22
Enzymology – assay protocol and materials 24
Crystallization of SpvC WT & H106N mutants 25
Data collection and processing 26

4. RESULTS 28

Crystallography 28

Structural aspects of SpvC WT (4HAH) as compared to the previously published structure (2Z8M) **28**
Structural aspects of SpvC H106N (PDB ID: 4H43) 32

5. DISCUSSION 38

6. CONCLUSION 42

REFERENCES 44

LIST OF TABLES

Table Page

1. Published crystal structures of enzymes comprising the phosphothreonine lyase class. Shown are the PDB ID's, the resolution of the structures, the space group and the ligands. 16

2. Primers used to clone *spvC* into the prk793 variant and carry out site directed mutagenesis, mutating H106 to N106. The mutation (H106N) highlighted in red. 21

3. Data collection and refinement statistics for SpvC WT. 29

4. Average B factors of side chains and backbone of several residues of SpvC H106N (4H43) that form the movable loop. It is seen that the residues, which form the loop, have high b-factors as compared to the overall b-factor of the protein (46.9). 35

LIST OF FIGURES

Figure Page

1. . Immunoblot assay used to identify the substrate of OspF.[6] It was noted that the WT induced activation of the MEK1 kinase, but the subsequent activation of Erk was absent as detected by western blot using antibodies specific to phosphorylated MEK1 and Erk. The remaining pathways were unaffected. 5

2. Immunoblot shows that OspF is inhibited by Vanadate. Also seen is a decrease in intensity of the phosphothreonine bands, even in presence of the inhibitors, implying OspF catalyzes dephosphorylation of both threonine and tyrosine residues. 6

3. A stereo view image of the picomolar inhibitor okadaic acid bound to PP1. The key residues involved in binding of the inhibitor to the protein are labeled.[10] As seen, okadaic acid forms several H-bonds to various active site residues. 7

4. Mass Spectra indicating the loss of 98 Da as shown by Shao et al.[7] The 98Da could not be attributed to dephosphorylation by either a PTP or a DuSP. 8

5. NMR spectra comparing peptide isolated following incubation with SpvC and a control experiment lacking the enzyme. The vinylic signal has been attributed to ß-methyldehydroalanine.[11] 9

6. Reaction mechanism put forth by Shao et al. An elimination reaction resulting in a loss of inorganic phosphate ($H_2PO_4^-$ - 97 Da) and a proton (1 Da) which is abstracted from the alpha carbon atom..... 10

7. Ligand-free structure of spvC WT (PDB ID: 2P1W) showing a tertiary structure similar to a cupped hand. 10

8. Electrostatic potential surface of SpvC WT with synthetic peptide (PDB ID: 2Q8Y). Pocket A & Pocket B are two highly electropositive pockets that bind the phosphothreonine and phosphotyrosine respectively. 11

9. Immunoblot assay of WT OspF and various mutants with Erk2 substrate. An alanine screen was used to identify key catalytic residues.[8] It was seen that K102, H104 and K134 were inactive or extremely slow in catalyzing the reaction. 12

10. In-vitro interactions between recombinant Erk and OspF WT and OspF mutants. 3 and 5 are inputs and pull-down respectively for OspF WT, Lane 6-8 and 11-13 are input and pull-down of OspF mutants K102R, H104A, K134A respectively.[8] Pull-down shows that the mutants bind to the substrate with similar affinity as the wild type. 13

11. Stereo view of SpvC K136A, shown in light blue, bound to phosphopeptide, shown in orange, PDB ID: 2Q8Y. 14

12. Proposed reaction mechanism of phosphothreonine lyase class of enzymes. The K136 behaves as the catalytic base that abstracts the proton followed by the H106 that donates a proton to the leaving phosphate group. 15

13. Sequence alignment of SpvC, VirA, OspF and Edwardsiella putative virulence protein shows highly conserved residues (marked by '*'), specifically the catalytic acid H106, base K134 and several residues involved in binding of substrate 15

14. Possible mechanisms for elimination reactions shown herein schematically. Concerted (14.1), E1cB (14.2) and E1 (14.3) shown herein with phosphoester as starting material resulting in vinylic product and inorganic phosphate as products. 18

15. Tertiary structure of SpvC WT, PDB ID 2Z8M in purple, when compared to our structure of ligand-free SpvC WT, 4HAH in teal, shows that the structures are very similar, with a RMSD of 0.16 for 215 residues. 30

16. Comparisons of the positions of the active site residues of SpvC WT in the structure 4HAH (teal) obtained in this work, compared with a previously reported structure, 2Z8M, in purple. It is clearly seen that the residues from the two structures superimpose each other. 31

17. pY binding pocket residues compared between structures obtained through this work against previously published structures. Overlaid above are 4HAH, in teal, and 2Z8M, in purple. 32

18. Shown above the structural alignment of SpvC WT (4HAH), in teal, on SpvC H106N (4H43) in green. The RMSD for 3D overlap of the two structures for 215 residues is 0.33. 33

19. Active site of SpvC H106N (PDB ID 4H43), in green, structurally aligned to SpvC WT (PDB ID: 4HAH), in teal. The mutation H106 ➔ N106 has been labeled. 34

20. Recognition site of SpvC H106 (PDB ID: 4H43), in green, aligned with SpvC WT (PDB ID: 4HAH), in teal. This site is known to bind to the phosphorylated tyrosine on the substrate. Residue E215 is displaced by 1.8 Å, attributed to the mobile loop it is a part of. 34

21. The movable loop of ligand-free SpvC WT (4HAH), in teal, aligned to SpvC K136A (2Q8Y) in light purple, with substrate bound, in orange. Upon substrate binding the R220 carrying loop changes its conformation to the closed form, such that the R220 is placed in proximity to the phosphothreonine residue (P219) of the substrate. The change also positions E215 in close proximity to the amide carbonyl of the substrate backbone. 36

22. The movable loop carrying the R220 and E215 residues compared between the structures of SpvC WT shown in light blue and the SpvC H106N, shown in green. Residues 200 to 225 compared using SSM [27a] indicated a Q-score of 0.90 for all 26 residues. 37

CHAPTER 1

INTRODUCTION

This report presents work on the investigation of mechanistic aspects of the reaction carried out by effector protein Salmonella plasmid virulence C (SpvC) with the aid of x-ray crystallography. SpvC, an effector protein, catalyzes an elimination of phosphate from phosphorylated MAPK enzyme Erk2.

1.1 Classification of Dephosphorylating Enzymes

Phosphorylation and dephosphorylation are universal mechanisms employed by biological systems to activate or deactivate enzymes. Addition (phosphorylation) and removal (dephosphorylation) of phosphate moiety at cytoplasmic pH require catalysis by enzymes. Enzymes that carry out phosphorylation can be broadly classified as kinases. Enzymes that dephosphorylate can either be classified as phosphatases or phosphothreonine lyases.

Phosphatases refer to a class of enzymes that catalyze the dephosphorylation of the substrate by cleavage of the oxygen-phosphorus bond. This results in formation of an enzyme-phosphate complex and the release of the substrate carrying a free hydroxyl group. The enzyme phosphate complex breaks down in the following step to release inorganic phosphate and regeneration of the enzyme active site.

The phosphothreonine lyase class of enzymes is a set of effector proteins released by various pathogens like Salmonella and Shigella. This class presently comprises of SpvC, OspF, and VirA. These effector proteins are known to directly aid the virulence of the pathogens. The activity of these enzymes involves the disruption of the MAPK pathway, preventing the expression of chemokine interleukin-8. Host cells use this chemokine to trigger the immune

response against the invading pathogen. The lack of interleukin-8 allows the pathogen to proliferate almost without detection through the host intestinal epithelial cells.

1.2 Phosphothreonine lyases released by Salmonella & Shigella

These pathogens are mainly attributed with causing dysentery and are considered a growing threat to humans due to fast developing antibiotic resistant strains. The past few years have seen over 35 widespread outbreaks of salmonellosis in U.S.A. alone,[1] and affecting over one billion people worldwide.[2]

Salmonella carries 5 genes on loci named: *salmonella plasmid virulence* (*spv*). The proteins expressed by these genes, namely *spv R, A, B, C,* and *D*, have been attributed to aid the virulence of the pathogen. While the understanding of these proteins is limited, it has been established that SpvB and SpvC are directly responsible for the post invasion virulence of Salmonella.[3]

Shigella releases over 30 different effector proteins, of which the function of about five proteins are known.[4] Of the proteins released, some have been classified under '*Outer shigella protein'* (*Osp*).[5] OspF is a protein released by Shigella, which is homologous to SpvC from Salmonella with over 70% sequence similarity.

The functions of these enzymes were discovered recently and were initially mistaken to be dual specificity phosphatases (DuSP),[6] and later correctly classified as phosphothreonine lyases.[7] Given that these enzymes are essential for the virulence of the pathogen, phosphothreonine lyases form an important drug target. The mechanism of catalysis and the transition state structure is often used in rational drug development. The mechanism by which phosphothreonine lyases catalyze the elimination of phosphate is unknown. Our objective has

been to decipher the mechanism using analytical tools like deuterium isotope exchange and x-ray crystallography.

CHAPTER 2

LITERATURE REVIEW

The phosphothreonine lyase class of enzymes is a recently discovered set of enzymes. Arbibe et al. [6] initially classified these enzymes as dual specificity phosphatases. Li et al. [7-8] correctly deciphered the activity of OspF and SpvC, and classified these as phosphothreonine lyases. The results and conclusions of these publications are discussed herein.

2.1 OspF classified as a Dual specificity phosphatase

Arbibe et al. [6] were the first group to successfully identify the substrate of OspF. Based on their experimental results, the authors concluded that OspF is a dual specificity phosphatase (DuSP) that dephosphorylates both tyrosine and threonine residues on Erk2. Some of their key conclusions are summarized below.

2.1.1. Erk2 identified as the substrate of OspF

To identify the role of OspF in vivo, the authors incubated HeLa cells with the wild type (WT) and non-virulent (–VP) strains of shigella in separate experiments. Antigens specific to MEK1, Erk, c-Jun and their phosphorylated form pMEK1, pErk, pc-Jun respectively were used to analyze the pool of proteins extracted from the HeLa cells. From the western blot obtained, shown in Figure 1, the authors concluded the following:

- Blot showed presence of inactive MEK1, and a gradual increase in the dually phosphorylated pMEK1 over time. The WT strain induced the activation of MEK1 to pMEK1

- Although MEK1 was phosphorylated, Erk2 remained inactive over 180 minutes. Other pathways involving the activation of NF-**k**B activation remained unaffected.

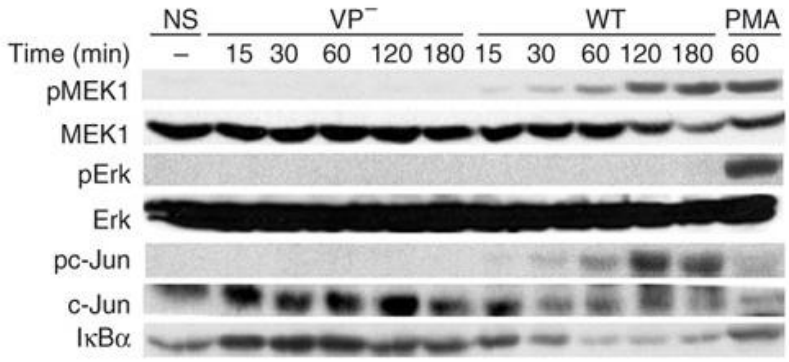

Figure 1. . Immunoblot assay used to identify the substrate of OspF.[6] It was noted that the WT induced activation of the MEK1 kinase, but the subsequent activation of Erk was absent as detected by western blot using antibodies specific to phosphorylated MEK1 and Erk. The remaining pathways were unaffected.

This led them to believe that OspF was able to deactivate Erk2, thus altering the factors affecting the transcription of genes involved in immune response.

2.1.2. OspF classified as dual specificity phosphatase

The MAPK enzyme Erk2 has two phosphorylated residues in its active state – phosphotyrosine and phosphothreonine. To identify the specificity of OspF, the authors used immunoblot assays coupled with inhibition studies. They incubated the enzyme with its substrate Erk2 in presence of two inhibitors :

- Vanadate – an inhibitor of protein tyrosine phosphatases.[9]
- Okadaic acid – a picomolar inhibitor of serine/threonine phosphatase.[10]

These were then subjected to Western Blot, shown in Figure 2, with antigens specific to Erk2, Erk2 with phosphorylated tyrosine (pY) and Erk2 with phosphorylated threonine (pT).

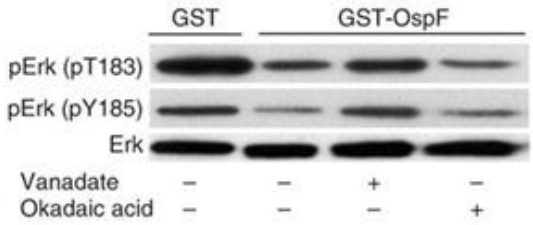

Figure 2. Immunoblot shows that OspF is inhibited by Vanadate. Also seen is a decrease in intensity of the phosphothreonine bands, even in presence of the inhibitors, implying OspF catalyzes dephosphorylation of both threonine and tyrosine residues.

Based on the intensity of the bands of pT and pY, the authors concluded that OspF was inhibited by vanadate, implying specificity for pY. They also observed a decrease in intensity of the pT bands in presence of either inhibitor. This led them to conclude that OspF was dephosphorylating both threonine and tyrosine residue, and hence classified it as a dual specificity phosphatase.

In the course of our literature review, we find that Arbibe et al. [6] failed to consider the significance of the pT-X_{AA1}-pY motif, which was addressed by another research group discussed later. The skewed conclusion could be attributed to several factors, e.g. the crystal structure of okadaic acid bound to PP1, shown in Figure 3. Crystal structure, Figure 3A, indicates the positions of various active site residues with respect to the inhibitor okadaic acid. A schematic view of the hydrogen bonds is shown in Figure 3B, indicating all hydrogen bonds between the protein and the inhibitor.

The OspF lacks any of the conventional sequence motifs of phosphatases. This aspect of OspF would account for the lack of inhibition by okadaic acid. This result combined with the

inhibition of OspF by vanadate points towards phosphotyrosine being essential in binding of the substrate.

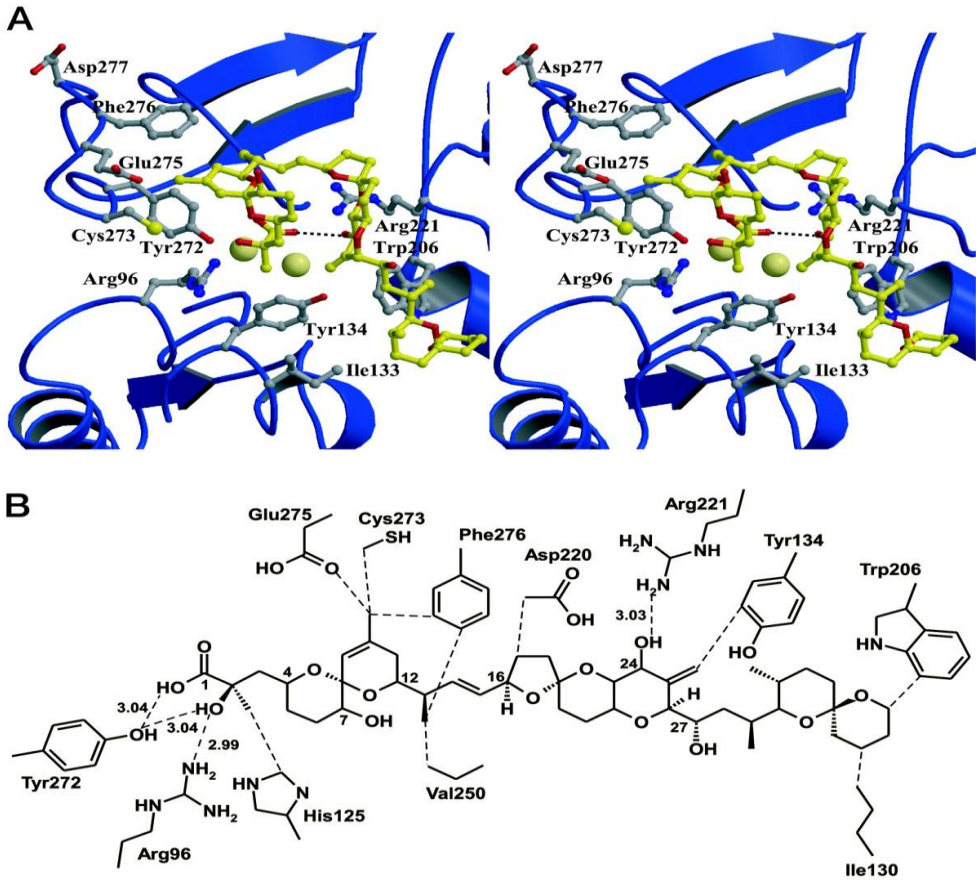

Figure 3. A stereo view image of the picomolar inhibitor okadaic acid bound to PP1. The key residues involved in binding of the inhibitor to the protein are labeled.[10] As seen, okadaic acid forms several H-bonds to various active site residues.

2.2 Structural and mutational studies show that OspF and SpvC dephosphorylate threonine residue

Li et al.[7-8] addressed the shortcomings of the previous publications by using a biochemical approach to determine the activity and the specificity of OspF and SpvC. The group used a combination of analytical techniques including mass spectrometry (MS), nuclear

magnetic resonance (NMR) and X-ray crystallography to study the protein and the substrate. Based on their results, reviewed in this section, the authors concluded that OspF and SpvC catalyze the cleavage of phosphate by a unique elimination reaction without the involvement of any co-factors.

2.2.1 OspF and SpvC dephosphorylate phosphothreonine residue on Erk2

The authors used MS to analyze the product from the reaction between OspF and a synthetic phosphopeptide. Their results, shown in Figure 4, indicated a loss of 98 Da. Dephosphorylation by a phosphatase would result in a loss of 78 Da, hence dephosphorylation of both pY and pT would result in a decrease of 156 Da. The mass spectra also showed that the phosphate group on tyrosine was intact, implying that these enzymes were specifically targeting the phosphothreonine residue on the substrate.

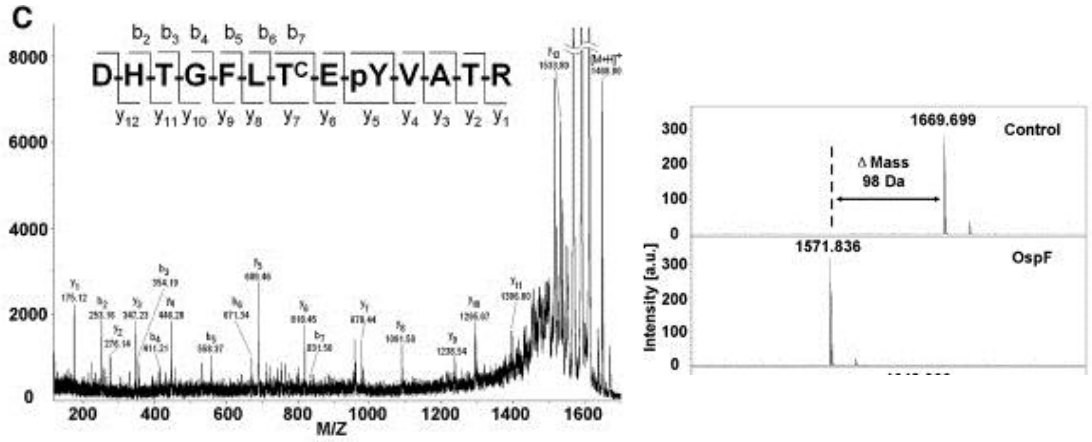

Figure 4. Mass Spectra indicating the loss of 98 Da as shown by Shao et al.[7] The 98Da could not be attributed to dephosphorylation by either a PTP or a DuSP.

To identify the product formed post reaction they subjected the reaction product to NMR spectroscopy. Then NMR spectrum of the product (Figure 5), when compared to that of

the starting material, showed a new vinylic signal. Kuipers et al.[11] had attributed this signal to methyl dehydroalanine, formed by dehydration of threonine.

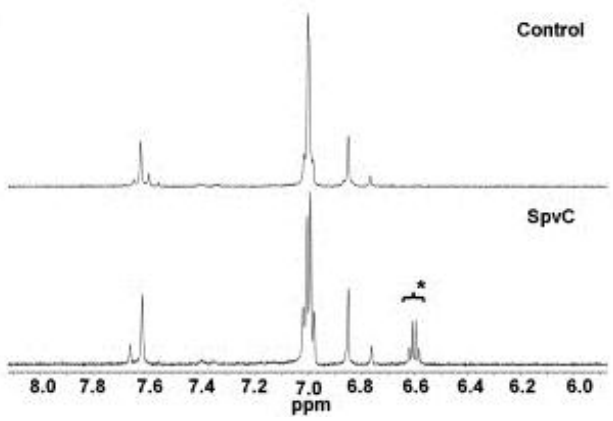

Figure 5. NMR spectra comparing peptide isolated following incubation with SpvC and a control experiment lacking the enzyme. The vinylic signal has been attributed to ß-methyldehydroalanine.[11]

To test if their conclusion about the new vinylic product was correct, they incubated the product with a kinase. It was seen that the synthetic peptide was not rephosphorylated on the threonine residue.

Based on the evidence gathered, they concluded that SpvC/OspF was irreversible dephosphorylating Erk2 by a unique elimination mechanism. Their proposed mechanism is schematically shown in Figure 6. A general catalytic base abstracts a proton from the alpha carbon of the threonine resulting in the formation of a double bond and the cleavage of the carbon-oxygen bond linking threonine to the phosphoester.

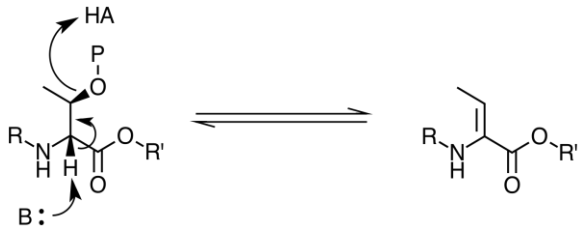

Figure 6. Reaction mechanism put forth by Shao et al. An elimination reaction resulting in a loss of inorganic phosphate ($H_2PO_4^-$ - 97 Da) and a proton (1 Da) which is abstracted from the alpha carbon atom.

2.2.2 X-ray crystallography and kinetic studies point to K136 and H106 as the catalytic acid and base of SpvC, respectively

To identify the key residues involved in the catalysis and the role of the phosphotyrosine on the substrate, the authors carried out crystallographic studies. X-ray crystallography showed a unique tertiary structure of SpvC, shown in Figure 7.

Figure 7. Ligand-free structure of spvC WT (PDB ID: 2P1W) showing a tertiary structure similar to a cupped hand.

The structure resembled a cupped hand with two highly electropositive centers on the concave side. A structure with a synthetic peptide bound showed that the two phosphorylated resides bound to these electropositive cavities, as shown in Figure 8.

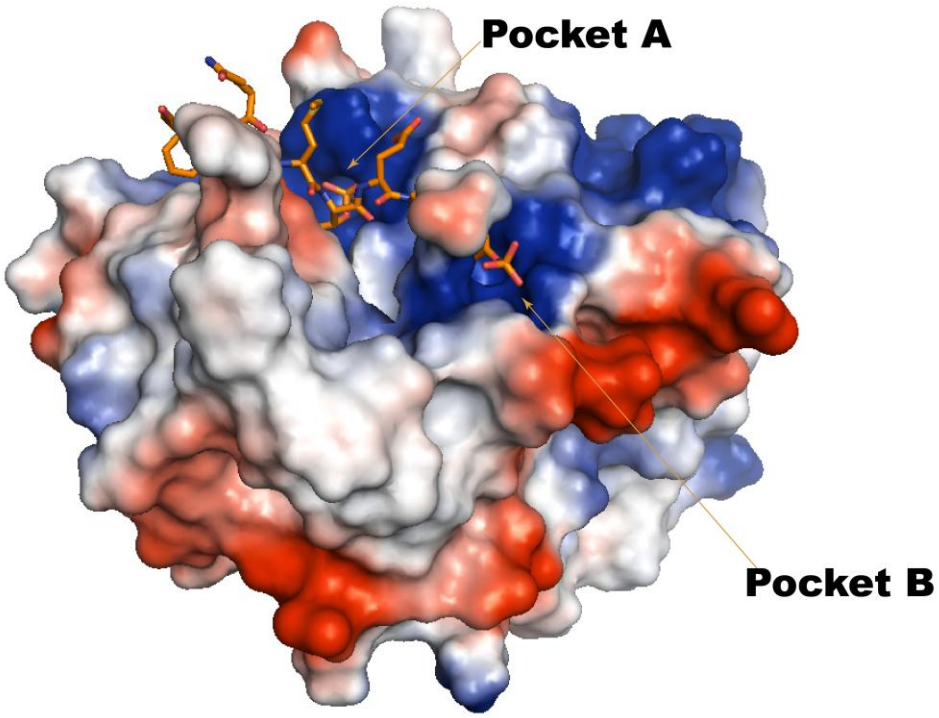

Figure 8. Electrostatic potential surface of SpvC WT with synthetic peptide (PDB ID: 2Q8Y). Pocket A & Pocket B are two highly electropositive pockets that bind the phosphothreonine and phosphotyrosine respectively.

The group used an alanine screen to identify the role of the residues present in the two electropositive pockets. Residues of OspF were mutated to alanine or arginine one by one and the activity of these mutants were tested with Erk2 substrate. Using Western Blot and antigens specific to Erk2 and pT-Erk2, the authors identified the specific mutants that were inactive or extremely slow to catalyze the reaction. The blot obtained by them, shown in Figure 9, showed

that the mutants K102R, H104A and K134A were inactive. The corresponding residues of SpvC were identified as K104, H106 and K136.

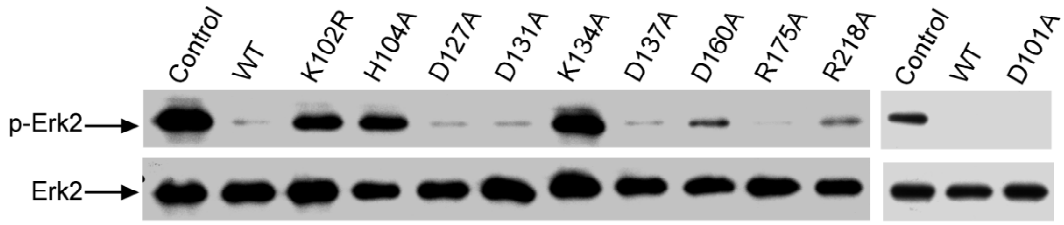

Figure 9. Immunoblot assay of WT OspF and various mutants with Erk2 substrate. An alanine screen was used to identify key catalytic residues.[8] It was seen that K102, H104 and K134 were inactive or extremely slow in catalyzing the reaction.

To identify if the mutations had resulted in loss of binding affinity, the authors carried out a Maltose Binding Protein pull-down assay. A fusion protein of Maltose Binding Protein (MBP) with Erk2 was purified. The purified MBP-Erk2 was immobilized onto amylose beads. These beads were then incubated with OspF. The beads were then washed extensively, and the flow-through of these washes was subjected to SDS-PAGE and coomassie staining to qualitatively detect if OspF was bound to the MBP-Erk on the beads. Results obtained by the group are shown in Figure 10. Lane 3 and 5 are inputs and pull-down for wild type OspF respectively. Lane 6-8 and Lane 11-13 are inputs and pull-down for OspF K102R, H104A and K134A mutants, respectively.

Based on the results obtained from the MBP pull down, they concluded that these mutants were capable of binding to the substrate. To identify the roles of these residues in catalysis the authors carried out kinetic studies to identify the pH rate profile. They found that the WT SpvC shows a bell shaped pH rate profile with an optimum of pH 8.5.

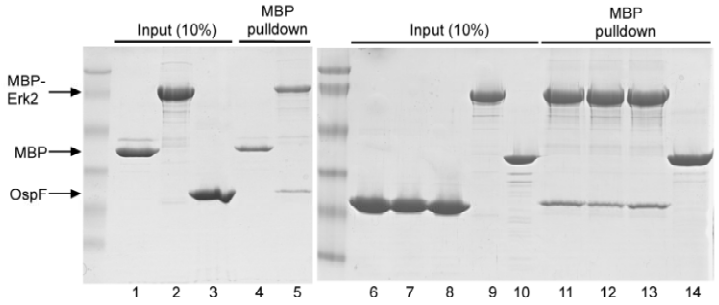

Figure 10. In-vitro interactions between recombinant Erk and OspF WT and OspF mutants. 3 and 5 are inputs and pull-down respectively for OspF WT, Lane 6-8 and 11-13 are input and pull-down of OspF mutants K102R, H104A, K134A respectively.[8] Pull-down shows that the mutants bind to the substrate with similar affinity as the wild type.

The acidic group was determined to have a pK_a of 7.6. Through mutational and kinetic studies of the H106 and Y158 residues, it was determined that the K136 residue was responsible for the acidic limb. Mutation of H106 to lysine resulted in the loss of the basic limb of the curve, implying that the H106 was the proton donor. The pK_a of H106 was determined to be 9.4; the elevated values was ascribed to hydrogen bonding between Nε and the D201, which increases the basicity of the Nδ by stabilizing the conjugate acid form. Mutation of the Y158 did not inactivate the enzyme completely, indicating that this residue was not involved in the catalytic reaction. This, along with the mutational studies of the K104 and H106 residues, indicated that the K136 residue behaves as the base. K136 abstracts the proton from the alpha carbon of phosphothreonine. The K104 residue was deemed responsible for H-bonding with the substrate backbone amide carbonyl, thus reducing the pKa of the proton on alpha carbon of phosphothreonine.

The authors also determined a crystal structure of SpvC K136A bound to a synthetic peptide. The peptide sequence derived from Erk2 contained the pT-X_{AA1}-pY motif in the sequence and was co-crystallized with the inactive SpvC K136A mutant. A stereo image of the structure is shown in Figure 11.

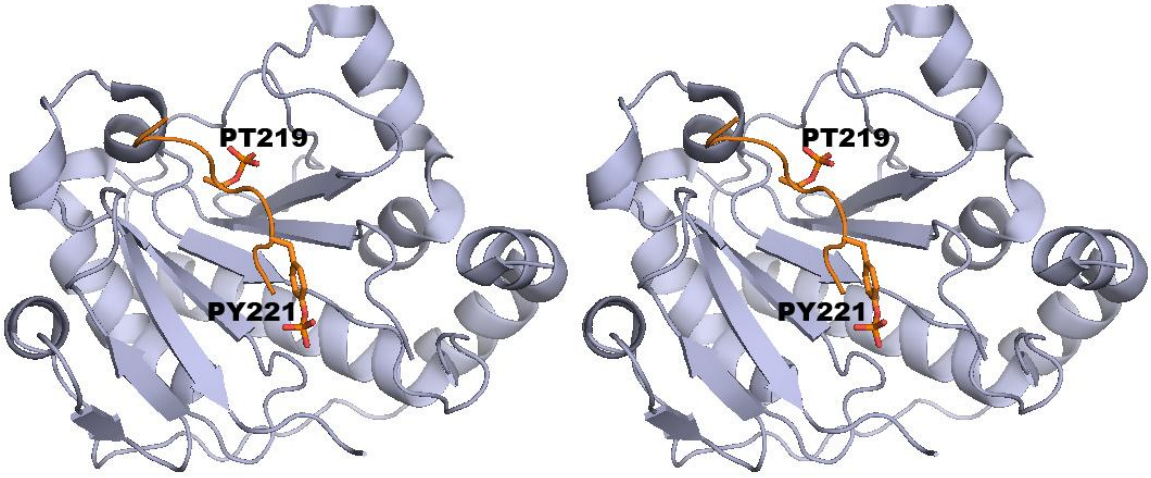

Figure 11. Stereo view of SpvC K136A, shown in light blue, bound to phosphopeptide, shown in orange, PDB ID: 2Q8Y.

Based on the peptide bound structure of SpvC K136A, the WT structure and the data obtained by the group through mutational and kinetic studies, Shao et al. have proposed the mechanism for the dephosphorylation carried out my OspF and SpvC, shown in Figure 12.

2.2.3 Structural data on phosphothreonine lyase class of enzymes is limited

A BLAST of the sequence of SpvC results in 4 proteins: SpvC (released by Salmonella entrica); OspF (released by Shigella flexneri); VirA (released by Chromobacterium violaceum); and FL6-60 - a putative virulence protein (released by Edwardsiella tarda). A multi-sequence alignment of these proteins is shown in Figure 13.

Structures obtained by various groups using X-Ray crystallography have been deposited in the Protein Data Bank. A summary of the structures obtained for members of the phosphothreonine lyase class of proteins is given in Table 1.

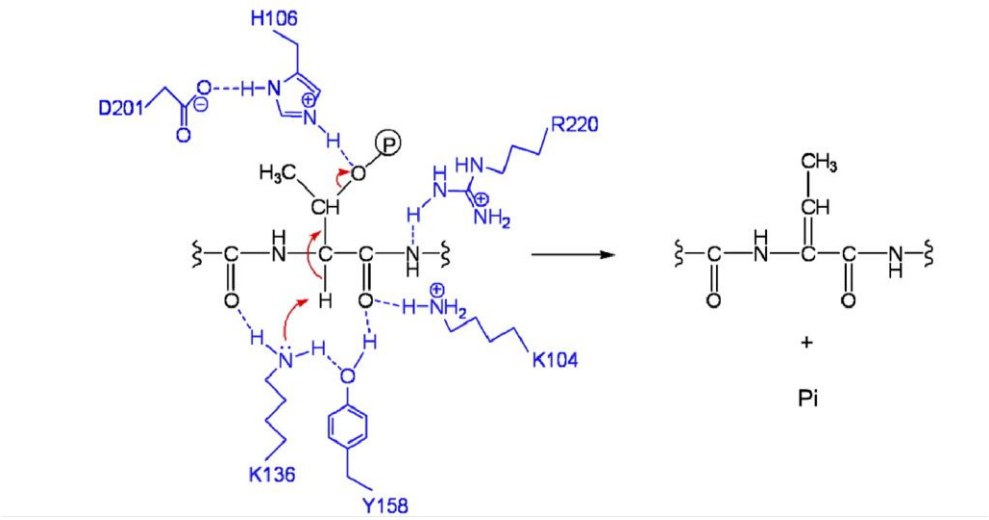

Figure 12. Proposed reaction mechanism of phosphothreonine lyase class of enzymes. The K136 behaves as the catalytic base that abstracts the proton followed by the H106 that donates a proton to the leaving phosphate group.

```
spvC            MPINRPNLNLNIPPLNIVAAYDGAEIPSTNKHLKNNFNSLHNQMRKMPVSHFKEALDVPD 60
VirA            ----------------------GPPQMSATNEDLKTNFHSLHNQMRQMPMSHFREALDAPD 39
Edwardsiella    ---------------------------MNESLKSNFDTLHRQMRQMPLSHFTVEPNAPD 32
OspF            ----------------------GPQMLSANERLKNNFNILYNQIRQYPAYYFKVASNVPT 38
                                        *: **.**. *:.*:*: *  :*    :.*

spvC            YSGMRQSGFFAMSQGFQLNNHGYDVFIHARRESPQSQGKFAGDKFHISVLRDMVPQAFQA 120
VirA            YSGMRQSGFFAMSQGFQLESHGGDVFMHAHRENPQCKGDFAGDKFHISVQREQVPQAFQA 99
Edwardsiella    YSGIRQSGFFALSQGFRLSNASDDFFIHARREAPQYRGEFIGDKFHISVQEQQVAQAFQA 92
OspF            YSDICQS-FSVMYQGFQIVNHSGDVFIHACRENPQSKGDFVGDKFHISIAREQVPLAFQI 97
                **.: ** * .: ***:: . . *.*:** ** ** :*.* *******: .: *. ***

spvC            LSGLLFSEDSPVDKWKVTDMEKVVQQARVSLGAQFTLYIKPDQENSQYSASFLHKTRQFI 180
VirA            LSGLLFSVDSPIDKWKVTDMERVDQQSRVAVGAQFTLYVKPDQENSQYSASSLHNTRQFI 159
Edwardsiella    LSGLLFSEDSPIDKWKVTDMARVDQQSRVGEGAQFTLYVKPDRGDSQYSATALHKTRQFI 152
OspF            LSGLLFSEDSPIDKWKITDMNRVSQQSRVGIGAQFTLYVKSDQECSQYSALLLHKIRQFI 157
                ******* ***:****:*** :* **:**. *******:*.*:  *****  **: ****

spvC            ECLESRLSENGVISGQCPESDVHPENWKYLSYRNELRSGRDGGEMQRQALREEPFYRLMT 240
VirA            ECLESRLSESGLMPGQYPESDVHPENWKYVSYRNELRSGRDGGEMQSQALREEPFYRLMA 219
Edwardsiella    ASLESRLSEQGIIPGRVPESDVHPDSWRYISYRNELRSERGGGEMQSQALREEPFYRLMT 212
OspF            MCLESNLLRSKIAPGEYPASDVRPEDWKYVSYRNELRSDRDGSERQEQMLREEPFYRLMI 217
                 .***.* .. : .*. * ***:*:.*:*:******** *.*.* * * **********

spvC            E 241
VirA            E 220
Edwardsiella    E 213
OspF            E 218
                *
```

Figure 13. Sequence alignment of SpvC, VirA, OspF and Edwardsiella putative virulence protein shows highly conserved residues (marked by '*'), specifically the <u>catalytic acid H106</u>, <u>base K134</u> and several residues involved in binding of substrate

Table 1. Published crystal structures of enzymes comprising the phosphothreonine lyase class. Shown are the PDB ID's, the resolution of the structures, the space group and the ligands.

Protein	PDB ID	Resolution (A)	Space group	Ligand
SpvC WT [7]	2P1W	2.30	$P4_2\ 2_1\ 2$	None
SpvC WT[12]	2Z8M	2.00	$P2_1\ 2_1\ 2_1$	None
SpvC WT [12]	2Z8N	1.80	$P2_1$	Sulfate Ion
SpvC WT [12]	2Z8O	2.40	$P4_1\ 2\ 2$	Tartaric Acid
SpvC K136A [12]	2Z8P	1.80	$P2_1\ 2_1\ 2_1$	Dually phosphorylated synthetic peptide
SpvC K136 A [7]	2Q8Y	2.0	$P2_1\ 2_1\ 2_1$	Dually phosphorylated synthetic peptide
OspF WT [13]	3I0U	2.70	$P4_1\ 2_1\ 2_1$	(4S)-2-Methyl-2,4-Pentanediol
VirA WT [14]	3B06	1.40	$P2_1$	None

2.2.4 The mechanism of the elimination reaction is yet to be determined

While the efforts of the above mentioned groups have allowed us to decipher the activity of these enzymes, what is yet to be answered in the mechanism used by these enzymes is uncertain. Previous studies of elimination reactions catalyzed by enzymes have shown that the preferred mechanism is E1cB, involving the formation of a carbanion intermediate. Considering the novelty of these enzymes and the reaction catalyzed by these enzymes, SpvC/OspF may or may not follow this mechanism; hence other possibilities, shown in Figure 14, have to be considered.

Enzyme catalysis involving elimination has been observed in various biological systems. One such enzyme is serine/threonine ammonia lyase, commonly known as serine/threonine dehydratase. These are classified as PLP-dependent enzymes due to the involvement of the PLP co-factor in catalysis of the reaction. This is unlike phosphothreonine lyases, wherein the

enzyme does not utilize any co-factor to dephosphorylate the substrate. Nonetheless, the reactions catalyzed by both enzyme classes involve the abstraction of a proton to form a vinyl group along with elimination of a leaving group. In case of the serine/threonine dehydratase, the leaving group is water.

β-Elimination can be carried out by any one of the mechanisms listed below:

1. Concerted – where in the proton abstraction and the phosphate elimination takes place simultaneously, shown in Figure 14.1.
2. E1cB – Reaction involving formation of a carbanion by abstraction of proton, shown in Figure 14.2. This reaction mechanism is inherently stepwise and two possibilities arise:
    i. $E1cB_R$: The proton abstraction is reversible, which implies the elimination of the phosphate is slow and hence the rate determining step. This indicates that k_{-1} is comparable or higher than k_2. Hence the intermediate will pick up a proton to return to the starting material.
    ii. $E1cB_F$: The proton abstraction is the rate-determining step, following which the elimination of the phosphate occurs rapidly.
    iii. This implies that k_1 is significantly lower to k_2, hence any intermediate formed converts to product.
3. E1 – Where in the phosphate group leaves first creating a carbocation followed by abstraction of the proton, shown in Figure 14.3.

In the case of concerted and E1cB mechanisms, it is possible that the elimination of the phosphate is irreversible. In order to account for all possibilities we depict the reaction to be

reversible. This report presents the body of work carried out to identify if this class of enzymes follows an E1cB mechanism involving the reversible formation of a carbanion at the active site.

Analytical techniques like isotope exchange when coupled with crystal structures of the protein concerned provide more details of the reaction involved.[15] It is essential that the mechanism of the beta-elimination be deduced so as to elucidate the transition state of the substrate-enzyme complex. This, in turn will aid inhibitor design for the enzymes.

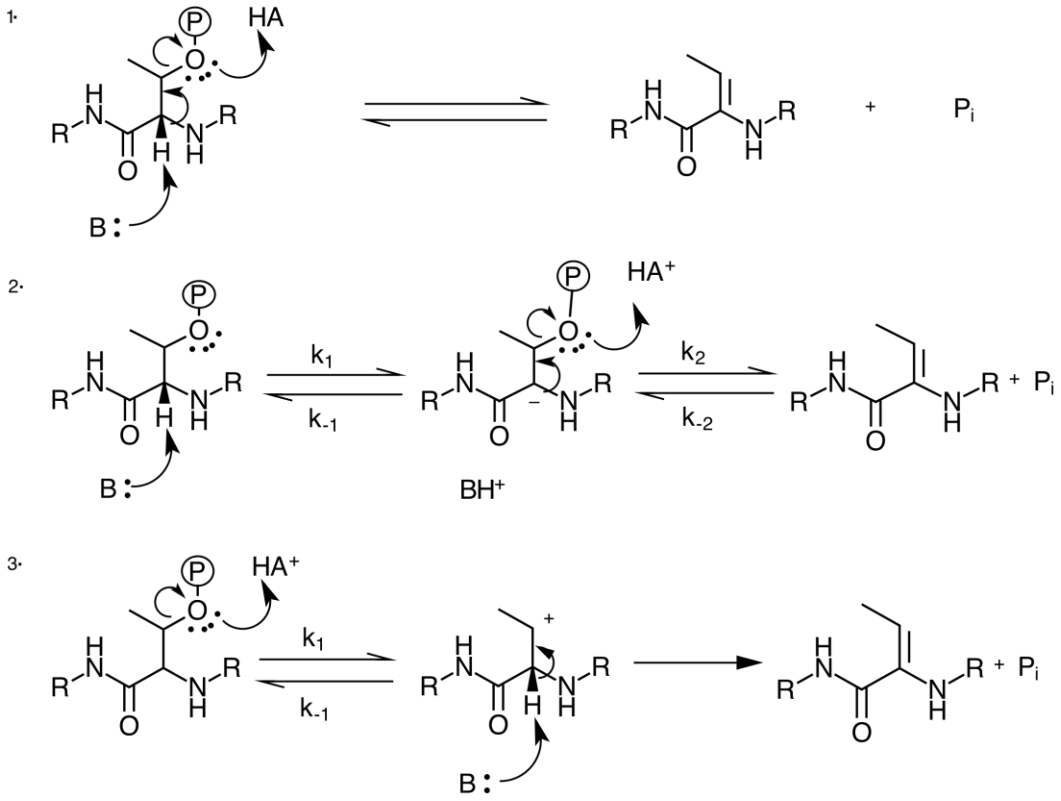

Figure 14. Possible mechanisms for elimination reactions shown herein schematically. Concerted (14.1), E1cB (14.2) and E1 (14.3) shown herein with phosphoester as starting material resulting in vinylic product and inorganic phosphate as products.

To determine if the enzymes use the E1cB mechanism, we chose to use deuterium isotope exchange experiment. Deuterium isotope exchange experiments are used to identify the

rate-determining step in multi-step reversible elimination reactions wherein proton abstraction is involved. The incorporation of an isotopically labeled atom from the solvent into the substrate is often used as a method to decipher the mechanism. We used this technique to identify whether a carbanion intermediate is formed in the mechanism used by SpvC to catalyze the elimination of phosphate. This involved carrying out the reaction in a deuterated solvent followed by isolation of the substrate peptide. This peptide would undergo M.S. analysis to determine incorporation of deuterium on the c-alpha carbon.

To carry out the above experiment, it was necessary that any carbanion intermediate formed at the active site did not convert to the vinylic product. Creating a mutant that could not protonate the leaving group would provide an enzyme in which any carbanion created at the active site would revert back to its protonated state by equilibrating with the solvent. Thus we chose the H106N mutant, which was earlier proven to be inactive.[7] The H106N mutant lacked the histidine that was determined to be the general acid, donating a proton to the leaving phosphate group. The work presented herein involves structural studies of this mutant to determine if the mutation resulted in structural perturbations that could account for the lack of activity.

CHAPTER 3

MATERIALS AND METHODS

3.1 Genomics – cloning and site directed mutagenesis

Dr. Feng Shao (of National Institute of Biological Sciences, Beijing, China) graciously provided us with the plasmids for MBP-OspF and MBP-SpvC, containing a GST tag. SpvC was cloned into a variant of prk793 vector such that the final sequence resulted in 10x histidine tagged SpvC with TEV cleavage site in between.

The gene for SpvC was amplified to generate multiple copies of the original gene sequence using the primers SpvC Forward (SF) and SpvC Reverse (SR), given in Table 2. These primers also incorporated a TEV cleavage site before and a BAMHI restriction site at the end of the SpvC sequence respectively. The product of this PCR was isolated using electrophoresis and Qiagen QIAquick PCR kit. A second PCR was carried out to insert a KpnI restriction site upstream of the TEV cleavage site. The PCR product was digested using KpnI and BamHI restriction enzymes (Fermentas) and ligated into a modified prk793 vector containing 10xhistidine tag upstream of the maltose binding protein (MBP) domain. The plasmid from this ligation – His-TEV-SpvC-prk793 (HP-SpvC) was purified using the protocol and materials provided in the Qiagen QIAquick PCR Purification Kit (Qiagen).

The *HP-SpvC* plasmid was used to transform DH5-α strain of *E-coli* cells and plated on sterilized agar media plates with ampicillin and chloramphenicol. These were incubated at 37˚C overnight. Colonies from the plate were selected to inoculate 10mL aliquots of LB-media with aforementioned antibiotics. These aliquots were placed in 37˚C shakers overnight. The cells were pelleted and the plasmids extracted using the protocol and materials provided in QIAprep

Spin Miniprep kit. The solution containing the plasmids was aliquoted into eppendorf tubes and stored at -20˚C for future use.

Table 2. Primers used to clone *spvC* into the prk793 variant and carry out site directed mutagenesis, mutating H106 to N106. The mutation (H106N) highlighted in red.

Primer Name	Primer Sequence (5' to 3')	Description
SpvC Forward (SF)	GAAAACCTGTATTTTCAGGGCATGCCCATAAATAGGCCTAATC	Forward primer used to amplify the SpvC gene.
SpvC Reverse (SR)	CGCGGATCCTTACTCTGTCATCAAACGATAAAACGG	Reverse primer used to amplify the SpvC gene.
SpvC KpnI (SKF)	CGGGGTACCGAAAACCTGTAT	Primer used to insert the KpnI restriction site.
SpvC SDM Forward (SSF)	GCCGGTGACAAGTTCAACATCAGTGTGCTCAGG	Forward primer used in site directed mutagenesis
SpvC SDM Reverse (SSR)	CCTGAGCACACTGATGTTGAACTTGTCACCGGC	Reverse primer used in site directed mutagenesis.

The plasmids were used to transform BL-21 codon+ strain *E-coli* cells. Sterilized baffled flasks containing 10mL of LB media, containing 100mg/mL of ampicillin and chloramphenicol each, were inoculated using the transformed cells. These were shaken at 37˚C for 18 hours until an optical density at 600nm (OD_{600}) of 1.5 to 1.8 was achieved. These growths were then mixed with 50% sterilized glycerol solution, in 4:1 ratio, and frozen at -80˚C for future use.

The mutant SpvC H106N was created by site directed mutagenesis (SDM) using the protocol and materials provided in the Quick Change Lightning Kit (Agilent Technologies). The vectors SSF and SSR were used to insert the mutation. The plasmids were purified by Qiagen QIAquick PCR kit. Copies of the HP-SpvC-H106N plasmid were created using E-coli DH5-α cells, isolated and sequenced to confirm correct mutation.

The plasmids were used to transform BL-21 codon+ strain E-coli cells with ampicillin and chloramphenicol resistance. Sterilized baffled flasks containing 10mL of LB media, containing 100mg/mL of ampicillin and chloramphenicol each, were inoculated using the transformed cells.

These were shaken at 37˚C for 18 hours until an optical density at 600nm (OD_{600}) of 1.5 to 1.8 was achieved. These growths were then mixed with 50% sterilized glycerol solution, in 4:1 ratio, and frozen at -80˚C for future use.

3.2 Purification of the enzymes SpvC and SpvC H106N

Four liters of 2x YT media was prepared in baffled flasks and sterilized by autoclaving. This was then inoculated E-Coli BL21 codon+ strain, transformed with the HP-SpvC-WT or HP-SpvC-H106N plasmid. The broth was shaken at 37˚for up to 16 hours such that the optical density of the broth was noted at 1.5 OD_{600}.

The broth was then induced with stock solution of 100 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) such that final concentration of IPTG was 1mM. At this point the broth was moved from 37˚C shakers to room temperature (r.t.) shakers and left overnight (up to 12 hours) to produce a good yield of the protein. The cells were then pelleted by centrifugation in a Fiberlite* F10-4x1000 LEX Rotor (Thermo Scientific), mounted on Sorvall RC-6 Plus Superspeed centrifuge, at 5000 rpm (9800 g) and stored at -80˚C for future use. The buffer solutions used to purify His-SpvC were prepared using Nano-purified water (desalinated water) and adjusted to the required pH at 4˚C. Each buffer was filtered through a Milipore 20µm filter before use.

The frozen stock of cell from -80˚C was thawed and warmed to 0˚C and re-suspended in the lysis buffer (pH 8.0 tris(hydroxymethyl)aminomethane (Tris-HCl) 20mM, Imidazole 20mM, NaCl 500mM, Glycerol 5% and ß-Mercaptoethanol 5mM) along with protease inhibitors aprotinin, leupeptin and pepstatin. The cells were broken by sonication and the debris was pelleted by centrifugation in SS-34 rotor (mfg. Thermo Scientific) mounted on a Sorvall RC-6 Plus

Superspeed Centrifuge, at 20,500 rpm (50,228 g) for 45 minutes. The supernatant was then incubated with 15 mL nickel resin (GE Nickel Sepharose High Performance) for a period of 2 hours at 4˚C and placed on a rotary shaker. This was followed by separating the resin from lysate by filtration and re-suspension of the nickel resin in lysis buffer. The resin was washed with 4 – 6 bed volumes of lysis buffer to ensure removal of non-specifically bound proteins and debris.

The elution was carried out by running a gradient of elution buffer (pH 8.0, Tris-HCl 20mM, Imidazole 500mM, NaCl 500mM, Glycerol 5% and ß-Mercaptoethanol 5mM) mixed with decreasing amounts of lysis buffer over a span of 120mL. The fractions were selected based on purity of the protein contained, determined by SDS-PAGE. The fractions were combined and loaded into 10kD dialysis tubing, which was suspended in TEV-cleavage buffer (pH 8.0, Tris-base 50mM, EDTA 0.5mM, Dithiothreitol (DTT) 1mM, Glycerol 5%), for 4 hours. The buffer (TEV-cleavage) was changed at regular intervals to eliminate imidazole from the combined fractions.

TEV-protease, an enzyme efficient in cleaving fusion proteins at sites containing the E-X_{aa}-X_{aa}-Y-X_{aa}-Q-(G/S) sequence, was used to cleave the His-tag from SpvC. During the cloning stage, this cleavage site was inserted at the N-terminus of the protein that the final SpvC sequence was left with minimal residues from the tag and the cleavage site. The dialysis tube containing the combined fractions and TEV-protease was incubated in the TEV-Cleavage buffer for approximately 8 hours. Completion of the cleavage was checked using SDS-PAGE.

After the total cleavage of the fusion protein, it was necessary to separate the cleaved His-tag and TEV protease from SpvC. The contents of the dialysis tube were incubated with clean Ni-resin for 2 hours. SpvC, now in the flow through, was collected along with the washing of the Ni-Resin. SDS-PAGE was carried out to confirm the absence of the cleaved His-tag in the flow-through.

The pure SpvC was then concentrated using 10kD conical tube concentrators. The final concentrate, 3-5 mL of clear solution containing pure SpvC, was used for 120mL Superdex Gel-Filtration. The gel-filtration was carried out using the Superdex buffer (pH 8.5, Tris-HCl 50mM, NaCl 150mM, Glycerol 5%), and the fractions collected. This provided SpvC in purity exceeding 95%, which was confirmed by SDS-PAGE.

SpvC H106N mutant was purified using the same protocol as the WT. It was observed that the purification of the mutant using the above protocol afforded similar yields and purity exceeding 95%, confirmed using SDS-PAGE.

3.3 Enzymology – assay protocol and materials

Earlier work [7] had shown that the histidine residue at position 106 functions in the catalytic reaction as the acid, donating a proton to the leaving inorganic phosphate moiety. We first tested the activity of the WT enzyme to confirm the integrity of the purified protein and to test the optimum reaction conditions. To do this, we incubated the enzyme at various pH values with a synthetic peptide of the sequence H-Glu-Met-pThr-Gly-pTyr-Val-Ala-OH. The release of phosphate was measured by using a colorimetric assay.[16]

The peptide was dissolved in the reaction buffer containing 50 mM Tris and 1 mM DTT pH 8.5. The same buffer was used to prepare a series of peptide dilutions in the range of 5-210 μM. 90 μL of substrate peptide were aliquoted in the wells of flat bottom 96 well plate (Costar, Fisher) and 10 μL of SpvC WT were added to the final concentration of 9.4 nM to start the reaction. After 60 seconds the reaction was quenched by addition of 230 μL of malachite green assay reagent. After 15 minutes the absorbance at 625 nm was measured and converted to inorganic phosphate concentration via a standard curve obtained using a phosphatate standard

diluted with reaction buffer. Blank readings obtained from wells to which 10 µL of water were added in place of the enzyme were subtracted to account for non-enzymatic phosphate release. Calculated rates were plotted against peptide concentrations and fitted to the Michaelis-Menten equation to obtain kinetic parameters k_{cat} and K_M of 3.2 ± 0.16 sec^{-1} and 52 ± 6.0 µM, respectively.

Having confirmed the activity of the WT, site directed mutagenesis was carried out successfully resulting in the H106N – an inactive mutant of SpvC. The purification strategy for the H106N mutant was same as the WT, resulting in protein of 95% purity and higher.

To test for the E1cB mechanism, a deuterium isotope exchange experiment was carried out. The synthetic peptide was incubated with SpvC H106N at pH 8.5 overnight in deuterated water based buffer. This was then treated with trifluoroacetic acid to denature and precipitate the enzyme. An eppendorf concentrator 10 kDa pore size was used to remove the enzyme from the quenched reaction mixture. The filtrate was then lyophilized overnight and then resuspended in H_2O. The same was carried out for a blank, wherein no enzyme was added to the reaction mixture. M.S. analysis of the isolated peptide showed that the experiment did not result in deuterium exchange on the α-carbon atom of the phosphorylated threonine.

3.4 Crystallization of SpvC WT & H106N mutants

The purified SpvC WT was first concentrated to 30 mg/mL (extinction coefficient of 23045 cm^{-1} M^{-1} calculated using ProtPram of Expasy Tools [17]), confirmed by Micro-Volume Full-Spectrum Fluorospectrometer (Mfg. by NanoDrop Products). This concentrated protein was then used to manually set up vapor diffusion sitting drop trays by hand. A grid of 6x4 with varying concentrations of PEG and Ammonium Chloride was created to screen for the optimum

conditions to crystallize the WT protein. The trays were set and maintained at room temperature (289 K). Crystals appeared after a week and were of square bi-pyramidal shape. These were frozen using liquid nitrogen after soaking the crystals in the well conditions containing increasing quantity of glycerol (up to 15%). The optimum conditions for the WT that afforded reproducibility were determined as following: Protein to Well Condition ratio of 1:1, 20% PEG 3350 (a 50% w/v solution), 0.1M Ammonium Chloride (1M stock solution).

The mutant SpvC H106N was concentrated to 14 mg/mL of solution. A robot was used to screen for various conditions using kits produced by various companies like Hampton Research, Molecular Dimensions, etc. The Morpheus Protein Crystallization Screen[18] manufactured by Molecular Dimensions provided the conditions required to crystallize the protein, which was further optimized to afford reproducibility. Crystals of monoclinic shape were obtained which were frozen using liquid nitrogen after soaking the crystals in the well conditions containing increasing quantity of glycerol (up to 15%). The conditions were determined to be as follows: Protein to Well Condition ratio of 3:1, 20% PEG 550 MME, 10% PEG 20000 (30% w/v solution), 0.1M Amino Acids Mix (1M stock solution mfg. by Molecular Dimensions), 0.1M Buffer System 1 [18] at pH 6.5.

3.5 Data collection and processing

Data for the WT crystal were obtained by using home source generator and detector (Rigaku RU-200/Raxis IV++). Data was indexed and processed using d*TREK[19] in the suite Crystal Clear. Molecular replacement was carried out using AutoMR / Phaser [20] of the Phenix suite.[21] The search model was the 1.4Å structure of SpvC (Protein Data Bank (PDB) entry 3BO6) complexed with sulfate ion. Refinements were carried out with Phenix Refine tool.[22] The data

for the H106N crystal was collected at Stanford Synchrotron Radiation Lightsource (SSRL). Data processing was carried out on SSRL servers followed by molecular replacement using AutoMR / Phaser [20] of the Phenix suite.[21] Refinement was carried out using Phenix Refine.[22] TLS parameters were selected using the TLS Motion Determination server.[23] In both cases Coot[24] and MolProbity[25] were used to carry out model building and validation. The crystallographic data and the statistics are given in Table 5 (SpvC WT) & Table 6 (SpvC H106N). All figures depicting SpvC crystal structures were prepared using PyMol.[26] The structures of SpvC WT and H106N mutant have been submitted to PDB under the ascension IDs 4HAH and 4H43, respectively.

CHAPTER 4

RESULTS

4.1 Crystallography

The two structures obtained have been submitted to the Protein Data Bank (PDB) under the ascension IDs 4HAH (SpvC WT) and 4H43 (SpvC H106N mutant). The wild type was crystallized under conditions previously published,[7] and then further optimized for reproducibility. The SpvC H106N mutant was initially crystallized using Morpheus Protein Crystallization Screen (mfg. Molecular Dimensions) with the aid of ARI Crystal Gryphon robot. The two structures have been compared to previously published structures to determine the structural integrity of the mutant. The statistical data for SpvC WT (4HAH) and SpvC H106N (4H43) have been presented in Table 3.

4.2 Structural aspects of SpvC WT (4HAH) as compared to the previously published structure (2Z8M)

The structure used to verify the integrity of the crystallized SpvC WT was a 2Å ligand-free structure (PDB ID: 2Z8M).[12] The structures were overlapped in PyMol[26] to compare the tertiary and secondary structures of the two ligand-free proteins. As shown in Figure 15, the tertiary structure of the two proteins closely resembled each other in terms of positions of loops, helices and β-sheets. PDBeFOLD[27] was used to compare the structures by assessing the overlap of the backbone and the side chains of all the amino acids present in each of the structures. The result (RMSD = 0.16) is presented in terms of root mean square deviation (RMSD). The lower the RMSD, the better the overlap of the two structures.

Table 3. Data collection and refinement statistics for SpvC WT.

Values in brackets correspond to those in outer resolution shell.

Data Collection	***SpvC WT***	***SpvC H106N***
Beamline	Homesource	SSRL
Wavelength (Å)	1.5418	0.97591
Resolution Range (Å)	32 – 1.80	50 – 2.3
Outer Shell (Å)	1.86 – 1.80	2.38 – 2.30
No. of Reflections		
Total	331429	405022
Unique	45887	35594
Average Redundancy	7.22 (4.35)	11.4 (11.5)
Mean I/ σ (I)	9.0 (2.5)	33.655 (4.4)
Completeness (%)	98.4 (86.8)	99.8 (99.5)
R_{sym} (%)	0.107 (0.485)	0.076 (0.614)
Space Group	$P2_12_12_1$	$P3_2$
Unit Cell Dimensions (a, b, c) (Å)	64.53 71.91 106.49	70.989 70.989 142.972
Refinement		
R_{work} / R_{free} (%)	0.1935 / 0.2310	0.1749 / 0.2161
Atoms in structure		
Protein	3496	3478
Water	468	191
Average B-factors (Å^2)		
Protein	22.01	46.88
Water	30.91	47.3
Root mean square deviation		
bonds (Å) / angles (degrees)	0.007 / 1.024	0.007 / 0.979
Protein geometry		
Ramachandran Outliers (%)	0.00	0.0
Ramachandran Favored (%)	99	98.6
Rotamer Outliers (%)	0.5	0.52
PDB Entry	4HAH	4H43

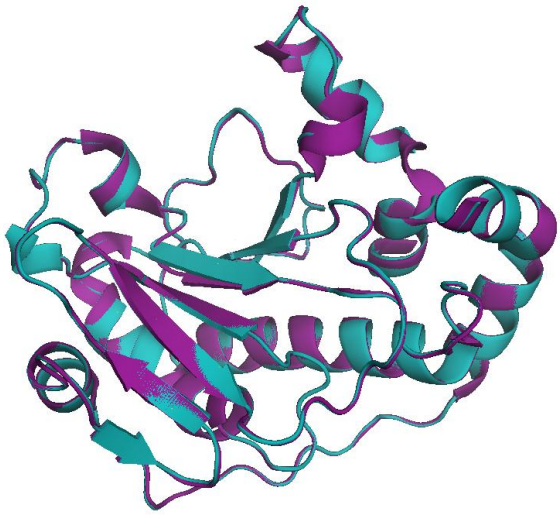

Figure 15. Tertiary structure of SpvC WT, PDB ID 2Z8M in purple, when compared to our structure of ligand-free SpvC WT, 4HAH in teal, shows that the structures are very similar, with a RMSD of 0.16 for 215 residues.

A close inspection of the structure 4HAH compared to 2Z8M showed that crucial residues important to catalysis are overlapping, as shown in Figure 16.

In addition to the catalytic site, SpvC has another binding pocket for the second phosphoryl group (pY) of the enzyme's native bis-phosphorylated substrates. The residues in this pocket responsible for binding phosphotyrosine were also found to have retained their 3-dimensional positions when compared to 2Z8M, as shown in Figure 17.

To support the structural analysis, a chemical assay was carried out to confirm the ability of the purified enzyme to dephosphorylate the substrate. This was done using a truncated peptide, containing the pT-X_{AA}-pY motif, H-Q-M-pT-G-pY-V-A-OH. The enzyme on incubation with the peptide at pH 8.5 resulted in release of inorganic phosphate, which was detected using phosphomolybdate assay.[16] Having confirmed the activity of the WT, we pursued the crystallization of the H106N mutant, previously found to be inactive[7] and used by us in a

deuterium isotope exchange experiment designed to probe the chemical mechanism. Because this experiment resulted in no deuterium isotope exchange, the goal of the structural study was to investigate if the mutation resulted in structural perturbations that would prevent the ability of the protein to bind to the substrate.

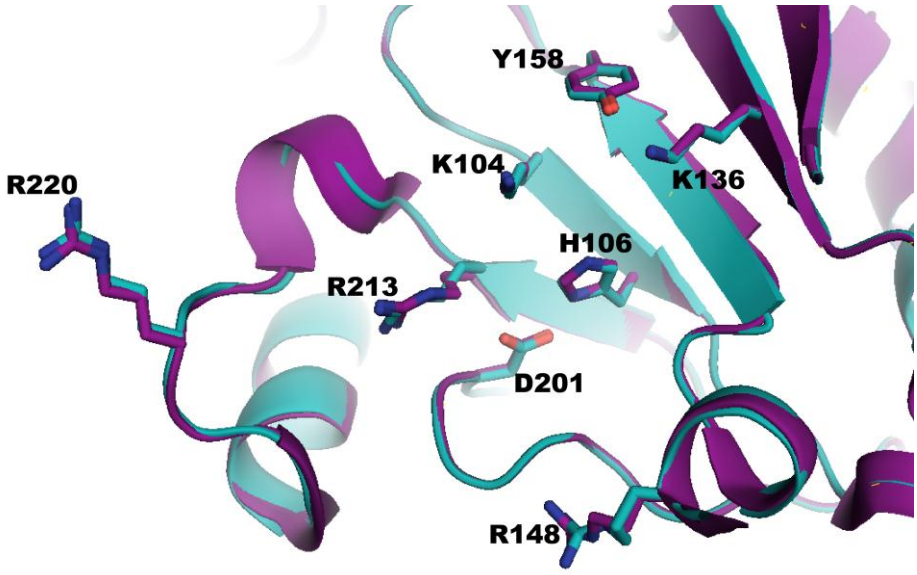

Figure 16. Comparisons of the positions of the active site residues of SpvC WT in the structure 4HAH (teal) obtained in this work, compared with a previously reported structure, 2Z8M, in purple. It is clearly seen that the residues from the two structures superimpose each other.

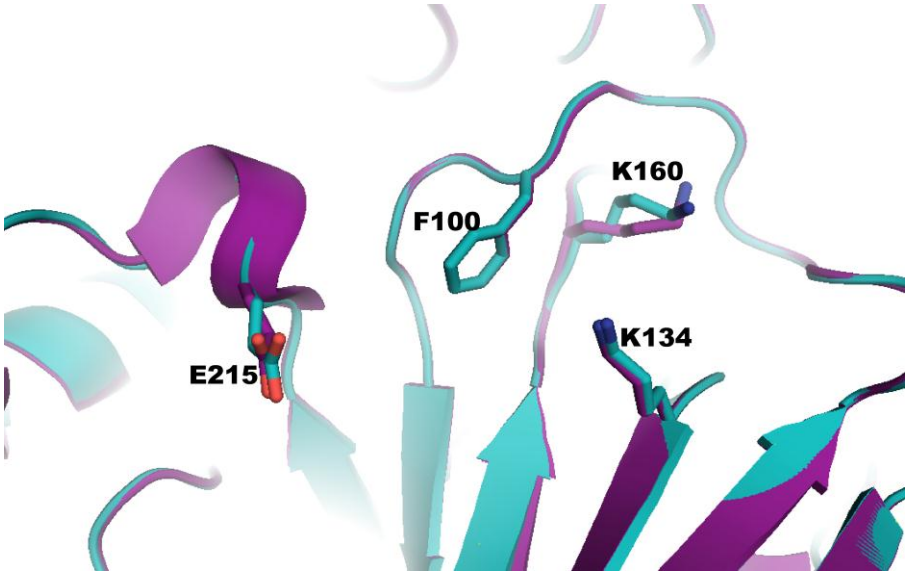

Figure 17. pY binding pocket residues compared between structures obtained through this work against previously published structures. Overlaid above are 4HAH, in teal, and 2Z8M, in purple.

4.3 Structural aspects of SpvC H106N (PDB ID: 4H43)

The H106N mutant has been reported previously, but to date no structure exists in the Protein Data Bank. In order to correctly interpret the results of mechanistic investigations with this mutant, an x-ray crystal structure of the protein was necessary. This would allow us to determine if the mutation of the wild-type protein resulted in structural perturbations that might affect the binding.

Sequencing confirmed the presence of the correct mutation following which the purified enzyme was used to obtain crystals in sitting-drop experiments set up by hand. PDBeFOLD[27] was used to compare this crystal structure, PDB ID 4H43, with the SpvC WT structure, PDB ID 4HAH. Calculations showed a RMSD of 0.33Å. Tertiary structures were compared in PyMol,[26] as shown in Figure 18.

The mutant was aligned with the WT to compare the positions of active site residues, as shown in Figure 19. The conclusion is that the mutation of the histidine residue on position 106 to asparagine does not affect the secondary or tertiary structure of the protein. This supports the conclusions reported from the kinetic assay work that the lack of activity of SpvC H106N is due to the mutation of the catalytic general acid, and not due to structural deformations induced by the mutation.

Figure 18. Shown above the structural alignment of SpvC WT (4HAH), in teal, on SpvC H106N (4H43) in green. The RMSD for 3D overlap of the two structures for 215 residues is 0.33.

The structural alignment of the WT and the H106N mutant was used to verify the position of residues involved in the two pockets which bind the phosphorylated tyrosine and threonine residues, shown in Figure 20. It was seen E215 was displaced by 1.8Å and R220 was displaced by 1.4 Å. It has been determined that the H215 residue forms H-bonds with the backbone of the substrate. This segment of the protein, a loop carrying the catalytically important residue R220, seems to be extremely mobile with a high B-factor as compared to the average B-factor of the whole structure.

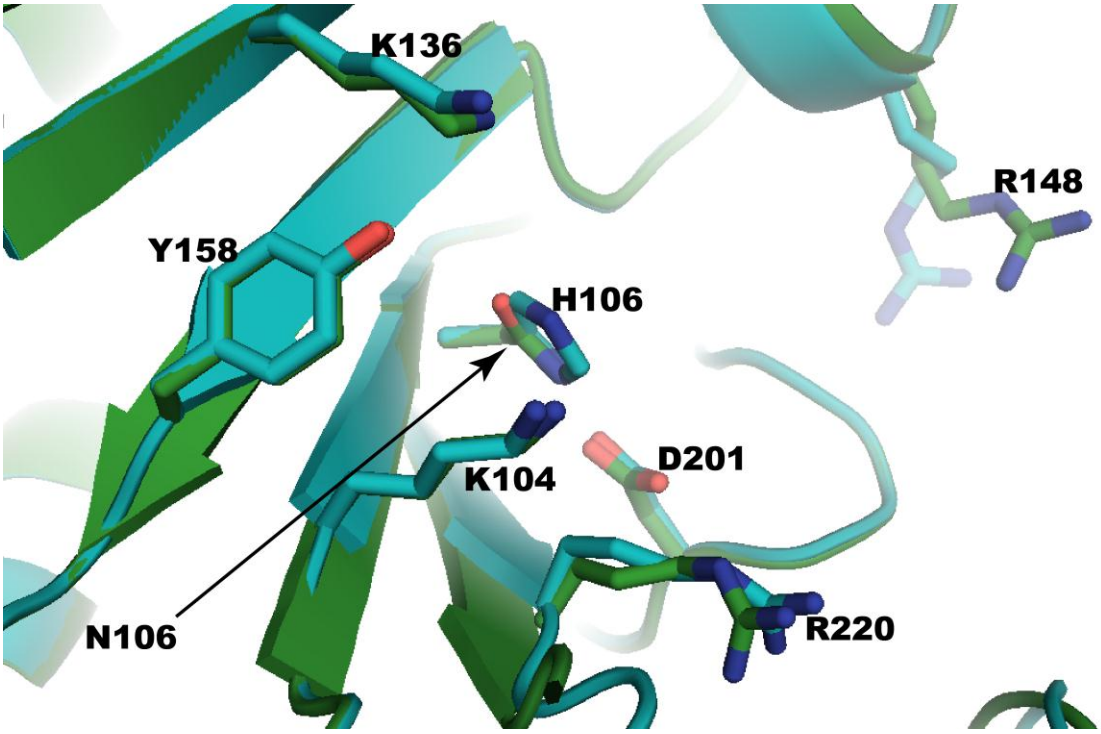

Figure 19. Active site of SpvC H106N (PDB ID 4H43), in green, structurally aligned to SpvC WT (PDB ID: 4HAH), in teal. The mutation H106 ➔ N106 has been labeled.

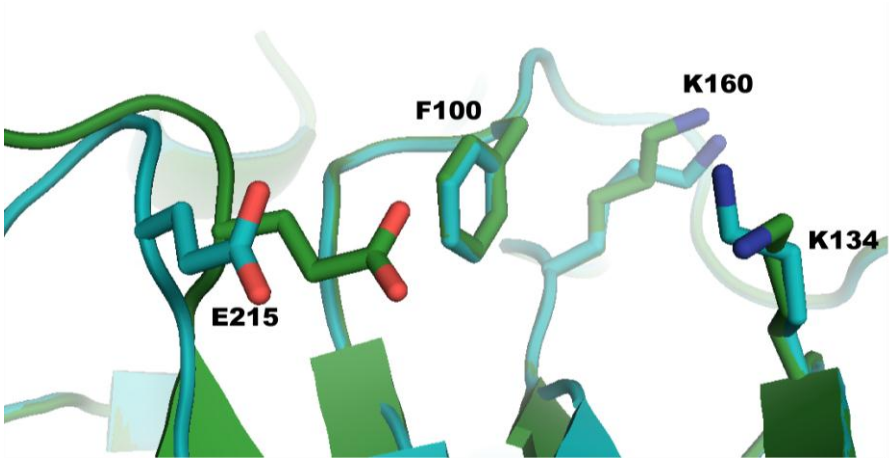

Figure 20. Recognition site of SpvC H106 (PDB ID: 4H43), in green, aligned with SpvC WT (PDB ID: 4HAH), in teal. This site is known to bind to the phosphorylated tyrosine on the substrate. Residue E215 is displaced by 1.8 Å, attributed to the mobile loop it is a part of.

B-Factors provide a numerical representation of the electron density detected at specific positions during the x-ray diffraction. High B-factor values imply high mobility of the residues, hence providing a reduced electron density. A study of the B-factors of the residues forming the R220 loop showed that the backbone and the side chains for the residues were significantly higher than the average B-factor of 46.9, as shown in Table 5. Average B-factors were calculated using StrucTools.[28] This implies that the loop carrying R220 and E215 residues is mobile.

Figure 21 shows that the loop carrying the two residues R220 & E215 undergoes a conformational change upon binding of the substrate. The R220 residue forms H-bonds with the phosphate moiety on the phosphothreonine of the substrate. The E215 backbone amide N-H forms H-bonds with the carbonyl of the backbone amide of the phosphotyrosine of the substrate backbone. The mobility of the loop allows a conformational change to occur upon substrate binding, bringing the respective residues in close proximity.

Table 4. Average B factors of side chains and backbone of several residues of SpvC H106N (4H43) that form the movable loop. It is seen that the residues, which form the loop, have high b-factors as compared to the overall b-factor of the protein (46.9).

Residue (Amino Acid)	Side Chain	Backbone
215 (E)	78.74	61.28
216 (L)	78.11	77.0
217 (R)	78.79	89.45
218 (S)	108.54	104.64
219 (G)	0.0	108.75
220 (R)	115.83	100.50
221 (D)	105.99	86.25

The respective positions of the movable loop in the WT and the H106N mutant are compared in Figure 22. It is seen that they closely resemble each other. A comparison using Secondary Structure Matching (SSM) [27a] showed that residues of the two loops had a Q-Score of 0.90, implying a close overlap of the two loops.

The structures of SpvC WT (PDB ID 4HAH) and SpvC H106N (PDB ID 4H43) closely resemble each other in the catalytic site, the pY binding site and the positions of the loop carrying the R220 and E215 residues. The H106N structure shows a high mobility of the loop, which we believe is responsible for the need for crystallization conditions different than that of SpvC WT. No significant deformations to the structure of the SpvC are induced due to the H106N mutation.

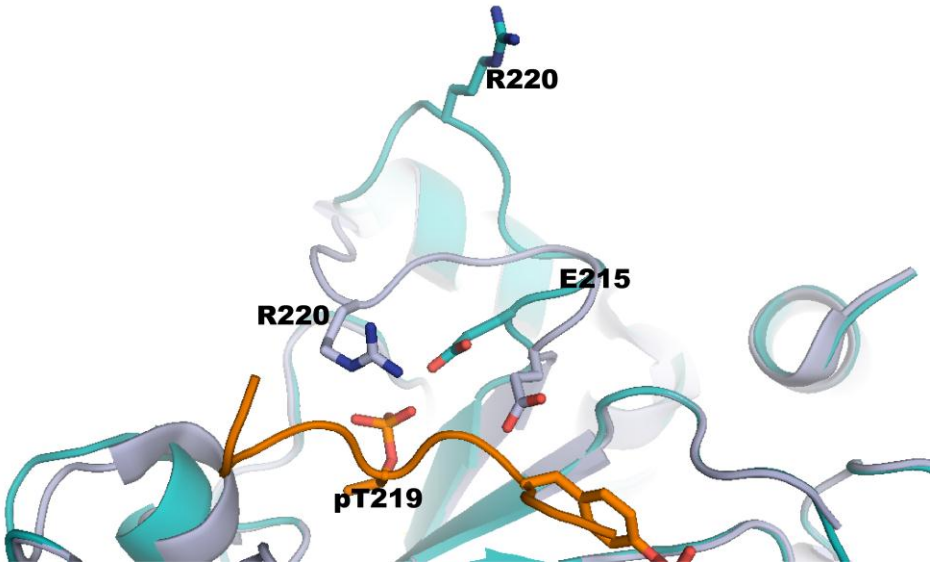

Figure 21. The movable loop of ligand-free SpvC WT (4HAH), in teal, aligned to SpvC K136A (2Q8Y) in light purple, with substrate bound, in orange. Upon substrate binding the R220 carrying loop changes its conformation to the closed form, such that the R220 is placed in proximity to the phosphothreonine residue (P219) of the substrate. The change also positions E215 in close proximity to the amide carbonyl of the substrate backbone.

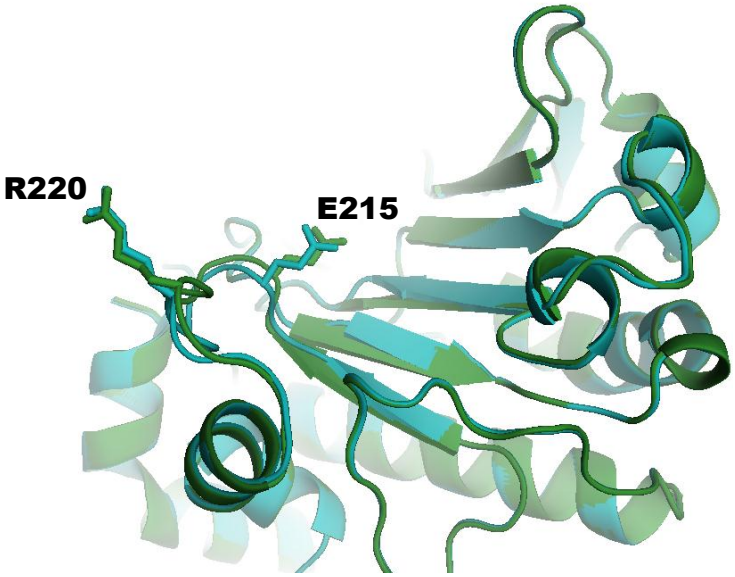

Figure 22. The movable loop carrying the R220 and E215 residues compared between the structures of SpvC WT shown in light blue and the SpvC H106N, shown in green. Residues 200 to 225 compared using SSM [27a] indicated a Q-score of 0.90 for all 26 residues.

CHAPTER 5

DISCUSSION

5.1 Results indicate the absence of a carbanion intermediate

The objective of the presented research was to provide structural information to assist in the interpretation of kinetics and mutagenesis experiments about the type of mechanism followed by phosphothreonine lyases. As described earlier, elimination reactions can take any one of the several pathways – E1, E1cB$_R$, E1cB$_F$ and concerted, mechanisms of which were shown in Figure 14. The research presented herein seeks to support the findings of the aforementioned deuterium isotope exchange experiment with structural data. The objective behind the research was two-fold:

- Determine the crystal structure of SpvC WT purified using a previously unpublished protocol.
- Purify the H106N mutant and determine the structures. Compare the structures of the wild type (4HAH) and the H106N mutant (4H43) to previously published structure (2Z8M), with the goal of determining if the mutation results in any structural perturbation that might affect substrate binding or catalysis.

The comparison of the two structures against previously published active WT structures allows us to verify the integrity of the proteins purified and crystallized in this study. The inactivity of the H106N mutant may arise due the structural changes induced due to the mutation, which might affect the binding of the substrate. This would result in the observed lack of deuterium exchange at the expected C-α, thus giving us a false positive result. In addition, the previously reported kinetic experiments with this mutant concluding that H106 is the catalytic

general acid would be invalidated if significant structural perturbations were found. Our work seeks to eliminate this possibility and support the deuterium exchange experiment in eliminating the E1cB mechanism of the various possibilities discussed earlier.

To test for the E1cB type of mechanism, it was necessary to create mutant that lacked the catalytic acid, responsible for donating a proton to the leaving phosphate group. The H106 residue was identified as the catalytic acid previously and was successfully created using site directed mutagenesis. Another lab member, Vyacheslav Kuznetsov, carried out deuterium isotope exchange experiments using this mutant.

Since the mutant lacked the catalytic acid, the reaction would remain incomplete. In the case of an E1cB mechanism being employed by SpvC (Equation 2, Table 1) depending on the rate-limiting step, different results would be observed. As stated earlier, it was our objective to test for the formation of a carbanion intermediate at the active site. An incomplete reaction due to lack of a catalytic acid would provide sufficient time for the reverse reaction, wherein the deprotonated substrate would pick up a deuterium atom from the solvent (buffered D_2O).

The substrate was incubated with the mutant enzyme in buffered D_2O. The substrate was then isolated and analyzed using mass spectrometry to determine the incorporation of deuterium on he C-α. The presence of deuterated C-α would indicate that a carbanion intermediate forms, and that the enzyme is capable of stabilizing such a species in the native reaction. Such a carbanion intermediate might undergo elimination rapidly (in other words, carbanion formation is rate-determining) or slowly, with elimination rate determining. If the first possibility is the case, then deuterium exchange might be observed even with the wild type enzyme. The H106N mutation hinders the elimination step, thus if the enzyme does form a carbanion intermediate, its only path is in the reverse direction, reprotonation by an

exchangeable proton (or deuteron) from the solvent. If, on the other hand, a concerted elimination reaction takes place with no carbanion intermediate, then no deuterium exchange will be observed even in the H106N mutant.

The lack of deuterium exchange at the C-α presents us with one of two possibilities –

- That SpvC does not dephosphorylate the phosphothreonine residue using carbanion intermediate, hence the lack of deuterium exchange at the C-α.
- The mutation of the H106N residues created a structural anomaly that renders the enzyme incapable of binding to the substrate peptide, thus resulting in lack of deuterium exchange.

Using the same purification strategy as the wild type SpvC, we successfully purified the SpvC H106N mutant. It was found that the mutant enzyme on incubation with the synthetic peptide did not release inorganic phosphate, as determined by phosphomolybdate assay[16]. Since we had created a mutant that altered an active site residue, it was necessary to verify the structural integrity of the purified mutant.

As described in the previous chapter, X-ray crystallography was used to determine the structure of our WT structure (4HAH) and mutant (4H43). The wild type structure obtained in this project was compared with the previously published structure of ligand free SpvC WT (2Z8M). It has been confirmed that the new structures retain their structural integrity despite the use of new purification and crystallization protocols. Our experimentation has shown that the SpvC WT, purified and crystallized by in this study, is active and is able to bind to the substrate effectively. Structural studies of SpvC H106N mutant have shown that the WT and the H106N mutant structure have no significant structural differences. The detailed comparison of the mutant with the wild type structures described in the preceding section reveal no structural changes in H106N mutant that would significantly affect binding of the substrate. While the

present structural investigation does not directly measure the ability of the H106N mutant to bind the substrates used in either the kinetics or deuterium exchange experiments, the present results document the lack of any structural consequences of the mutation.

The structural data obtained in this project therefore supports the conclusion that the lack of deuterium on the phosphopeptide implies that the enzyme class does not utilize the mechanism involving the formation of a carbanion (E1cB). While there exists the possibility of several other mechanisms, through our presented work, we have successfully eliminated the E1cB type reaction mechanism. These structural results also support the conclusion of an earlier study that the lack of activity of the histidine mutant implicates it as a catalytic residue,[8] and is not due to structural perturbations.

CHAPTER 6

CONCLUSION

Phosphothreonine lyases form a class of enzymes that are relatively new and are of great significance in development of drugs against disease causing pathogens like Salmonella and Shigella. Our objective has been to expand the understanding of this class of proteins by understanding the mechanism used by the proteins to carry out dephosphorylation.

Through our work we have gained some knowledge about the mechanistic aspects of the enzyme SpvC, which we believe will aid the scientific community in understanding phosphothreonine lyases. We have:

- Created a new plasmid carrying the protein with 10x His-tagged recombinant SpvC with TEV cleavage site. This creates a recombinant SpvC with only one residue from the cleavage site on the N-terminus as compared to previously published recombinant SpvC. The presence of fewer amino acids from tags and cleavage sites results in lesser packing irregularities during crystallization.
- Developed a new and easily reproducible purification protocol that results in OspF and SpvC of purity greater than 95%.
- Crystallized and solved the previously unreported structure of the SpvC H106N mutant In doing so, we have verified that the H106N (PDB ID 4H43) mutant retains all structural aspects of the active wild type protein (PDB ID 4HAH) with only minor differences in the position of the movable R220 loop.

- Results from our deuterium isotope exchange experiment indicate that phosphothreonine lyase catalyzes the dephosphorylation without the formation of a carbanion intermediate.

Our work allows us to narrow down the mechanistic possibilities that can be involved in the elimination of the phosphate group by phosphothreonine lyases. This will aid future research in developing mechanism-based inactivators of phosphothreonine lyases.

REFERENCES

1. C.D.C. Salmonella Outbreaks. http://www.cdc.gov/salmonella/outbreaks.html.

2. Diao, J.; Zhang, Y.; Huibregtse, J. M.; Zhou, D.; Chen, J., *Nat. Struct. Mol. Biol.* **2008,** *15* (1), 65-70.

3. Guiney, D. G.; Fierer, J., *Front Microbiol* **2011,** *2*, 129.

4. Parsot, C., *Curr. Opin. Microbiol.* **2009,** *12* (1), 110-6.

5. Buchrieser, C.; Glaser, P.; Rusniok, C.; Nedjari, H.; D'Hauteville, H.; Kunst, F.; Sansonetti, P.; Parsot, C., *Mol. Microbiol.* **2000,** *38* (4), 760-71.

6. Arbibe, L.; Kim, D. W.; Batsche, E.; Pedron, T.; Mateescu, B.; Muchardt, C.; Parsot, C.; Sansonetti, P. J., *Nat. Immunol.* **2007,** *8* (1), 47-56.

7. Zhu, Y.; Li, H.; Long, C.; Hu, L.; Xu, H.; Liu, L.; Chen, S.; Wang, D. C.; Shao, F., *Mol. Cell* **2007,** *28* (5), 899-913.

8. Li, H.; Xu, H.; Zhou, Y.; Zhang, J.; Long, C.; Li, S.; Chen, S.; Zhou, J. M.; Shao, F., *Science* **2007,** *315* (5814), 1000-3.

9. Brandao, T. A.; Hengge, A. C.; Johnson, S. J., *J. Biol. Chem.* **2010,** *285* (21), 15874-83.

10. Maynes, J. T.; Bateman, K. S.; Cherney, M. M.; Das, A. K.; Luu, H. A.; Holmes, C. F.; James, M. N., *J. Biol. Chem.* **2001,** *276* (47), 44078-82.

11. Kuipers, O. P.; Rollema, H. S.; Yap, W. M.; Boot, H. J.; Siezen, R. J.; de Vos, W. M., *J. Biol. Chem.* **1992,** *267* (34), 24340-6.

12. Chen, L.; Wang, H.; Zhang, J.; Gu, L.; Huang, N.; Zhou, J. M.; Chai, J., *Nat. Struct. Mol. Biol.* **2008,** *15* (1), 101-2.

13. Singer AU, S. T., Lam R, Gordon R, Nocek B, Kagan O, Edwards AM, Chirgadze NY, Parsot C, Arbibe L, Savchenko A, **2009**.

14. Brennan, D. F., Roe, S.M., Barford, D., Structure and Mechanism of the Chromobacterium violaceum VirA Phosphothreonine Lyase Type III Secretion System Effector. 2008.

15. Takashi Furuta, H. T., Hiromi Shibasaki, Ynadsuji Kasuya, *J. Biol. Chem.* **1992,** *267* (18), 12600-12605.

16. Lowry, O. H.; Lopez, J. A., *J. Biol. Chem.* **1946,** *162*, 421-8.

17. Wilkins, M. R.; Gasteiger, E.; Bairoch, A.; Sanchez, J. C.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F., *Methods Mol Biol* **1999,** *112*, 531-52.

18. Gorrec, F., *J Appl Crystallogr* **2009,** *42* (6), 1035-1042.

19. Pflugrath, J., *Acta Crystallographica Section D* **1999,** *55* (10), 1718-1725.

20. Zwart, P. H.; Afonine, P. V.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; McKee, E.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Storoni, L. C.; Terwilliger, T. C.; Adams, P. D., *Methods Mol Biol* **2008,** *426*, 419-35.

21. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 213-21.

22. Afonine, P. V.; Grosse-Kunstleve, R. W.; Adams, P. D., *Acta Crystallographica Section D* **2005,** *61* (Pt 7), 850-5.

23. Painter, J.; Merritt, E. A., *J Appl Crystallogr* **2006,** *39* (1), 109-111.

24. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., *Acta Crystallographica Section D* **2010,** *66* (4), 486-501.

25. Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., 3rd; Snoeyink, J.; Richardson, J. S.; Richardson, D. C., *Nucleic Acids Res.* **2007,** *35* (Web Server issue), W375-83.

26. Schrödinger, L.L.C. *The PyMOL Molecular Graphics System*, Ver. 1.2r3pre.

27. (a) Krissinel, E.; Henrick, K., *Acta Crystallographica Section D* **2004,** *60* (12 Part 1), 2256-2268; (b) Velankar, S.; Alhroub, Y.; Best, C.; Caboche, S.; Conroy, M. J.; Dana, J. M.; Fernandez Montecelo, M. A.; van Ginkel, G.; Golovin, A.; Gore, S. P.; Gutmanas, A.; Haslam, P.; Hendrickx, P. M.; Heuson, E.; Hirshberg, M.; John, M.; Lagerstedt, I.; Mir, S.; Newman, L. E.; Oldfield, T. J.; Patwardhan, A.; Rinaldi, L.; Sahni, G.; Sanz-Garcia, E.; Sen, S.; Slowley, R.; Suarez-Uruena, A.; Swaminathan, G. J.; Symmons, M. F.; Vranken, W. F.; Wainwright, M.; Kleywegt, G. J., *Nucleic Acids Res.* **2012,** *40* (Database issue), D445-52.

28. This study utilized the high-performance computational capabilities of the Helix Systems at the National Institutes of Health, Bethesda, MD (http://helix.nih.gov).