

STRUCTURE-FUNCTION RELATIONSHIP OF SUGAR

KINASES FROM *Vibrio cholerae* O395

by

Rakhi Paul

Enrollment No. LIFE05200904001

Saha Institute of Nuclear Physics

Kolkata

A thesis submitted to the

Board of Studies in Life Sciences

In partial fulfillment of requirements

For the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE



November, 2015

Homi Bhabha National Institute¹

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Ms. Rakhi Paul entitled "Structure Function Relationship of Two Sugar Kinases from *Vibrio cholerae* O395" and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Prof. Dipak Dasgupta	Retired in June 2015
Chairman – Prof. Subrata Banerjee <i>Subrata Banerjee</i>	Date: 17/3/16.
Convener (Guide) – Prof. Udayaditya Sen <i>Udayaditya Sen</i>	Date: 17/3/16
Examiner – Prof. Samudrala Gourinath <i>Samudrala Gourinath</i>	Date: 17/3/2016.
Member 1- Prof. Sampa Biswas <i>Sampa Biswas</i>	Date: 17/3/2016
Member 2- Prof. Abhijit Chakrabarti <i>Abhijit Chakrabarti</i>	Date: 17/3/2016

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.

I hereby certify that I/we have read this thesis prepared under my/our direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 17/3/16

Place: Kolkata

Udayaditya Sen
<Signature>

Guide

¹ This page is to be included only for final submission after successful completion of viva voce.

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI. Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Rakhi paul

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

Place- Kolkata

Date-30.11.2015

Rakhi Paul

List of Publications Arising from the Thesis

Journals

a. Published:

- i) Paul R, Patra MD, Sen U. Crystal Structure of Apo and Ligand Bound Vibrio cholerae Ribokinase (Vc-RK): Role of Monovalent Cation Induced Activation and Structural Flexibility in Sugar Phosphorylation. *Adv Exp Med Biol.*2015; 842:293-307.
- ii) Paul R, Patra MD, Banerjee R, Sen U. Crystallization and preliminary X-ray analysis of a ribokinase from Vibrio cholerae O395. *Acta Crystallogr F Struct Biol Commun.* 2014 Aug; 70 (Pt 8):1098-102.
- iii) Paul R, Nath S, Sen U. Cloning, expression, purification, crystallization and preliminary X-ray analysis of a fructokinase from Vibrio cholerae O395. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2012 Dec 1;68(Pt 12):1564-7.

b. Accepted: NA

c. Communicated: 1

d. Other Publications: NA

DEDICATIONS

To my parents

and my well-wishers

ACKNOWLEDGEMENTS

I take this opportunity to convey my gratefulness to one and all those who have supported and guided me through the entire tenure of PhD work.

I feel highly privileged to express my sincere gratitude to my esteemed supervisor, Prof. Udayaditya Sen, who spent a lot of time to read my thesis and gave me a lot of useful suggestions and criticisms, being patient with me through all the times. He introduced me to the theoretical and practical aspects of crystallography and guided me at every stage of my research work and gave constant encouragement. Indeed, I am honored to be one of his students.

I am extremely thankful to Department of Atomic Energy, Govt. of India and Saha Institute of Nuclear Physics for providing financial and academic support to conduct of this piece of research. I express my special gratitude to C&MB Division for providing instrumental and other academic supports.

The members of my supervisory committee, Prof. Dipak Dasgupta, Prof. Abhijit Chakrabarti, Prof. Sampa Biswas and Prof. Subrata Banerjee provided with many useful suggestions and comments to me.

It gives me immense pleasure to acknowledge Prof. J.K. Dattagupta for the highly motivational and inspiring lectures during Post M.Sc period, which helped me to choose crystallography as my PhD topic. I express my sincere thanks to Prof. Jhimli Dasgupta for her valuable comments and suggestions involving research works. Without her assistance, I probably would not be able to crystallize fructokinase. I am grateful to Abhijit Bhattacharyya and Bikram Nath for the technical supports they provided.

I also wish to thank my senior and junior lab colleagues for providing constructive supports throughout the work. I am also thankful to Bipin Bose for his help in lab maintenance.

I am deeply indebted to my parents, Sri Ramkinkar Paul and Smt. Chhanda Paul and my loving sister who enlighten the way of my life with their painstaking efforts and hard work, care, love, and contribution for shaping and moulding my scientific career.

I would like to deliver my special thanks to Dr. Tamal Basak who continuously encouraged me to be strong and optimistic in life.

Though, all cannot be mentioned, but none is forgotten.

CONTENTS

	Page No.
SYNOPSIS	viii-xxi
LIST OF FIGURES	xxii-xxiii
LIST OF TABLES	xxiv
CHAPTER 1 INTRODUCTION AND REVIEW OF LITERATURE	1-11
1.1 General aspects of bacterial metabolism	2
1.2 Embden-Meyerhof-Parnas (EMP) pathway or glycolysis	2
1.3 Pentose Phosphate Pathway	3
1.4 Type of sugars participate in metabolism	4
1.4.1 Ribose metabolism	5
1.4.2 Fructose metabolism	6
1.5 Phosphorylation of sugar is the key step before metabolization	7
1.6 Bacterial kinase: sequence and structure classification	7
1.7 Bacterial sugar kinase families	8
1.8 Objective	9
1.9 Determination of protein structure by X-ray crystallography	10
CHAPTER 2 MATERIALS & METHODS	12-27
2.1 Cloning and over-expression of proteins with 6× His-tag at the N-terminal	13

2.1.1	Amplification of gene sequence by PCR	14
2.1.1.1	Steps and ingredients involved in Polymerase Chain Reaction	14
2.1.1.2	Site directed mutagenesis (SDM)	15
2.1.2	Digestion and ligation of PCR product	16
2.1.3	Transformation of plasmid by electroporation	16
2.1.4	Checking for successful cloning	17
2.1.5	Overexpression of cloned gene	17
2.1.6	Storage for future use	18
2.2	Purification of recombinant protein followed by concentrating it for crystallization.	18
2.2.1	Cell Lysis and Removal of Crude Lysate	18
2.2.2	Removal of Non-Specific Proteins by Ni ²⁺ -NTA Affinity Chromatography	18
2.2.3	Removal of 6×-Histidine tag by Thrombin Cleavage	19
2.2.4	Removal of Impurity by Size Exclusion Chromatography	19
2.3	Crystallization, Data Collection, Data Processing and Structure Solution	20
2.3.1	Crystallization	20
2.3.2	Data collection	22
2.3.3	Data Processing	23
2.3.4	Initial Structure Determination	24
2.3.5	Refinement and model building step	24
2.3.6	Structure analysis	26

2.4	Determination of enzymatic activity	26
CHAPTER 3	CLONING, OVER-EXPRESSION, PURIFICATION, CRYSTALLIZATION, DATA COLLECTION AND DATA PROCESSING OF VcRK and VcFK AND ENZYME KINETICS OF Vc-FK.	28-44
3.1	Cloning and over-expression of VcRK and VcFK	29
3.2	Purification of 6×Histidine tagged proteins	30
3.3	Removal of Histidine tag by thrombin: standardization of cleavage time	30
3.4	Checking concentration of final protein	31
3.5	Crystallization of concentrated proteins	32
3.5.1	Crystallization of VcRK	32
3.5.2	Crystallization of VcFK	34
3.6	Data collection of apo and ligand bound VcRK and VcFK	36
3.6.1	Data collection of VcRK	36
3.6.2	Data collection of VcFK	37
3.7	Initial structure determination of VcRK and VcFK by molecular replacement	39
3.8	Structure refinement and interpretation of final models of VcRK and VcFK.	40
3.9	Results of enzymatic activity of VcFK	42

CHAPTER 4	ANALYSIS OF CRYSTAL STRUCTURE OF TWO	45-74
	SUGAR KINASES	
4.1	Structure of Ribokinase from <i>Vibrio cholerae</i> (VcRK)	46
4.1.1	Overall monomeric structure of VcRK	46
4.1.2	Packing of molecules inside the asymmetric unit of VcRK apo and VcRK complex crystals	48
4.1.3	Domains involved in dimerization and formation of β -clasp	49
4.1.4	Difference in dimeric structure of Apo and substrate-ADP bound structure	50
4.1.4.1	The Open homodimeric form of apo VcRK	50
4.1.4.2	Formation of halfway closed dimer upon ribose-ADP binding	52
4.1.5	Location of active site in VcRK	52
4.1.5.1	The five-membered ring form of ribose is the substrate for VcRK	53
4.1.5.2	The ribose-binding pocket	54
4.1.5.3	Residues involved in ribose binding	55
4.1.5.4	ADP binding site of VcRK	57
4.1.6	Divalent and monovalent cation binding	57
4.1.7	Sequence alignment VcRK with ribokinases from other sources	59
4.2	Structure of fructokinase from <i>Vibrio cholerae</i> (VcFK)	60
4.2.1	Overall monomeric structure	60
4.2.2	Dimerization of VcFK	61

4.2.3	Monovalent cation creates ATP binding site	65
4.2.4	Fructose binding site and associated structural changes	67
4.2.5	Loop movement associated with sugar binding	70
4.2.6	ADP binds in a pocket formed upon monovalent cation binding	71
4.2.7	Divalent cation binding site	72
4.2.8	BeF ₃ bound structure	73
CHAPTER 5	DISCUSSION	75-89
5.1	Monomeric structures of VcRK and VcFK display some common feature with other ribokinase superfamily members	76
5.2	Tight and specific binding of sugar induces large structural changes	79
5.2.1	Ribose binding triggers the lid closure in VcRK, not ADP	79
5.2.2	Sugar binding site of fructokinase is highly selective	81
5.2.2.1	Fructokinase phosphorylates fructose only at O6' position but not at O1'	81
5.2.2.2	Specificity of fructokinase towards fructose over AiRs	81
5.3	Mode of dimerization and its role in phosphotransfer fidelity	82
5.4	Formation of anion hole: stabilization of transition state	88

SYNOPSIS



Homi Bhabha National Institute

SYNOPSIS OF Ph. D. THESIS

- 1. Name of the Student:** Rakhi Paul
- 2. Name of the Constituent Institution:** Saha Institute of Nuclear Physics
- 3. Enrolment No. :** LIFE05200904001
- 4. Title of the Thesis:** Structure-Function Relationship of Two Sugar Kinases from *Vibrio cholerae* O395
- 5. Board of Studies:** Life Sciences

SYNOPSIS

1 INTRODUCTION

Carbohydrate metabolism is crucial for bacterial survival. Extracellular polysaccharide hydrolysis generates various hexoses, pentoses and their derivatives. These substances support the growth of many bacteria and archaea. Glucose is the most abundant sugar participating in metabolic pathways like glycolysis and TCA cycle, which are essential for generation of adenosine tri-phosphate (ATP), the main energy source of the living organisms. Beside glucose, other hexose monosaccharides like fructose (fruit sugar), mannose, galactose and pentose sugars

like ribose, arabinose are available to bacteria from different sources and they need to be phosphorylated before entering into the metabolic processes. These phosphorylation reactions are catalyzed by different kinases, specific of each sugar, where ATP serves as the phosphoryl group donor. Ribokinase (RK) phosphorylates ribose before entering into pentose phosphate pathway or nucleotide salvage pathway [1] whereas Fructokinase (FRK) phosphorylates fructose to fructose-6-phosphate before participating in glycolysis [2].

Ribokinase (RK)

Phosphorylation of ribose by RK is essential for entry of ribose into pentose phosphate pathway, a crucial step of carbohydrate metabolism. The product ribose-5-phosphate is then used for the synthesis of nucleotides and amino acids like tryptophan and histidine [1]. RK belongs to a well-known protein family called Ribokinase superfamily whose folding patterns and substrate specificity have been studied extensively during the last decade [3]. Apart from RK, this family includes FRK, 1-phosphofructokinase, 6-phosphofructo-2-kinase, tagatose-6-phosphate kinase, inosine kinase, adenosine kinase and pyridoxal kinase. Members of the RK family possess similar secondary structure, but their substrate binding site, overhanging lid or flap region and mode of oligomerization differ. An aspartic acid residue is thought to be the base responsible for activating O5' of ribose, which in turn attacks of the γ phosphate of ATP. An anion hole near the N-terminus of helix $\alpha 8$ is thought to stabilize the transition state. Monovalent cations with ionic radii similar to K^+ , NH_4^+ activate the enzyme before phosphoryl group transfer [4]. Divalent cations like Mg^{2+} , Ca^{2+} , Mn^{2+} are essential for the catalysis to stabilize the multiple negative charges of ATP.

Fructokinase (FRK)

FRK belongs to ribokinase-like superfamily or PfkB superfamily of kinases. FRK phosphorylates fructose to fructose-6-phosphate, and thereby allows alternate entry of the sugar into EMP pathway and murein biosynthesis. Although several biochemical experiments, mainly from different plant sources, have been done for FRK, similar data from bacterial sources are limited. Structural study on FRK is limited to the crystal structure of apo form fructokinase from *H. orenii* [2]. However, no substrate-cofactor bound structure of fructokinase has been reported yet.

2. OBJECTIVE:

To bring more light into sugar phosphorylation of *Vibrio cholerae*, we have cloned, over-expressed, purified, crystallized and solved the structures of *Vibrio cholerae* RK (Vc-RK, gene name *rbsK*) and FRK (Vc-FRK, gene name *cscK*) in their apo form and substrate-cofactor bound form.

3. RESULTS AND DISCUSSION

3.1 Ribokinase:

Vc-RK (306 amino acids; M.W. 32 kDa) was cloned in pET28 vector and expressed as 6×His tagged protein. The protein was purified through Ni²⁺-NTA chromatography, cleaved the 6×His tag by thrombin and finally the protein was purified from contaminants by size exclusion chromatography. Although fresh protein samples appeared to be homogeneous, a non-reducing 10 % (w/v) SDS-PAGE of few-days-old samples indicated a heterogeneous population containing higher oligomeric species, which disappeared on a reducing gel (**figure 1a**). We

therefore added 5mM β ME (β -mercaptoethanol) in all protein samples prior to crystallization to obtain good quality crystals (shown in fig. 1b).

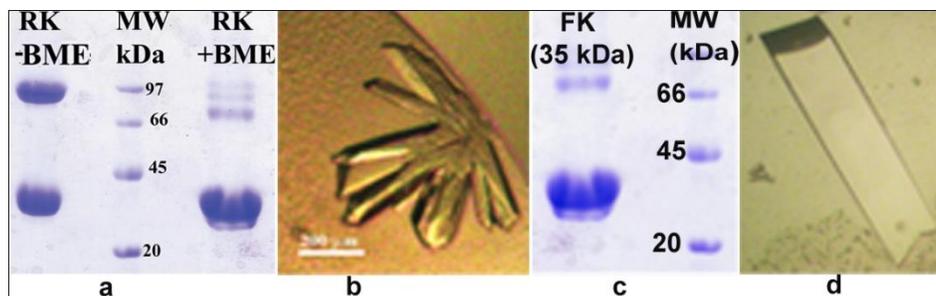


Figure 1. (a) SDS-PAGE of Vc-RK. (b) Crystals of Vc-RK. (c) SDS-PAGE of Vc-FRK and (d) crystal of Vc-FRK.

We have solved the crystal structures of apo (3.4\AA), sugar+ADP bound (1.75\AA) and sugar+ADP+Cs⁺ bound (2.37\AA) Vc-RK. All structures have been solved by molecular replacement (MR) [5, 6] method. Initially the structure of sugar+ADP bound form was solved using the coordinates of *E. coli* RK [1] (PDB code: 1RK2; sequence identity 54%) as search model. The coordinates of Vc-RK thus obtained was subsequently used for solving the structure of apo Vc-RK and Cs⁺ bound form.

Overall structure of Ribokinase

Gel filtration chromatography indicated a dimeric Vc-RK in solution, which was also supported by PDBePisa [7]. Each monomers of Vc-RK (fig. 2a) has two domains: a central globular α/β domain that contains most of the catalytic residues and a β -sheet region, distinctly protruded from the former, which participate in dimer formation (fig. 2b and 2c). Eight such dimers of apo Vc-RK in the asymmetric unit are arranged in a double-helical fashion (fig. 2d, 2e). For sugar + ADP -bound form and sugar + ADP + Cs⁺- bound form, each asymmetric unit contains two such dimers.

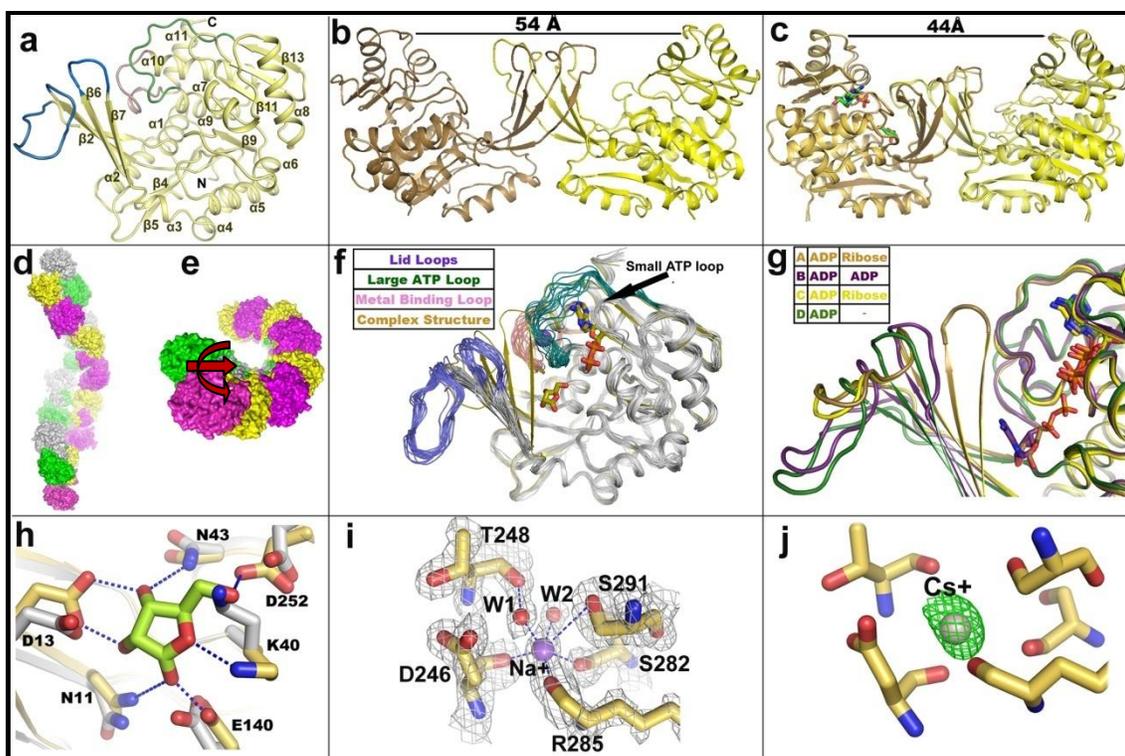


Figure 2. (a) Ribbon representation of the secondary structure of Vc-RK monomer. The dimerization loop (blue), ADP binding loop (green) and the metal binding loop (burgundy) are indicated. (b, c) Ribokinase dimer in apo and substrate bound form; apo dimer is symmetric about β -clasp axis and more open than the ternary complex structure. The distances indicated were measured between C α atoms of residue Thr195 from the two subunits. Ternary complex is more closed towards ADP-ribose bound chain than towards ADP bound chain (d-e) Arrangement of 16 monomers in the form of eight dimers forming a double-stranded helical pattern; upper strand composed of four ‘yellow and magenta’ dimers and the lower strand consists of four ‘grey and green’ dimers (top), perpendicular view of the helix axis where spiral like look is evident (bottom). (f) Superposition of sixteen apo monomers (gray) with complex structure (yellow): loops showing high flexibility in apo structure are highlighted. (g) Differences in lid closure upon ligand binding. (h) Residues of Vc-RK (yellow sticks) involved in binding sugar (green sticks). Hydrogen bonds between sugar and Vc-RK are shown as dashed line (blue). Amide nitrogen atoms of G39 and N11 are not shown. The structure of apo Vc-RK (grey) is overlaid on it to show the local structural change upon sugar binding. (i) 2Fo-Fc map contoured at 1.6 σ (light blue mesh) around the monovalent cation (Na⁺ ion, violet sphere) binding site. Residues interacting with the metal ion (yellow sticks) and two water molecules (W1 and W2, small red spheres) are also shown (j) Fo-Fc map contoured at 6.5 σ (green mesh) with Cs⁺ (CsCl) soaked crystal showing Cs⁺ (flesh sphere) occupying roughly the same position as Na⁺.

The catalytic α/β domain consists of a central nine stranded twisted β -sheet flanked on both faces by five α -helices (fig. 2a). The first six β -strands of this nine-stranded sheet together with the associated α helices have the topology of a typical Rossmann fold. The ribokinase fold is extended by three additional β strands (β_{11} – β_{13}) connected by short reverse turns. The protruding β -sheet domain is formed by two insertions into the central α/β fold, each of which forms two pairs of antiparallel β -strands: β_2 – β_3 (residue 13-36) and β_6 – β_7 (residue 95-107) (named as lid loops, fig 2f). A long loop (residues 18-31) joins β_2 and β_3 , which bends at Pro19 and Pro24. A superposition of all sixteen monomers of apo Vc-RK chain along with sugar+ADP bound Vc-RK chain (fig. 2f) indicate that although the central α/β region (gray) superposes well loops implicated in ATP binding (pink and green) and dimerization (violet) exhibit flexible nature.

Substrate, cofactor binding and change in dimerization:

The active site is located near a shallow trench between the β -flap and the central α/β domain (fig. 2f). This region consists of a ribose-binding site, an ATP binding site and an adjacent anion hole.

Open apo dimer: In the apo dimer α/β domain of each monomer stays away from the β -clasp region with a distance of 54Å between them (fig. 2b). This open form of the apo Vc-RK structure allows entry of the ribose into the sugar-binding pocket. During dimer formation, few residues from the larger β -turn of one monomer like Leu28, Gly30, and Tyr33 come close to the residues Ile109, Ser110 and Glu112 respectively from the shorter β -turn of another monomer. The inner part of the clasp region contains mainly hydrophobic residues like Met96, Val35, Pro22, Tyr33 and Phe21.

Ribose binding and halfway-closed asymmetric dimer formation: High resolution (1.75Å) electron density map clearly indicates that D-ribose adopts a five membered α -furanose form when bound to Vc-RK. This form allows the O5' atom of ribose to be available for phosphorylation. D-ribose molecule binds on top of the conserved 'GGK' motif and projects all of its non-ring 'O' atoms towards the α/β domain. It forms altogether 12 hydrogen bonds with the highly conserved side chains of Asn11, Asp13, Lys40, Asn43, Glu140 and Asp252 and the main-chain amide atoms of Gly39 and Asn11 (fig. 2h). Asp252 is believed to act as the base responsible for ionizing the O5' of ribose to make it better a nucleophile in attacking the γ -phosphate of ATP.

Crystals of Vc-RK complexed with ribose and ADP contain four molecules in the asymmetric unit. The four monomers (chains A, B, C, D) are arranged in the form of two dimers A:B and C:D. Out of the four monomers chain A and chain C bind both ribose and ADP, chain B binds two ADP molecules and chain D binds only one ADP molecule but no sugar. Upon ribose binding, the protruding lid (composed of β 2- β 3 and β 6- β 7 and the associated loops) acquires a closed conformation and folds back to the α/β domain with the sugar trapped underneath. Therefore, chain A and chain C assume a closed conformation due to ribose binding but, chain D remains in the open form as there is no sugar at the active site (fig. 2f, 2g). For chain B, ADP occupies the ribose site where the sugar part of ADP matches with the site normally occupied by ribose keeping the phosphates towards the α/β domain and the adenine base towards the lid. Larger size of ADP and/or its non-cognate nature may keep the lid in the open conformation. Therefore, for each dimer one chain (A or C) is in the closed conformation while the other chain (B or D) is in the open conformation leading to a 'half-way closed' asymmetric dimers about their symmetry axis (fig. 2c).

ADP binding: The nucleotide-binding site is located in a shallow groove in the α/β -fold domain, next to the substrate ribose but not covered by the lid. The adenine base is stacked between the main chain atoms in the long loop connecting strand β 13 and α 9 (residues Gly239 to Ala250). On one side, there are β 11, β 12 and the connecting short loop and the other side contains α 9 and α 10 (fig 2f). In contrast to the substrate ribose, the sugar of the nucleotide analogue has the more common C3'-endo pucker. The O2' atom of ribose forms one hydrogen bond with the His276 ND1 atom of Vc-RK. The phosphate group makes one hydrogen bond with protein, from O3' to N-atom of Gly222. The phosphate group head towards ribose residue at an approximate distance of 7.5 Å from the sugar O5' atom.

Monovalent cation binding: Monovalent cations like Na⁺, K⁺, Cs⁺ and NH₄⁺ ions act as an activator of RK [4]. All the four chains of Vc-RK are seen to be activated through monovalent cation binding and from a scrutiny of salts used during crystallization this cation was assigned as Na⁺. 2Fo-Fc map around the Na⁺ ion is shown along with the metal ion binding residues of Vc-RK and two water molecules (fig. 2i). Na⁺ is seen to be hexa-coordinated and is within polar contact with Asp246, Ser282, Arg285, Ser291 and two water molecules. For further confirmation of the position of Na⁺, we soaked the crystals of Vc-RK+ribose+ADP with a cryo buffer containing 40 mM CsCl and solved the structure to a resolution of 2.37Å. A Fo-Fc map calculated with the refined structure (with no monovalent cation at this site) shows a clear density for Cs⁺ located roughly in the same position of Na⁺ (fig. 2j).

3.2 Fructokinase

Vc-FRK (323 aa, M.W. 35 kDa) was cloned, expressed and purified using the same protocols followed for Vc-RK. Purified Vc-FRK (fig. 1c) was crystallized using PEG as precipitant (fig. 1d). Crystals of Vc-FRK diffract upto 2.45 Å (apo), 2.5 Å (fructose-bound), 1.75 Å (fructose-

ADP-Ca²⁺ bound) and 1.3 Å (fructose-ADP-Ca²⁺-BeF₃ bound). Crystal structures of apo Vc-FRK were solved by MR method [5, 6] using the coordinates of Aminoimidazole riboside kinase [8] (PDB code 1TZ6; 54% sequence identity) which was subsequently used to solve the substrate bound Vc-FRK structures.

Overall structure:

Like other members of RK superfamily Vc-FRK forms dimer and as expected, the apo Vc-FRK form an open dimer (fig. 3a) whereas substrate bound form has a closed structure (fig. 3b). The fold of Vc-FRK is quite similar to Vc-RK and they superposes with an RMSD of 1.57Å. Most of the structural changes occur in the active site lid region. Lid domain of Vc-FRK is smaller in length than RK, and appears less inter-twisted compared to Vc-RK. High inter-twisting helps the lid region to approach the adjacent monomer more easily whereas in Vc-FRK smaller lid domain folds back towards the same monomer (fig. 3a and 3b). Moreover, Vc-FRK contains insertion of an extra helix-turn (approximately 13 extra residues inserted) between residues 173-193. This helix-turn is near substrate binding site and blocks alternate entrance of fructose from solvent.

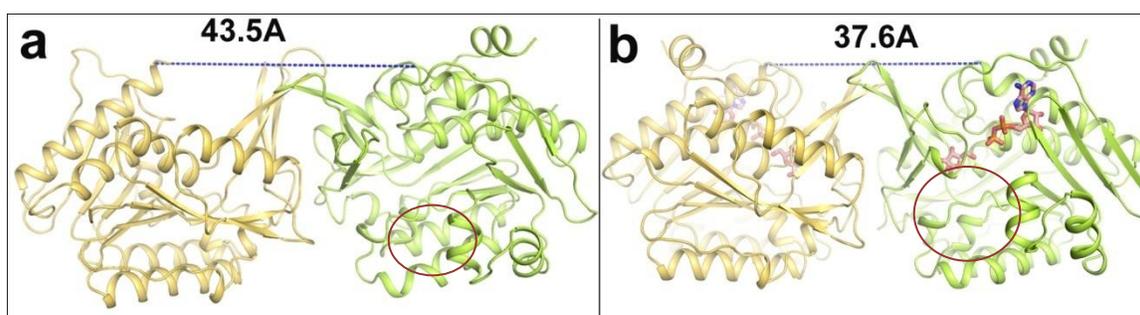


Fig. 3. (a) ‘Open’ dimer of apo Vc-FRK. (b) ‘Closed’ dimer in fructose+ADP+Ca²⁺ bound state. The distances indicated were measured between C^α-atom of residue 304 between two subunits. Shifting back of lid towards own monomer in substrate bound figure has been shown by dark circles.

Sugar, ADP and divalent cation binding:

As expected, there is a huge structural change upon sugar binding and the loop joining β 6- β 7 shifts 15Å to cover the fructose (fig. 4a). Fructose molecule is hydrogen bonded with the side chain nitrogen and oxygen atoms from Asp26, Asp30, Asn49, and Arg176 and with main chain nitrogen from Ala46. Besides, few hydrophobic residues like Phe113, Phe115 from the lid domain help substrate to be trapped into active site groove before phosphorylation (fig. 4b).

The ADP molecule is emplaced in a shallow trench, mostly unveiled to the solvent molecule. Its phosphate group is directed towards the O6'-atom of fructose whereas the adenosine group is partially covered by the small loop joining β 11 and β 12 on one side and by α 7, α 8 and the long loop joining β 13 and α 7 on the other side. The adenosine part of ADP is stabilized by hydrogen bonds with main chain atoms of Gly239 and Val256 and side chain of Asn296. A Ca^{2+} ion is positioned in the active site, in close proximity to the fructose and the β -phosphate of ADP residue. Ca^{2+} ion adopts an octahedral geometry with four water molecules forming the base and a water molecule and the O $^{\delta}$ atom of Asn174 occupying the vertices of the octahedron (fig. 4c, 4f). The negatively charged atmosphere created by nearby acidic acid residues prevent the phosphate groups from taking alternate positions and thereby motivates the cofactor towards the bound sugar molecule before phosphorylation (fig. 4d and 4e).

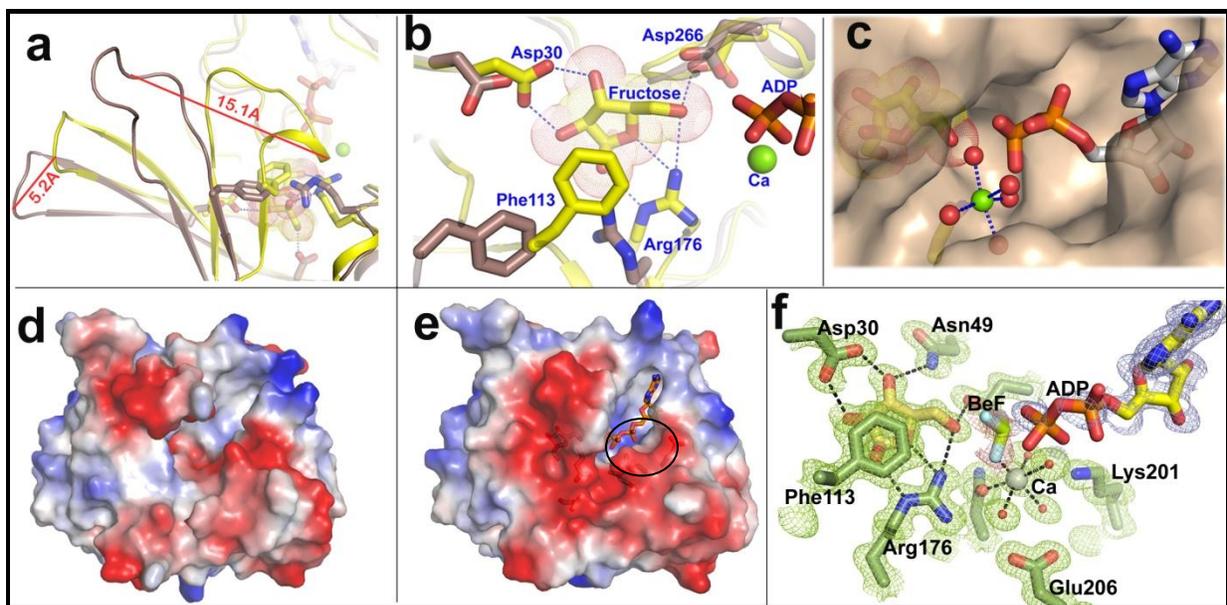


Fig. 4. (a) Open and closed “lid” formed in Vc-FRK before and after sugar binding (coffee color: apo structure and yellow: sugar+ADP+Ca²⁺ bound). (b) Structural changes upon fructose binding. (c) ‘Buried’ sugar with O6’ pointing ADP (partially buried in a shallow cleft) and octahedrally coordinated Ca²⁺. (d) Electrostatic surface of apo and (e) sugar+ADP bound structure with sugar buried. (f) 2Fo-Fc map contoured at 1.0 σ around BeF₃ (flesh color mesh), fructose, nearby protein residues and Ca²⁺ ion (green mesh) and ADP (violet mesh).

BEF3 binding: So far, several structures belonging to RK superfamily have been solved, either in free or sugar and/or ADP bound states. However, these structures do not provide the location and orientation of γ -phosphate during sugar phosphorylation. To identify this we have soaked the ADP+fructose bound Vc-FRK crystal with BeF₃ and solved the structure at 1.3Å. This high-resolution map allows us to locate precisely the BeF₃ molecule occupying exactly halfway (3.2Å from each) between the acceptor O6’ of fructose and the β -phosphate of ADP (fig. 4f). Two fluorine atoms points towards the anion hole and within H-bonding distance with the main chain amide of Gly263, Gly265 and Asp266. The other fluorine atom is hydrogen bonded with Arg176 and within polar contact with Ca⁺² atom. Overall, this structure provides a precise

picture of the geometry of phosphoryl transfer during fructose phosphorylation and role of nearby conserved residues to make the phosphorylation reaction with high fidelity. This structure reveals the general characteristics of metallo-fluorides as accurate analogs of the transition state and the ground state in the context of kinase mechanisms.

4 Summary

In order to dissect the various hierarchical steps in bacterial sugar phosphorylation, to identify motifs for substrate recognition, we have solved the crystal structures of two enzymes in apo and substrate-cofactor bound form. We are first to report the substrate bound structure of *Vibrio cholerae* fructokinase and Ribokinase: (i) Structural change of fructokinase upon sugar binding, (ii) Mimic the reaction intermediate steps with BeF_3 bound fructokinase structure, (iii) Asymmetric dimerization of Ribokinase depending on ribose binding, (iv) Binding of nucleotide in the position of ribose, which helps us to find out alternate substrate and potential ingredient for a new enzymatic action.

5 Significance

Ribose and fructose emerge as new model substrates for pentose phosphate pathway and glycolysis, as an alternate of glucose, in depth study of bacterial metabolism. Our results can be used to investigate sequences, structure, thermodynamic and kinetic aspects of not only *Vibrio cholerae* sugar phosphorylation but also for other enzymes related to carbohydrate metabolism for bacterial system. Our finding will open new quest for ‘functional role of monovalent and divalent cations’ in other ATP dependent kinases. Our structural study can be optimized for other substrate-cofactor-ligand complex systems. For fructokinase, the x-ray structure of BeF_3 bound enzyme show explicit enzyme–substrate interactions that discriminate between the ground and the transition states of the reaction. They also illustrate the partially dissociative geometry of the

transition state of phosphoryl transfer and demonstrate the potential applications of metallo-fluorides for the study of kinase mechanisms.

References

1. Sigrell, J. A, Cameron, A. D, Jones, T. A., Mowbray, S. L. (1998). *Structure* 6, 183-193.
2. Chua, T. K., Seetharaman, J., Kasprzak, J. M., Ng, C., Patel, B. K., Love, C., Bujnicki, J. M., Sivaraman, J. (2010). *J. Struct. Biol.* 171(3), 397-401.
3. Park, J. & Gupta, R.S. (2008). *Cell. Mol. Life Sci.* 65, 2875–2896.
4. Li, J., Wang, C., Wu, Y., Wu, M., Wang, L., Wang, Y., Zang, J. (2012). *J. Struct. Biol.* 177, 578-582.
5. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., Read, R. J. (2007). *J. Appl. Crystallogr.* 40, 658-674.
6. Potterton, E., Briggs, P., Turkenburg, M., Dodson, E. (2003) *Acta Cryst. D59*, 1131-1137.
7. Krissinel, E., Henrick, K. (2007). *J. Mol. Biol.* 372, Pages: 774-797.
8. Zhang, Y., Dougherty, M., Downs, D. M., Ealick, S. E. (2004). *Structure* 12(10), 1809-1821.

Publications in Refereed Journal:

- a. **Published:**
 - i) Paul R, Patra MD, Sen U. Crystal Structure of Apo and Ligand Bound Vibrio cholerae Ribokinase (Vc-RK): Role of Monovalent Cation Induced Activation and Structural Flexibility in Sugar Phosphorylation. *Adv Exp Med Biol.* 2015; 842:293-307.
 - ii) Paul R, Dandopath Patra M, Banerjee R, Sen U. Crystallization and preliminary X-ray analysis of a ribokinase from Vibrio cholerae O395. *Acta Crystallogr F Struct Biol Commun.* 2014 Aug; 70 (Pt 8):1098-102.
 - iii) Paul R, Nath S, Sen U. Cloning, expression, purification, crystallization and preliminary X-ray analysis of a fructokinase from Vibrio cholerae O395. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2012 Dec 1;68(Pt 12):1564-7.

b. **Accepted: NA**

c. **Communicated: 1**

d. **Other Publications: NA**

Signature of Student:

Date:

Doctoral Committee:

Sl.No	Name	Designation	Signature	Date
1.	Prof. Dipak Dasgupta	Chairman		
2.	Prof. Udayaditya Sen	Guide & convener		
3.	Prof. Abhijit Chakrabarti	Member		
4.	Prof. Sampa Biswas	Member		
5.	Prof. Subrata Banerjee	Technology Adviser		

LIST OF FIGURES

Figure No.	Name of Figures	Page No.
1.1	Flow diagram of metabolic conversion of glucose to pyruvic acid and generation of ATP in Glycolysis	3
1.2	Steps involved in Pentose phosphate pathway are illustrated by flow diagram	4
1.3	Hexose and pentose sugars	5
1.4	Steps involved in fructolysis and its final products	6
2.1	The phase diagram in protein crystallization	22
3.1	Size comparison of DNA fragments before and after cloning in 1% agarose gel	29
3.2	12 % SDS-PAGE profile showing the Ni-NTA purification profile	30
3.3	Standardization of cleavage time using thrombin	31
3.4	SDS PAGE profile for checking concentration and impurity level	32
3.5	VcRK crystals in (a) apo form, (b) ribose +ADP bound form	33
3.6	VcFK crystals in (a) apo form, (b) sugar bound form, (c) sugar +ADP +Ca ²⁺ bound form	34
3.7	X-ray diffraction image of VcRK crystals	36
3.8	X-ray diffraction image of VcFK crystals	38
3.9	Results of enzymatic activity of VcFK	43
4.1.1	Cartoon representation of VcRK monomer	47
4.1.2	Spiral arrangement of apo VcRK molecules in the crystal	48

4.1.3	Crystallographic arrangement of sugar-ADP bound VcRK	49
4.1.4	The ‘hand-shake motif’ or ‘ β -clasp’ that forms the dimer of VcRK	50
4.1.5	The (a) apo and (b) substrate bound dimeric form of Vc-RK	51
4.1.6	Structural flexibility of the active site loops in apo VcRK	53
4.1.7	All five forms of D-ribose in solution	54
4.1.8	Location of ribose in VcRK and surrounding region	55
4.1.9	Details of ribose binding in VcRK	56
4.1.10	Zoomed view of ADP binding site	57
4.1.11	Monovalent cation binding site in VcRK	58
4.1.12	Alignment of VcRK with ribokinases from other sources	59
4.2.1	Secondary structure of VcFK	61
4.2.2	Cartoon representation showing the dimeric structure of VcFK in apo form and in sugar-ADP bound form	63-64
4.2.3	Hydrophobic residues involved in dimerization	64
4.2.4	Structural changes due to monovalent cation binding in VcFK	66
4.2.5	Pentavalent geometry of the monovalent cation Na ⁺ .	67
4.2.6	Entrapment of fructose molecule in VcFK	68
4.2.7	Detailed interaction of fructose with VcFK	69
4.2.8	Movement of large lid loop on sugar binding	70
4.2.9	The nucleotide-binding site of apo and sugar-ADP bound VcFK structure	71
4.2.10	Octahedral geometry of calcium ion in between sugar and β -phosphate of ADP	72

4.2.11	Overall active site and geometry of phosphotransfer	73
5.1	Figure 5.1: Structure and sequence alignments of ribokinase family of proteins	78
5.2	Relative movement of β -clasp region for different chains	80
5.3	Phosphorylation at O1' and O6' positions of fructose by three different kinases	81
5.4	Specificity of VcFK towards fructose	82
5.5	Covering ADP and ribose by the lid domain during dimerixation of VcRK Electrostatic surface representation of VcFK	84
5.6	Difference in dimerization of E.coli RK (EcRK) and VcFK	85
5.7	Electrostatic surface representation of VcFK	86
5.8	Role of active site glutamic acid residues and their conservation throughout the superfamily	87
5.9	Stabilization of negative charge via anion hole	88

LIST OF TABLES

Table No.		Page No.
2.1	Primer sequences for PCR of wild type and mutant proteins	14
2.2	Composition of PCR reaction mixture	15
2.3	Calculation of the two-step PCR reaction mixture	16
3.1	Detailed crystallization conditions for VcRK and VcFK crystals in apo and ligand bound form	35
3.2	Data-collection and processing parameters of VcRK crystals	37
3.3	Data-collection and processing parameters of VcFK crystals	39
3.4	Selection of initial model for molecular replacement of VcRK	40
3.5	Selection of initial model for molecular replacement of VcFK	40
3.6	Refinement statistics and description of final model of VcRK	41
3.7	Refinement statistics and description of final model of VcFK	42

CHAPTER- 1

INTRODUCTION

AND

REVIEW OF LITERATURE

1.1 General aspects of bacterial metabolism

Metabolism is the sum of all chemical transformations that occur in a cell or organism to maintain its living state. It is a highly coordinated and directed cell activity. Metabolism can be conveniently divided into two categories: i) Catabolism, where the breakdown of nutrient molecules into smaller products produces energy in the form of ATP and NADH or NADPH. ii) Anabolism (or biosynthesis), which allows the synthesis of all compounds needed by the cells, such as fatty acids, proteins and other macromolecules. As catabolism is productive in terms of energy and reductive power, anabolism is demanding. It requires both energy and reductive power. Metabolic pathways can be linear, branched or cyclic type. Catabolism is a convergent process opposite to anabolism, which is in general divergent. Catabolism and anabolism are highly regulated process and are controlled by cell's necessity, to prevent wasteful reactions, i.e., synthesis and degradation of same molecule simultaneously [1, 2].

Though different highly specific pathways present in different species, few general pathways govern the central metabolism in almost all living cells. These include glycolysis or EMP pathway (whose final outcome is production of ATP from glucose), TCA cycle or Krebs cycle (used by all aerobic organisms to generate ATP through the oxidation of acetate derived from carbohydrates, fats and proteins) and pentose phosphate pathway (in which hexose sugars are oxidized to pentoses) [1].

1.2 Embden-Meyerhof-Parnas (EMP) pathway or glycolysis

Glycolysis includes a series of reactions (**figure 1.1**) that constitute the first phase of most carbohydrate catabolism. Glycolysis breaks down glucose molecule and forms pyruvate with the production of two molecules of ATP. The pyruvate, final product of glycolysis, can be used in either anaerobic respiration if no oxygen is available or in aerobic respiration via the TCA cycle which yields much more usable energy for the cell [3].

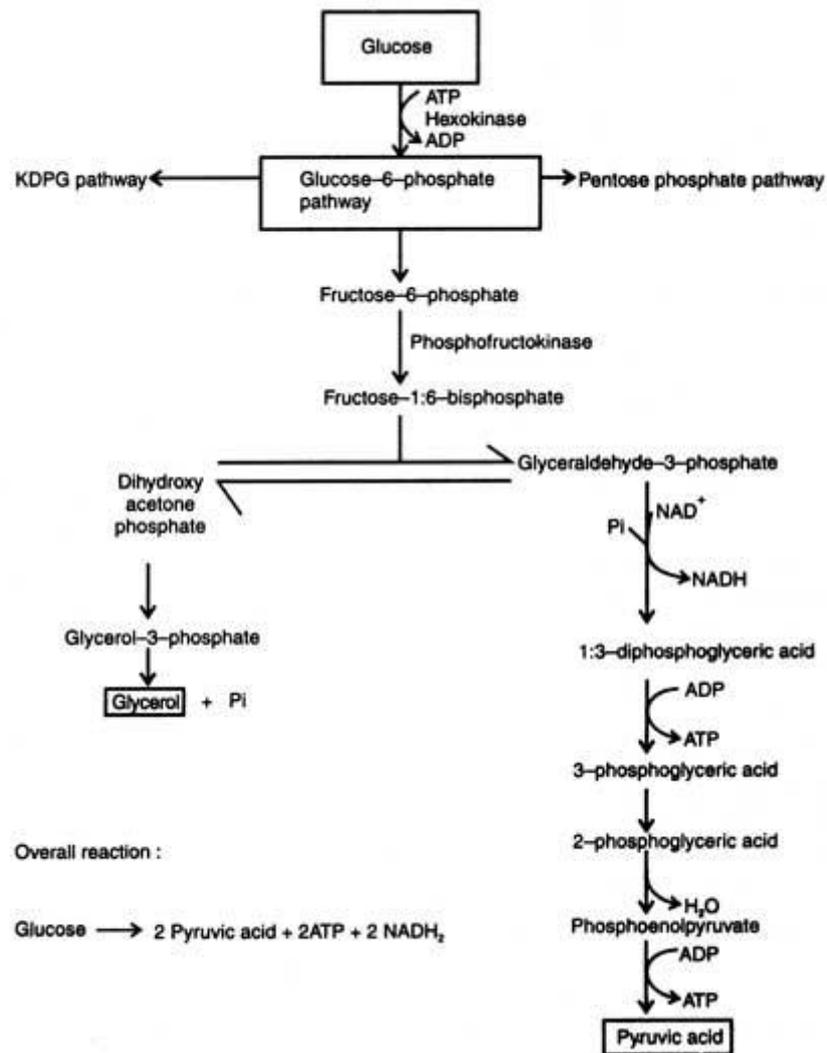


Figure 1.1: Flow diagram of metabolic conversion of glucose to pyruvic acid and generation of ATP in Glycolysis (figure adapted from *Bacterial Physiology and Metabolism*, 2008).

1.3 Pentose Phosphate Pathway

The pentose phosphate pathway converts a six-carbon containing glucose molecule into a five-carbon containing pentose sugar ribose, which is used as a component of the nucleic acids. This pathway consists of two phases: the oxidative generation of NADPH and the non-oxidative inter-conversion of sugars (**Figure 1.2**). In the oxidative phase, NADPH is generated when glucose 6-phosphate is oxidized to ribose 5-phosphate. This five-carbon

sugar and its derivatives are the components of RNA and DNA, ATP, NADH and coenzyme A [3].

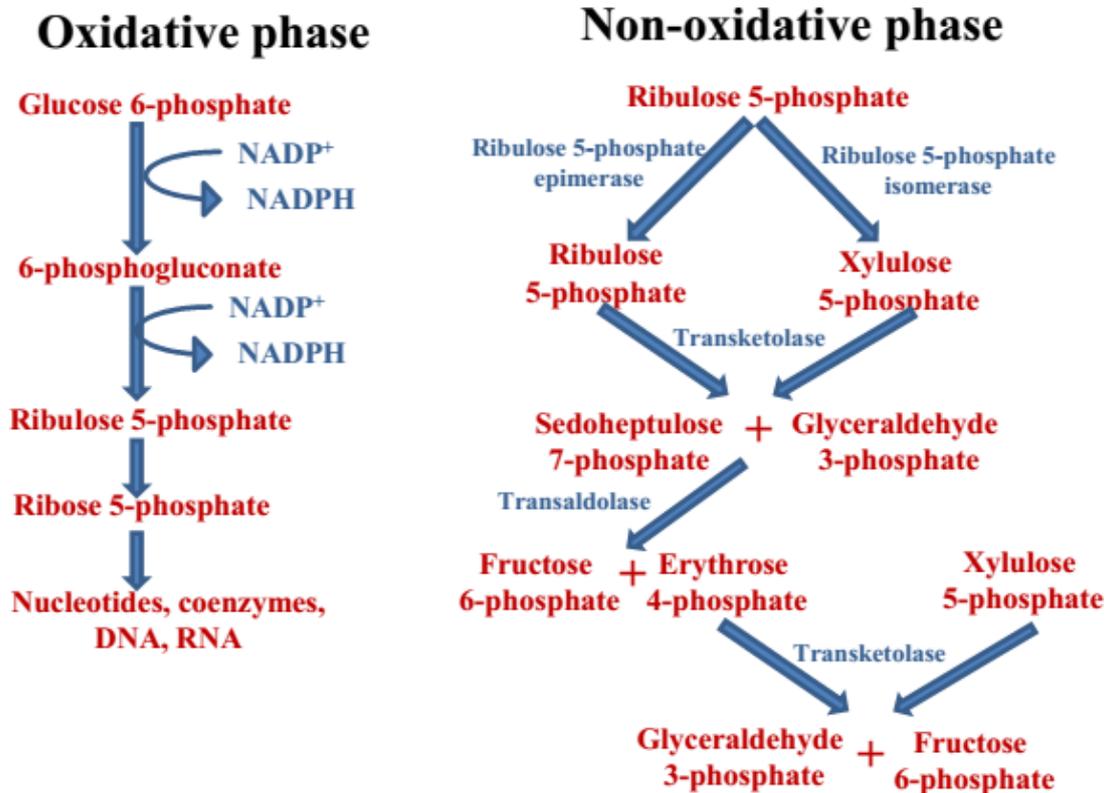


Figure 1.2: Steps involved in Pentose phosphate pathway are illustrated by flow diagram. Left side contains oxidative phase and right side illustrates the non-oxidative phase.

1.4 Type of sugars participate in metabolism

Glucose is the primary monosaccharide participant of bacterial metabolism. Glycolysis uses glucose as its main ingredient to generate ATP. Other than glucose, many other hexose and pentose sugars from different sources can participate in metabolism by taking entry in different steps of EMP and HMP pathway. Among these sugars, most important are hexose (like fructose, galactose, mannose and tagatose) and pentose sugars (like ribose, xylose, arabinose) as depicted in **figure 1.3**. Among these, fructose is adequately available in nature. When cell's glucose level becomes insufficient, fructose is converted to

glucose by Gluconeogenesis [1]. Ribose sugar is primary component of nucleic acids & nucleotides.

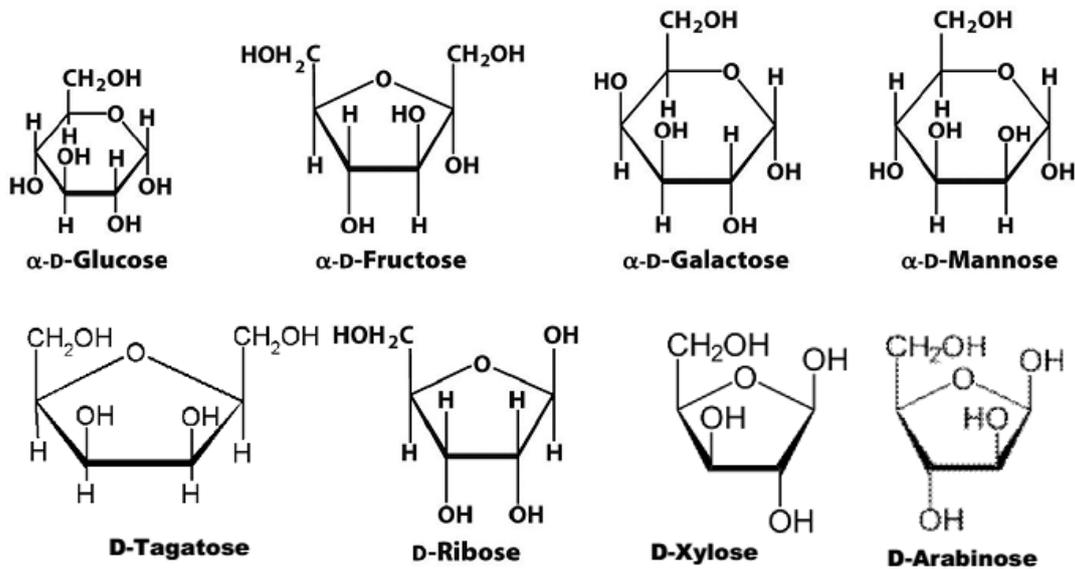


Figure 1.3: Hexose and pentose sugars. Different hexose sugars and pentose sugars are shown in their most common pyranose or furanose state.

1.4.1 Ribose metabolism

D-ribose is an abundant sugar that serves as an energy source as well as a component of RNA, DNA and many cofactors. It is one of the few compounds that bacteria can sense and actively seeks out through chemotaxis [4]. During uptake, ribose diffuses through the outer cell membrane into the periplasm, where it binds by the ribose-binding protein. The ribose-binding protein can then interact either with the membrane-bound chemotaxis receptor or with the permease that transports ribose across the inner membrane [4]. When D-ribose enters the cytoplasm of the cell, it must be phosphorylated to produce ribose-5-phosphate by Ribokinase (EC 2.7.1.15), which transfers the γ -phosphate group from ATP to ribose [5, 6, 7]. The product, D-ribose-5-phosphate, is then available for the synthesis of nucleotides, tryptophan and histidine, or for entry into the pentose phosphate pathway (figure 1.2). This enzyme is also important for recycling sugar produced by nucleotide breakdown.

Biochemical study on the catalytic mechanism of RK has revealed the requirement of divalent cations, such as Mg^{2+} , for the transfer of the phosphate group. Other divalent cations, such as Mn^{2+} , Co^{2+} , Ca^{2+} , Ni^{2+} , and Cu^{2+} , are able to play the same role as Mg^{2+} in the enzymatic reaction [8, 9, 10]. Deprived of these metal ions, the ribokinase, from either eukaryote or prokaryote, will lose its catalytic activity. In addition to divalent ions, monovalent cations have been known to play important roles in activation of ribokinase. It was demonstrated that *Escherichia coli* RK could be activated by monovalent cations such as potassium and cesium, with apparent K_d of 5 and 17 mM, respectively. In the absence of K^+ , the activity of RK is at least 60-fold lower. Unlike potassium and cesium, sodium and lithium have no effects on the enzymatic activity of RK [11].

1.4.2 Fructose metabolism

Fructose, an abundant hexose sugar found in many plants, is one of the main monosaccharides for bacterial growth in most ecosystems. Fructose provides an active participant of EMP pathway or glycolysis (**figure 1.1**) and fructolysis (The conversion of fructose to lactate; shown in **figure 1.4**) [12]. However, in contrast to glucose, fructose enters glycolysis at a step that bypasses the regulatory control exerted by phosphofructokinase. In addition, fructose can be used to generate glucose through gluconeogenesis. Fructose-6-phosphate provides precursor for murine (peptidoglycan) biosynthesis. For fructose to enter into the pathways of intermediary metabolism, it must first be phosphorylated to fructose-6-phosphate by fructokinase (EC 2.7.1.4).

The enzyme catalyzes ATP-dependent phosphorylation of fructose into fructose-6-phosphate but is not active with other sugars as phosphoryl acceptors. The fructokinase of *M. alcaliphilum* 20Z is most active in the presence of Mn^{2+} at pH 9.0 and 60°C, being inhibited by ADP ($K_i = 2.50 \pm 0.03$ mM). The apparent K_m values for fructose and ATP are 0.26 and

1.3 mM, respectively; the maximal activity is 141 U/mg protein. [13]. Hexokinase can also phosphorylate fructose, but it has a wide selectivity towards hexose sugars whereas fructokinase is highly specific towards fructose [14].

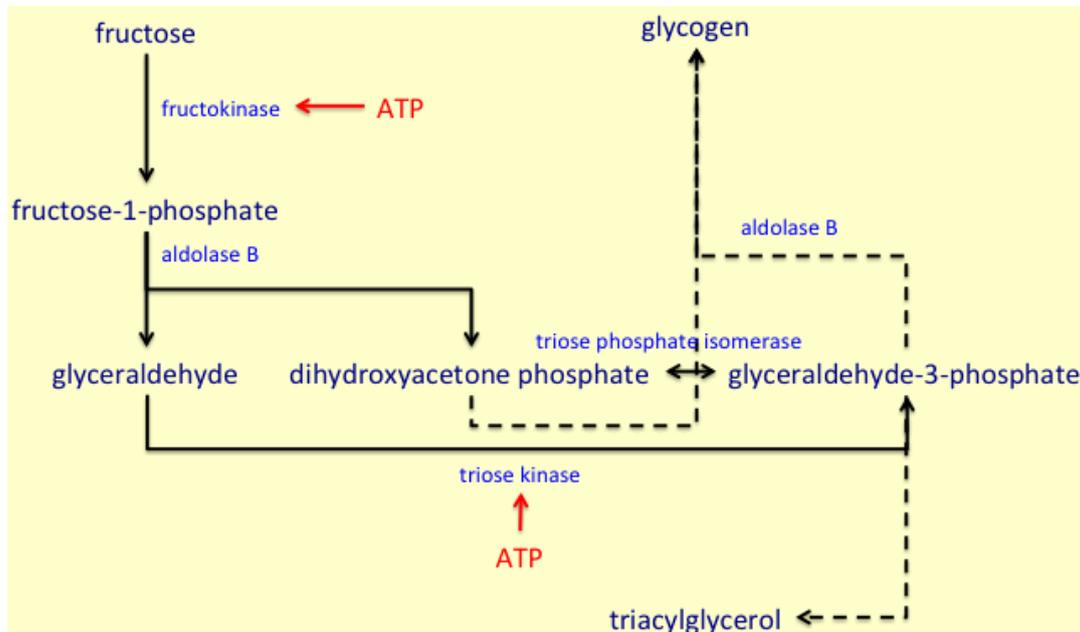


Figure 1.4: Steps involved in fructolysis and its final products.

1.5 Phosphorylation of sugar is the key step before metabolization

In most cases, Carbohydrates are phosphorylated before they are metabolized. The phosphorylation of sugars is the key step in energy production and biosynthesis. It is believed that phosphorylated intermediates are less likely to diffuse away through the cytoplasmic membrane [3]. Addition of the charged group prevents leakage of neutral sugars across biological membranes [15].

1.6 Bacterial kinase: sequence and structure classification

Kinases are a ubiquitous group of enzymes that participate in a variety of cellular pathways. By definition, the common name kinase is applied to enzymes that catalyze the transfer of the terminal phosphate group from ATP to an acceptor, which can be a small molecule, lipid, or protein substrate [16]. The cellular and physiological roles of kinases are

diverse. Many kinases participate in signal transduction pathways, in which these enzymes are essential components. Other kinases are involved centrally in the metabolism of carbohydrates, lipids, nucleotides, amino acid residues, vitamins, and cofactors. Some kinases have roles in various other processes, such as gene regulation, muscle contraction, and antibiotic resistance. Because of their universal roles in cellular processes, kinases are among the best-studied enzymes at the structural, biochemical, and cellular level. Despite the fact that all kinases catalyze essentially the same phosphoryl transfer reaction, they display remarkable diversity in their structures, substrate specificity, and in the number of pathways in which they participate [17].

According to several protein sequences and structural classification schemes such as SCOP [18], CATH [19], and Pfam [20] performed by Cheek, S. et.al., 2002, 2005 [17, 21], approximately 98% of over 17,000 kinase sequences have been cataloged into 25 families of homologous proteins, wherein 22 families (~98.8% of all sequences) for which three-dimensional structures are known fall into 10 fold groups. These fold groups not only include some of the most widely spread proteins folds, such as the Rossmann-like fold, ferredoxin-like fold, TIM-barrel fold, and antiparallel β -barrel fold, but also all major classes (all α , all β , $\alpha+\beta$, α/β) of protein structures.

1.7 Bacterial sugar kinase families

Although there is a rich diversity of structures, regulation modes, and substrate specificities among the sugar kinases, there are also common structural features. These conserved structural motifs provide clear indications as to how these enzymes manage to transfer the γ -phosphate of a purine nucleotide triphosphate to the hydroxyl groups of their sugar substrates. Depending upon overall structural folds, Bork et. al. (1993) [16] classified sugar kinases into three different superfamilies: the hexokinase family, ribokinase family

(also called the Phosphofructokinase-B (PfkB) family) and galactokinase family. The term superfamily is used to distinguish this broad collection of enzymes from smaller, more closely related subsets that have been commonly referred to as families. These enzymes use the γ -phosphate of ATP to phosphorylate hydroxyl oxygen of sugar atom. Ribokinase belongs to the ribokinase family, but fructokinase specificity has evolved independently in both the hexokinase and ribokinase families. This is an example of independent Darwinian adaptation of a structure to the same substrate at different evolutionary times [16].

According to the broad classification of kinases by Cheek et. al. [17, 21] discussed in **section 1.6**, ribokinase-like family belongs to the second fold group, named Rossmann-like group among the 10 fold groups. Hexokinase and galactokinase families belong to group IV (ribonuclease H-like) and group VI (GHMP kinase) respectively. The common structural feature of ribokinase-like family is that the architecture of their nucleotide-binding domain core is three layers ($\alpha/\beta/\alpha$) composed of $\beta\alpha$ repeats, with the central β -sheet mostly parallel. BLAST searches [22] suggest that, in order of sequence similarity to ribokinases, the other members are ketohexokinase, adenosine kinase, fructokinase, 2-dehydro-3-deoxy-gluconokinase, 1-phosphofructokinase and the minor 6-phospho-fructokinase (PfkB) [15].

1.8 Objective

Structure function studies of two sugar kinases, namely ribokinase and fructokinase from *Vibrio cholerae* O395, abbreviated as VcRK and VcFK respectively, which phosphorylate the pentose sugar ribose and hexose sugar fructose respectively have been discussed in this thesis. These two kinases are present in different bacterial species including *V. cholerae*. *V. cholerae* is a versatile bacterium that flourishes in diverse environments, including the human intestine, rivers, lakes, estuaries, and the ocean. In hospitable environment, they colonize to form biofilm, a surface-associated, three-dimensional,

structured bacterial community enclosed in an extracellular matrix [23]. Like other bacteria, *V. cholerae* uses glucose as universal energy source and other sugars like fructose, mannose and disaccharides such as sucrose and cellobiose. Bacterial surface colonization and biofilm formation are highly influenced by the type of carbohydrates available in the environment. These sugars are transported into cells by different transport system among which most important is phosphoenolpyruvate (PEP)-dependent sugar phosphotransferase system (PTS). It is a unique pathway in that it consists of a multienzyme phosphotransfer cascade that ultimately activates the transported sugar by phosphorylation [24].

Our aim is to solve the structure of VcRK and VcFK using X-ray crystallography. We have cloned, overexpressed, purified and crystallized these two proteins in apo form and in presence of different substrates with an idea to visualize substrate induced conformational changes and to identify key residues required for their enzymatic activity. Kinetic parameters have been determined for FK to establish the mechanism we proposed based on our structural studies.

1.9 Determination of protein structure by X-ray crystallography

3D structures of biological macromolecules are generally determined by X-ray crystallography, Nuclear Magnetic Resonance spectroscopy or Cryo-electron microscopy techniques. Among those, X-ray crystallography is the most widely used method to visualize protein structures at the atomic level. When a protein crystallizes, it is arrested in one of its most stable conformations. Interaction of protein with ligands, other proteins and even DNA and associated structural changes can also be explored by this method. Starting from the first crystal structure of Myoglobin by John Kendrew in 1962 [25], the Protein Data Bank (www.rcsb.org/pdb/) has been enriched by more than 1,13,000 macromolecular structures till now among which 89% are solved by X-ray crystallography. The use of protein structure

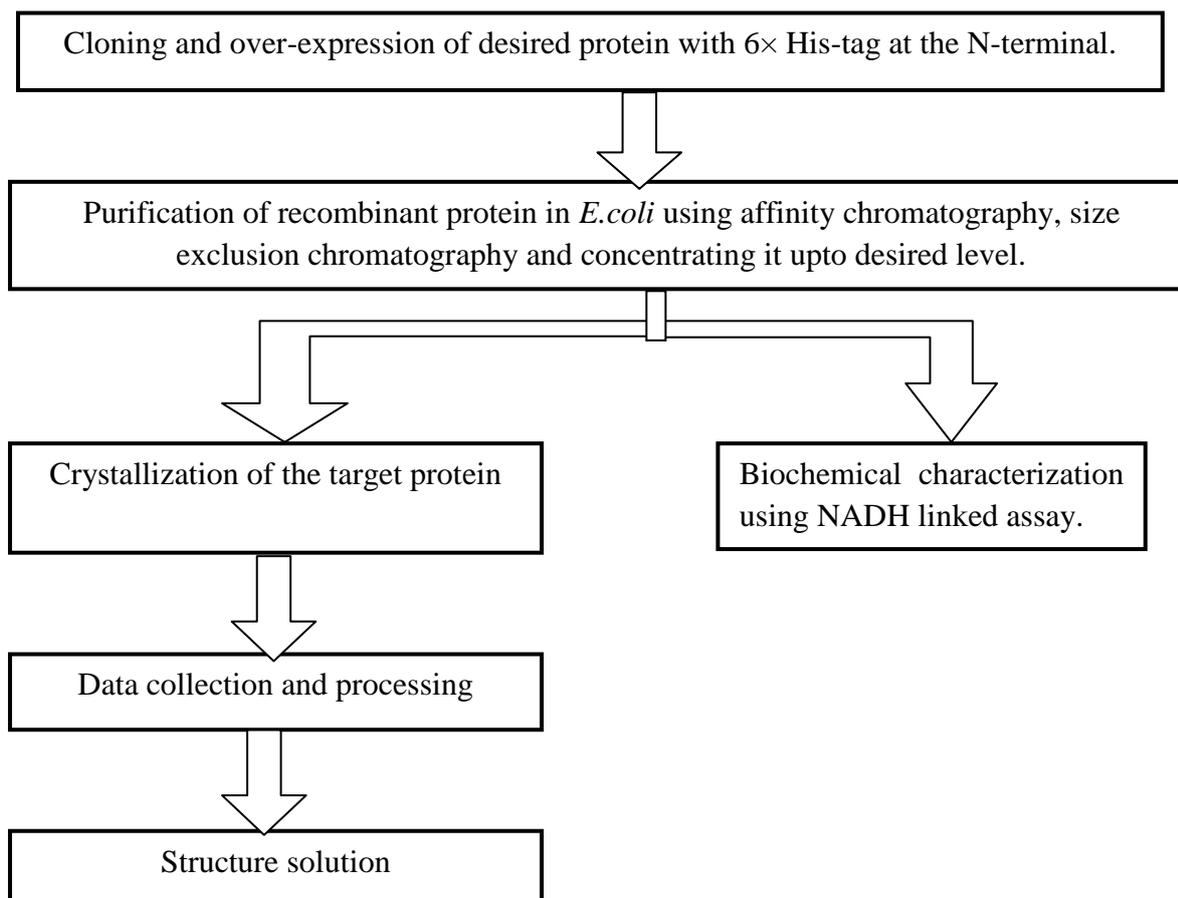
information is currently widely spread within many areas of science and industry, among which are biotechnology and pharmaceutical industry. Thousands of papers describing such structures have been published in the scientific literature, and 20 Nobel Prizes in chemistry or medicine have been awarded for discoveries based on macromolecular crystallography [26].

The steps involved in determination of crystal structure of a protein starting from the DNA level have been discussed thoroughly in **chapter II**. All that we need is a sufficiently large single crystal of the target protein, which produces diffraction data of good intensity and resolution [26]. The diffraction data actually contains the information of intensity of the crystal in reciprocal space. Combination of intensity and proper phase produces the electron density of actual crystal structure. There are different methods for determination of initial phase [27] such as (i) Ab initio phasing or direct method (suitable for very high-resolution data, better than 1.4 Å) [28], (ii) Molecular replacement (suitable for proteins having >25% sequence identity and <2.0 Å r.m.s.d with a known structure from PDB) [29], (iii) Isomorphous replacement (SIR, MIR) with two or more heavy atom (Au, Pt, Ag, Hg etc) derivatives [30] or (iv) Anomalous scattering (SAD/MAD) with anomalous scatterers at different wavelengths [31, 32]. The initial phase can be improved by coordinate refinement, restrained ADP (atomic displacement parameter), occupancy refinement and bulk solvent correction. The value of observed and calculated structure factors are compared by two R-factors (reliability factor), R_{work} (calculated from 95% data) and R_{free} (calculated from remaining 5% data) which reduce in gradual steps of refinement [33]. The model is validated to assess the values of bond lengths and bond angles within acceptable limit (according to Ramachandran plot). The final model (*.pdb) along with the scaled reflection file (*.mtz) is deposited in Protein Data Bank (PDB).

Chapter- 2

MATERIALS & METHODS

The following general scheme is followed to elucidate the structure and function relationship of two sugar kinases, namely Ribokinase (VcRK) and fructokinase (VcFK). Starting from the DNA level to the final protein structure (or function) each different experimental steps adopted are depicted with the help of this flow diagram.



2.1. Cloning and over-expression of proteins with 6× His-tag at the N-terminal

The two target genes from *V. cholerae* 0395, namely *cscK* (coding Fructokinase, VcFK; uniprot ID A5EZZ9) and *rbsK* (coding Ribokinase, VcRK; uniprot ID A5F1B7) were cloned into pET-28(a+) vector to get N-terminal 6×-His tagged protein construct.

2.1.1 Amplification of gene sequence by PCR

The genomic DNA of *V. cholerae* O395 was used as template. For wild type protein, the desired sequences were amplified by PCR [34] with the help of appropriate forward and reverse primers (Table 1). For mutant proteins, Site Directed Mutagenesis (SDM) experiments was performed by two-step PCR with suitable mutation forward and mutation reverse primers, in addition to gene forward and gene reverse primers.

Table 2.1: primer sequences for PCR of wild type and mutant proteins

Clone name	Primer sequences
VcRK (wild type)	G ^f P: 5'-GGAATTCC <u>CATATG</u> AATAAGTTGGTGGTTCTGGGTAGCGTC-3'
	G ^r P: 5'-GGGGATCCTTATAAATCAACCTATGAGTGTTCTGCTAAAAA-3'
VcFK (wild type)	G ^f P: 5'- GGAATTCC <u>CATATG</u> AAGGCTTTGGTTAGACTCAGC-3'
	G ^r P: 5'- GGGGATCCTTATAATTCTAAAAAGGCATACAGCGCCG -3'
VcFK (E110Q)	M ^f P: 5'- AAGTTTTTCGGAGCAGGAATTACAGTTT -3'
	M ^r P: 5'- AAAGTGTAAATTCCTGCTCCGAAAAGTT -3'
VcFK (E110Q)	M ^f P: 5'- AAGTTTTTCGGAGCAGGAATTACAGTTT -3'
	M ^r P: 5'- AAAGTGTAAATTCCTGCTCCGAAAAGTT -3'

G^fP-gene forwarding primer, G^rP-gene reverse primer; M^fP-mutation forward primer, M^rP-mutation reverse primer; Restriction enzyme used in forward- NDEI (CATATG) and reverse- BAMHI (GGATCC) primers are underlined.

2.1.1.1 Steps and ingredients involved in Polymerase Chain Reaction

For PCR, annealing temperature (T_m) was chosen approximately 5K below the theoretical melting temperature of both primers. Initially, gradient PCR was performed to choose best annealing temperature, which helped to obtain maximum yield at the second step

PCR. Composition of reaction mixture used for PCR has been given in **Table 2**. The following program was used for both steps of PCR.

- Initial denaturation at 95 °C for 5 minutes
- [Denaturation at 95 °C for 30 seconds
- Annealing at $X \pm 5$ °C for 30 seconds (X varies for different constructs)
- Polymerization at 72 °C for 45 seconds]^{30 cycles}
- Final extension at 72 °C for 10 minutes

Table 2.2: Composition of PCR reaction mixture

Components	Final concentration
Template- <i>V.cholerae</i> genomic DNA	~20ng
Forward primer	0.8 pmol
Reverse primer	0.8 pmol
dNTP mix- 100X	1X
10X PCR buffer (Fermentas)	1X
<i>Taq</i> polymerase (Fermentas)	2U
Autoclaved water to make final volume 100µl.	

2.1.1.2 Site directed mutagenesis (SDM)

SDM by two-step PCR method [35] was used to prepare VcFK single mutants E110Q and E205Q. In first step, proper annealing temperatures were determined by gradient PCR in two sets- one with ‘gene forward’ and ‘mutation reverse’ primers and other set containing ‘gene reverse’ and ‘mutation forward’. The PCR products from two sets were taken as templates for the second step PCR. The reaction mixtures were given in **Table (3)**.

Table 2.3 Calculation of the two-step PCR reaction mixture

Components	Final concentrations 1 st step PCR	Components	Final concentrations 2 nd step PCR
Vc-genomic DNA	~ 20ng	1 st step PCR product	~10ng each
Gene forward and mutation reverse primer-set-I	0.8 pmol each	Gene forward primer	0.8 pmol
Gene reverse and mutation forward primer-set II	0.8 pmol each	Gene reverse primer	0.8 pmol
dNTP mix- 100X	1X	dNTP mix- 100X	1X
10X PCR buffer (Fermentas)	1X	10X PCR buffer (Fermentas)	1X
<i>Taq</i> polymerase (Fermentas)	2U	<i>Taq</i> polymerase (Fermentas)	2U
Autoclaved water to make final volume 100µl.			

2.1.2 Digestion and ligation of PCR product

PCR product was purified with Qiagen ‘PCR purification kit’ followed by digestion with restriction endonucleases *NDEI* and *BAMHI* and purification again to create gene insert. pET-28(a+) vector (resistant to kanamycin) was digested with the same restriction enzymes and purified to create the ‘sticky ends’ at target vector. The gene insert was ligated into the digested pET-28(a+) vector.

2.1.3 Transformation of plasmid by electroporation

Ligation product was transformed into electro-competent cells *E. coli* XL1-blue (tetracycline resistant, Stratagene) by electroporation technique. The transformed cells were

cultured on agarose plates containing 30 µg/ml kanamycin and 12.5 µg/ml tetracycline. Colonies were obtained for *E. coli* XL1B cells that contain the cloned vector.

2.1.4 Checking for successful cloning

For further confirmation, plasmids were isolated from fresh culture of cloned *E. coli* XL1-blue cells by mini preps using 'alkaline lysis' method [36] and analysed on agarose gels after digestion with the corresponding enzymes (*NDEI* and *BAMHI*). The sequence of the clones has been confirmed by sequencing.

2.1.5 Overexpression of cloned gene

For over-expression of desired Histidine-tagged protein, pET-28(a+) plasmids (kanamycin resistant) isolated previously from *E. coli* XL1-blue cells were transformed into *E. coli* expression strain BL21(DE3) (no antibiotic resistance). The transformed cells were cultured in agarose plates containing 30µg/ml kanamycin.

Four single colonies from transformed cells were inoculated in individual falcons containing 2ml Luria Broth medium along with 30µg/ml kanamycin, and grown overnight at 310K. Next day, each of these overnight cultures was inoculated freshly in 1:50 ratio to 3ml Luria Broth medium along with 30µg/ml kanamycin. Cells were allowed to grow at 310K at 175 rpm shaking until OD600 reaches to ~0.4-0.5. Each culture was then divided into two equal aliquotes, one part was induced by IPTG (Isopropyl β-D-1-thiogalactopyranoside) with final concentration 1 mM, and the other part was kept as uninduced control. All aliquots were allowed to grow for another 4 hours at same temperature. Next, cells were harvested by centrifugation at 12000g for 2 min, and the pellets were suspended in 30µl de-ionized water, vortexed and mixed with 10 µl of 4X sample buffer for SDS (250 mM Tris-HCl pH 6.8, 10% SDS, 40% glycerol, 5% Bromophenol blue and 20% β-mercaptoethanol). Cells were lysed by

boiling for 5 min. in water bath, centrifuged at 12000g to separate the cell-debris, and then the clear supernatants of both induced and corresponding uninduced samples were loaded one beside another in 12% SDS Page run at 200V, 20 mA. After run, the gel was stained with Coomassie R-250 (Sigma). Over-expression of recombinant protein was confirmed by comparison of induced samples with corresponding uninduced controls and molecular weight marker.

2.1.6 Storage for future use

A part of overnight culture containing cloned *E. coli* BL21 (DE3) cells were treated with glycerol and stored at 193 K for future use.

2.2 Purification of recombinant protein followed by concentrating it for crystallization.

For structure solution and biochemical characterizations, the 6×-Histidine tagged protein was purified from a large scale culture of *E. coli* BL21 (DE3) cells and finally the 6×-Histidine tag was removed with the help of following protocols.

2.2.1 Cell Lysis and Removal of Crude Lysate

E. coli BL21 (DE3) cells containing cloned plasmid were grown overnight at 310 K in 5 ml Luria Broth medium (Himedia) in presence of 30 µg/ml kanamycin (SRL). Cells were inoculated freshly at 500 ml Luria Broth media with 30 µg/ml kanamycin and were grown in mid-log phase at 310 K until the OD₆₀₀ reached 0.5. Then cells were induced with 0.1 mM IPTG (SRL) prior to incubation for 16 h at 293 K. Cells were harvested by centrifugation at 4500g for 20 minutes and resuspended in lysis buffer containing 50 mM Tris pH 8.0, 300 mM NaCl at 273 K. The cells were lysed by the addition of lysozyme followed by sonication on ice. The cell debris was discarded by centrifugation at 14000 g for 45 minutes and the clear supernatant containing protein was collected.

2.2.2 Removal of Non-Specific Proteins by Ni²⁺-NTA Affinity Chromatography

The clear supernatant containing 6×His tagged protein was applied to a glass column containing Ni²⁺-NTA (Ni²⁺-NTA Agarose - QIAGEN) resin pre-equilibrated with lysis buffer and kept for few minutes. The weakly bound and non-specifically bound proteins were removed by washing the column with wash buffers containing 5 mM and 10 mM imidazole respectively. The strongly bound protein was eluted by elution buffer containing 150 mM imidazole (SRL), 50 mM Tris (MERCK) pH 8.0, 300 mM NaCl (SRL) by competitive binding with the Ni²⁺-NTA resin. The protein was concentrated and imidazole concentration was serially reduced below 10 mM by continuous dilution with lysis buffer and centrifugation using molecular weight cut-off membrane (MW cut off 10 kDa) at 277 K.

2.2.3 Removal of 6×-Histidine tag by Thrombin Cleavage

Thrombin recognizes the consensus sequence Leu-Val-Pro-Arg-Gly-Ser, cleaving the peptide bond between Arg and Gly. The protein contains 6× His-tag followed by this cleavage sequence prior to N-terminal amino acid residue. The enzymatic cleavage reaction is influenced by temperature, Ca²⁺ ion and is inhibited by imidazole and finite concentration of DTT/BME. Therefore, the cleavage time and thrombin concentration was standardized for both proteins. For 5-10 mg VcRK and VcFK, 2 units of thrombin (Novagen) removes the tag in 40 hr and 32 hr respectively.

2.2.4 Removal of Impurity by Size Exclusion Chromatography

After thrombin cleavage, the thrombin and cleaved tag was separated from the cleaved protein by size exclusion chromatography using Shephacryl S-100 (GE Healthcare Bioscience). The protein solution containing all impurities was directly added to a Shephacryl S-100 column of dimension 85 cm× 1.5 cm pre-equilibrated with respect to gel filtration buffer. The column was pre calibrated with protein mixture of different molecular mass. The elution profiles were recorded in a chart recorder based on the absorbance profile at 280 nm.

The eluted fractions were analyzed by SDS- PAGE, pooled and were concentrated to desired level. Protein concentration was checked by measuring OD₂₈₀, Bradford assay (Bio-Rad) and SDS-PAGE.

2.3 Crystallization, Data Collection, Data Processing and Structure Solution

X-ray crystallography enables us to visualize protein structures at the atomic level and enhances our understanding of protein function. Specifically we can study how proteins interact with other molecules, how they undergo conformational changes, and how they perform catalysis in the case of enzymes. Armed with this information we can design novel drugs that target a particular protein, or rationally engineer an enzyme for a specific industrial process [26].

The concentrated proteins obtained through preceding steps (**Sections 2.1 and 2.2**) were used for crystallization. We shall first mention the standard protocols adopted for crystallization. Details of the concentration of buffer, precipitant, substrate, ADP etc will be discussed for each individual case. The following standard protocols were used for cloning, purification and crystallization.

2.3.1 Crystallization

Crystallography technique is widely used to arrest the protein in its most stable conformation using various additives and to determine its structure by X-ray diffraction. The diffraction from a single molecule would be too weak to be measurable. Therefore, we use crystal, an ordered three-dimensional array of molecules, to magnify the signal. If the crystal is well ordered, then diffraction will be measurable at high angles or high resolution that can produce intricate structural insights.

For crystallization, we need the protein in pure and homogeneous form. Impurity results poor internal order of the crystals that diffracts to low resolution and consequently

yield less structural information. Non-reducing SDS PolyAcrylamide Gel Electrophoresis (SDS-PAGE, without β -mercaptoethanol, to determine the presence of Cysteine mediated dimerization) and native PAGE methods (to check the presence of oligomers) have been done prior to crystallization. The protein must remain in soluble and folded form at crystallization buffer. The degree of ordered secondary structures was tested with circular dichroism spectroscopy. Ideally, the sample needs to remain stable over that period. Occasionally good protein crystals will form overnight at room temperature, but usually it may take several days to one or two weeks before crystals of suitable size can grow. Our proteins remained stable for 3-4 weeks at 253 K.

Some of the factors, which may affect protein solubility (and protein crystallization) are listed here: i) pH, ii) Ionic strength, iii) Concentration of precipitant, iv) Concentration of macromolecule, v) Temperature, vi) Additives, effectors, and ligands, vii) Organism source of macromolecules, viii) Presence of substrates, coenzymes, inhibitors, ix) Reducing or oxidizing environment, x) Metal ions, xi) Rate of equilibration [37].

The concentrated protein was used for initial crystallization trial by vapor diffusion method using 24 Well Crystallization Plates made by Hampton Research. Vapor deposition methods are usually performed in two ways, by hanging drop and sitting drops.

In either vapor diffusion experiments a protein droplet containing purified protein, buffer and precipitant is in a closed well with a reservoir of similar buffers and precipitants, but whose concentrations are higher than in the protein droplet. The reservoir does not contain the protein. At the beginning of the experiment, the protein and precipitant in the protein droplet are in too low a concentration to cause the protein to crystallize, and the system is undersaturated. During the course of the experiment, water vaporizes from the protein droplet and collects in the reservoir. The decrease in water in the protein droplet

moves the protein droplet from a state of undersaturation to a state where the protein and precipitant concentration in the protein droplet are sufficiently high that protein crystallization occurs in the protein droplet. The process can be understood clearly with the help of the phase diagram of protein crystallization [38], depicted in **Figure 2.1**.

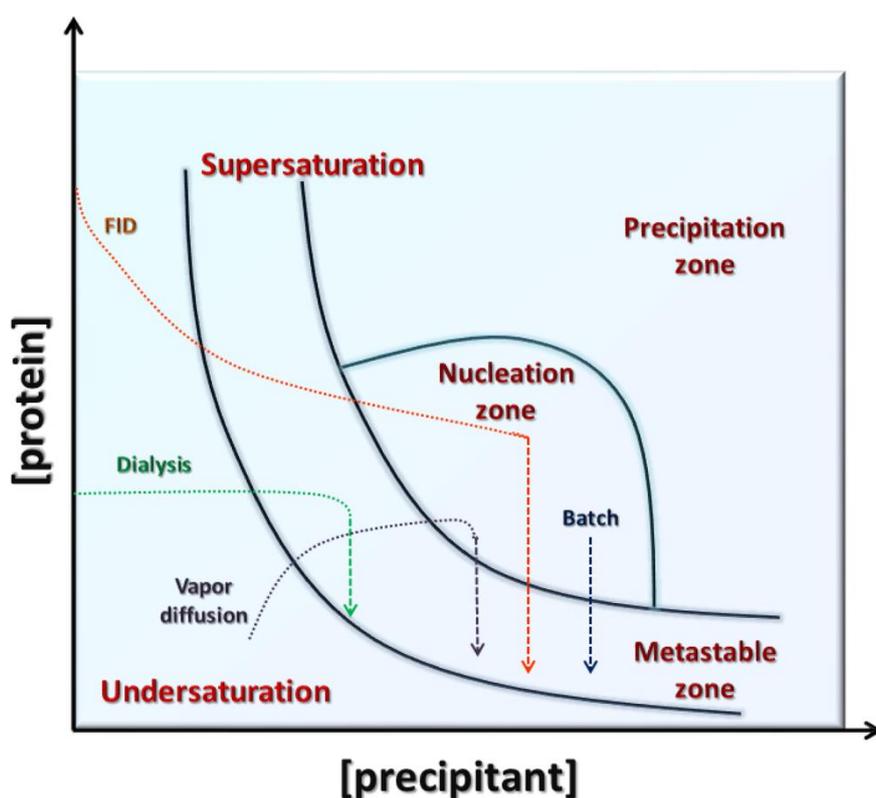


Figure 2.1: The phase diagram in protein crystallization. The curves represent variation of protein concentration with concentration of precipitating agent. The solubility curve separates supersaturation zone from undersaturation zone. Once a protein reaches nucleation zone, crystal nuclei are formed. Then protein concentration falls and phase point shifts downward. If it remains in metastable zone, crystal size increases. The figure has been adopted from **Int. J. Mol. Sci. 2013**.

2.3.2 Data collection

Single crystal were fished out from the crystallization drops with the help of a 10-20 micron nylon loop (Hampton Research, Laguna Niguel, California, USA) and soaked in appropriate cryoprotectant solution for cryogenic data collection, A cold nitrogen gas stream

(Oxford Cryosystem) keeps the crystal at 100 K throughout the experiment. For some crystals diffracted were collected using our home source X-ray diffractometer and for others diffraction data were collected at BM14 ESRF-EMBL-India beamline Synchrotron.

In home source, the loop containing crystal was flash-cooled to 100 K and mounted to goniometer head, which enables the crystal accurately positioned in the X-ray beam by means of a number of adjustment screws. Focused X-ray of wavelength 1.54 Å (Cu-K α) was generated by Bruker-Nonius FR591 rotating anode generator, equipped with Osmic MaxFlux confocal optics and running at 50 kV and 50 mA. The ray emerged from a narrow tube called a collimator and struck the crystal to produce a diffraction pattern that was recorded on the MAR Research image plate detector of diameter 345 mm.

Synchrotron diffraction data were collected on a MAR 225 CCD detector (1 sec readout time - 74 micron pixel - mosaic 9 chips) on BM14 at the European Synchrotron Radiation Facility (ESRF), Grenoble. In ESRF, crystals are loaded manually or via the sample changer (SC3) and maintained under a 100K cryo-stream jet (Oxford-CryoSystem). The incident X-ray beam had a wavelength of 0.97 Å and 360 frames were recorded with non-overlapping 1° oscillations. For apo VcRK crystals 12 s exposure per image was used, whereas for the co-crystals the exposure time was only 1 s.

2.3.3 Data Processing

Data were processed and scaled using *iMOSFLM* [39] and *SCALA* [40]. In *iMOSFLM*, diffraction images are accumulated for ‘Index’-ing to determine the possible crystal lattice parameters and to estimate crystal mosaicity. Then ‘cell refinement’ task was done to allow the refinement of cell parameters, crystal orientation and mosaicity based on a post-refinement procedure [41, 42] that provides values that are more accurate. This is applicable only when the resolution of the data is better than ~ 3.5 Å. Finally, ‘Integration’ step was

performed to refine various parameters [43]. The output file was then used in *SCALA* to merge multiple observations of reflections and produced a file that contains averaged intensities for each reflection.

2.3.4 Initial Structure Determination

In order to visualize the structure the phase problem needed to be solved. ‘Molecular Replacement’ (MR) method [29] solved this purpose where the coordinates of the highest homologue to the target protein was taken as a ‘model’ for phasing. In this method the Patterson map, which is an inter-atomic vector map, was derived from data obtained from the unknown structure and from the structure of a previously solved homolog. Now use of ‘Rotation Functions’ to determine the orientation of the molecule followed by a ‘Translation Function’ [44], these two maps were oriented in correct position in the unit cell. Following this correctly oriented and translated phasing models, phases were calculated to derive an electron density map, which was then used to build and refine in subsequent steps to get an atomic model of the unknown structure.

The model was used as input to the ‘Automated MR mode’ of the program *PHASER* [45] in CCP4 suite [46]. The Automated MR mode combines i) anisotropy correction i.e. removing directional dependence of the diffraction data, ii) likelihood enhanced fast rotation function, iii) likelihood enhanced fast translation function [47], iv) packing and refinement modes for multiple research models and v) a set of possible space groups to automatically solve a structure by MR.

2.3.5 Refinement and model building step

The initial phases are generally quite poor and it is often difficult to start building the model. Refinement of electron density map helps to improve the phases and better interpretation of structure. Refinement was done through statistical adjustment of the

coordinates to fit the diffraction data better. As a measure of the fitting the ‘working R-factor’ is generally used that measures how far the calculated amplitude differ from the observed amplitudes. The R_{work} can get trapped in local minima giving the false impression of having a good model. For this reason, R_{free} is mostly used to validate the refinement process [33]. The factor is calculated from 5% of the reflections that are excluded from refinement and it therefore gives an independent measure of the refinement progress. Subsequent refinement cycles were then carried out with *Phenix Refine* in Phenix suite [48] and with *REFMAC5* [49] in CCP4 suite to maximize the agreement between model and the X-ray data. The progress in refining the model is measured by R_{work} and R_{free} . Two methods widely used in refinement are i) Maximum likelihood and ii) Simulated annealing. Both methods use restraints to how an atomic model has to look like with respect to bond distances, angles, torsions and temperature factors (B-factors). In maximum likelihood, the phases are adjusted to minimize the R-factor. In simulated annealing the structure is heated to add randomness and slowly cooled and refined. The randomness reduces the probability of falling into wrong local minima. Additional methods like non Crystallographic Symmetry (NCS) [50] restraints average the phases in symmetrical parts (different molecules) of the asymmetric unit (AU) and are applied in case of multiple copies of molecules in the AU. Density modification aims to adjust the density to the expectations of how it should generally look like. The solvent does not diffract normally and the electron density should therefore be zero in the solvent region. In protein often the temperature factors are averaged in all dimensions (isotropic) instead of individual (anisotropic). TLS refinement for Translation, Liberation (small movements) and screw rotation of a group of atoms can give a good approximation of anisotropy with much fewer parameters [51]. The model building was done with COOT [52].

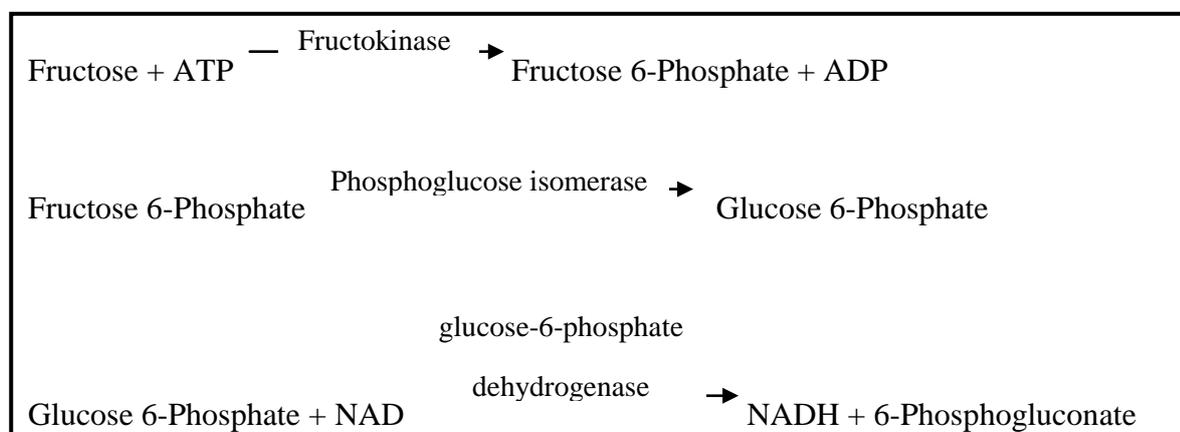
2.3.6 Structure analysis

Average B-factor for each residue was calculated using B_{average} in CCP4 [46]. The oligomeric state of the protein was analyzed using the PISA web server [53]. Sequence alignments of proteins were done by CLUSTALW [54, 55] and MULTALIN [56]. Figures were prepared using Pymol [57].

2.4 Determination of enzymatic activity

Kinetic parameters of VcFK and its mutants were calculated using the substrate fructose and ATP by a 3-step reaction as described below, whose outcome is reduction of NAD to equimolar amount of NADH and consequent increase in absorbance at 340 nm [58].

Fructose (Sigma Aldrich) at a concentration range of 0.05 to 1 mM is phosphorylated by adenosine triphosphate (ATP) (Sigma Aldrich) in the reaction catalyzed by fructokinase. Fructose 6-phosphate is then converted to glucose 6-phosphate (G6P) by enzyme Phosphoglucose isomerase (PGI). Glucose-6-phosphate (G6P) (Sigma Aldrich) is then oxidized to 6-phosphogluconate in the presence of glucose-6-phosphate dehydrogenase and consequently nicotinamide adenine dinucleotide (NAD) (Sigma Aldrich) is reduced to NADH. The reaction steps are depicted as following [58].



The reaction mix contained 0.025-1 mM fructose, 0.5-1 mM ATP (Sigma Aldrich), 0.5-1 mM MgCl₂ (Sigma Aldrich) 0.5 mM nicotinamide adenine dinucleotide (NAD) (Sigma Aldrich), 1U phosphoglucose isomerase (PGI) (Sigma Aldrich), 1U glucose-6-phosphate dehydrogenase (G6PDH). The reaction initiated by the addition of 56 μM VcFK enzyme. Absorbance data were collected at steady state (10 min. after initiation of reaction) at 340 nm in JASCO V630 spectrophotometer.

CHAPTER- 3

**CLONING, OVER-EXPRESSION,
PURIFICATION,
CRYSTALLIZATION, DATA
COLLECTION AND DATA
PROCESSING OF V_cRK AND V_cFK
AND ENZYME KINETICS OF V_c-FK.**

3.1 Cloning and over-expression of VcRK and VcFK

VcRK and VcFK were successfully cloned in pET-28a (+) and over-expressed as N-terminal 6× Histidine tagged recombinant proteins. The cloned plasmids from separate cultures were digested with the same restriction endonucleases used for cloning (*NdeI* and *BamH1*), and the resultant DNA fragments were compared with initial PCR product in 1% agarose gel. **Figure 3.1a** and **b** show the cloned DNA fragments of VcRK and VcFK respectively with corresponding PCR products. The over-expression profiles of both proteins are given in **figure 3.1c** and **d**.

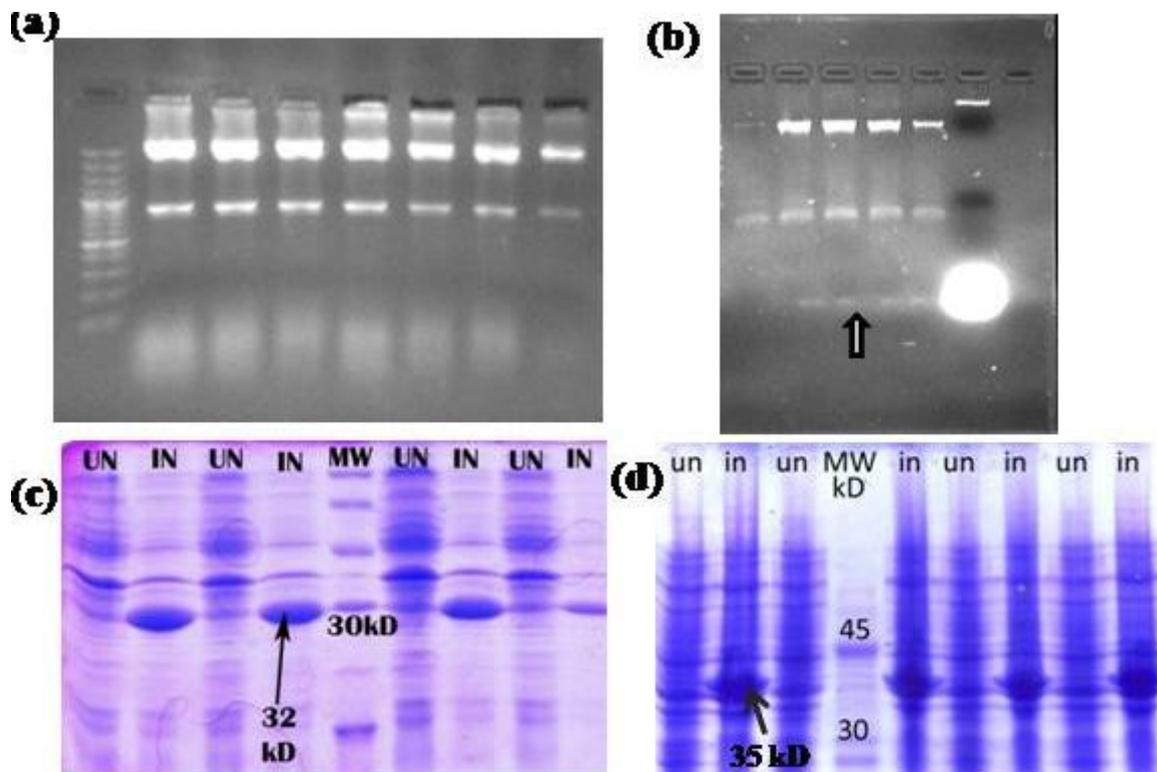


Figure 3.1: (a and b) Size comparison of DNA fragments before and after cloning in 1% agarose gel. (a) VcRK: Lane-1: marker, Lane 2-8: digested cloned plasmid samples containing the insert. (b) VcFK: Lane 2-5: digested cloned plasmid samples containing the insert, lane 6 is the PCR product of insert before clone. (c and d) 15% and 12% SDS PAGE gel profiles, illustrating over-expression of (c) VcRK (32 kDa) and (d) VcFK (35 kDa) respectively. Lanes marked as 'IN' denotes induced samples whereas corresponding uninduced samples are marked as 'UN'. The positions of induced bands are marked with respect to molecular weight (MW) marker.

3.2 Purification of 6×Histidine tagged proteins

The supernatant obtained after sonication and centrifugation contained Histidine-tagged over-expressed protein along with other impurities. In order to purify, the whole supernatant was loaded to a column containing Ni-NTA bead and was successfully eluted from the Ni-NTA column using buffers containing increasing concentration of imidazole. The elution profiles of VcRK and VcFK are given in **figure 3.2a** and **b** respectively.

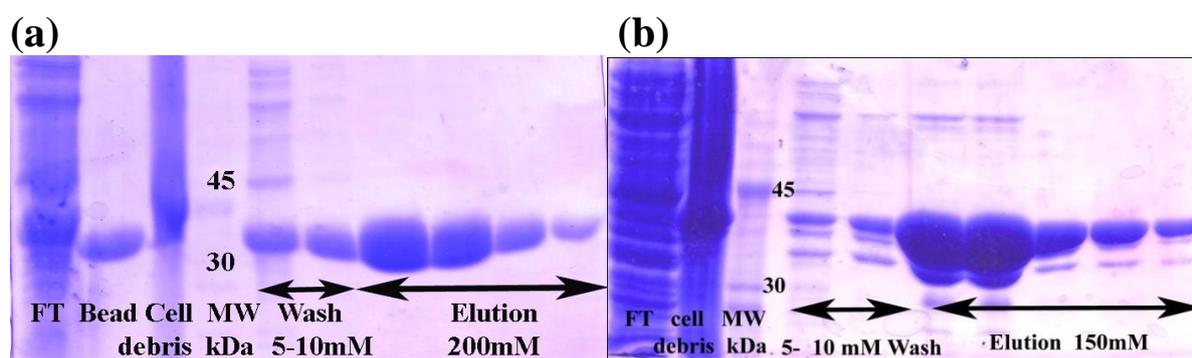


Figure 3.2: 12 % SDS-PAGE profile showing Ni-NTA purification profile of (a) VcRK and (b) VcFK. (a) Lane 1: flow through (FT) collected after passing the cell supernatant onto the buffer equilibrated Ni-NTA column. Lane 2: Ni-NTA bead after elution by imidazole. Lane 3: resuspended pellet after sonication and centrifugation. Lane 4: molecular weight marker. Lanes 5-6: samples collected after washing the column with buffer containing 5mM and 10mM imidazole respectively. Lanes 7-10: eluted fractions containing 200 mM imidazole buffer.

(b) Lane 1: flow through (FT) collected after passing the cell supernatant onto the buffer equilibrated Ni-NTA column. Lane 2: resuspended pellet after sonication and centrifugation. Lane 3: molecular weight marker. Lanes 4-5: samples collected after washing the column with buffer containing 5mM and 10mM imidazole respectively. Lanes 6-10: eluted fractions containing 150 mM imidazole buffer.

3.3 Removal of Histidine tag by thrombin: standardization of cleavage time

Prior to size exclusion chromatography using Sephacryl S-100 column, the 6×Histidine tag were removed by thrombin. Because each target protein presents the cleavage

site somewhat differently, it is recommended to test several thrombin concentrations, temperatures and/or incubation times to optimize specificity and efficiency of cleavage. In 277K, both proteins in thrombin cleavage buffer were incubated with thrombin and aliquots were taken at different time intervals to check the progress of the cleavage reactions. The SDS-PAGE profiles of the thrombin cleavage profile at different time are given below (**figure 3.3**).

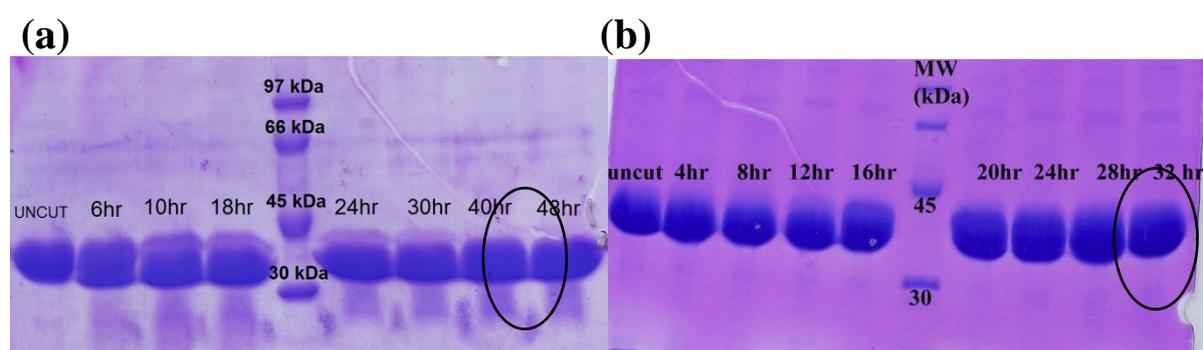


Figure 3.3: Standardization of cleavage time using thrombin in (a) VcRK and (b) VcFK.

(a) lane-1: sample collected before addition of thrombin. Lanes 2-4, 6-9 contains reaction mix collected at 6 hr, 10 hr, 18 hr, 24 hr, 30 hr, 40 hr and 48 hr respectively. Lane 5 contains molecular weight marker. Here cleavage completes at 40 hrs.

(b) lane-1: sample collected before addition of thrombin. Lanes 2-5, 6-10 contains reaction mix collected at 4 hr, 8 hr, 12 hr, 16 hr, 20 hr, 24 hr, 28 hr and 32 hr respectively. Lane 6 contains molecular weight marker. Here cleavage completes at 32 hrs.

3.4 Checking concentration of final protein

After gel-filtration and concentration, proteins were obtained in concentrated form. To check purity and state of cysteine mediated oligomerization of protein samples, SDS PAGE were performed in presence and absence of β -mercaptoethanol. In VcRK, dimer obtained in absence of β -mercaptoethanol, therefore β -mercaptoethanol was added to break disulfide bonds. Results have been in **figure 3.4**.

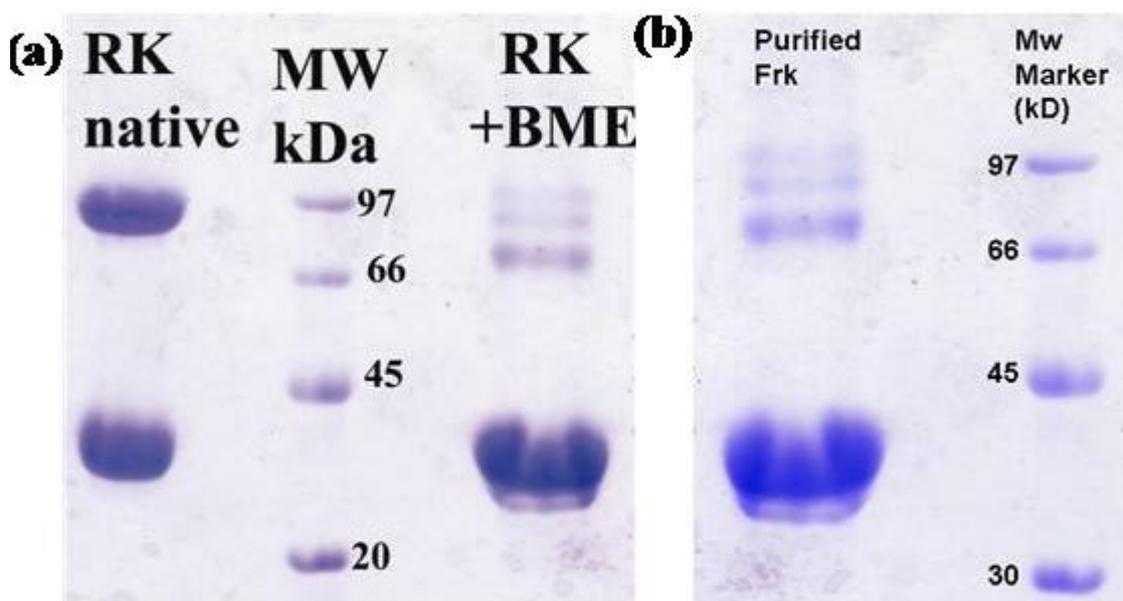


Figure 3.4: SDS-PAGE profile for checking concentration and impurity level (a) VcRK (32 kDa) and (b) VcFK (35 kDa).

3.5 Crystallization of concentrated proteins

Initial crystallization trials for both unliganded and sugar-ADP bound VcRK VcFK were performed by the hanging-drop vapour-diffusion method at 293 K and 277 K with Grid Screen Ammonium Sulfate, Grid Screen PEG 6000, Crystal Screen and Crystal Screen 2 from Hampton Research [59]. VcRK and VcFK were attempted to crystallize both in presence and in absence of 6×Histidine tag. Crystals of 6×Histidine tagged protein were of very poor quality and did not diffract, therefore all good quality crystals were obtained after removal of the tag.

3.5.1 Crystallization of VcRK

Initial trials using PEG 6000 as precipitant always produced very tiny crystals covered with protein skin, but the size and quality of the crystals could not be improved further by using different protein: precipitant ratios or reservoir strength. Although fresh protein samples

appeared to be homogeneous, a nonreducing 10% (w/v) SDS-PAGE of few-days-old samples indicated a heterogeneous population with the appearance of higher oligomeric species. However, these higher oligomers disappeared on a reducing gel (**figure 3.4a**). The amino-acid sequence of VcRK contains five cysteine residues and there is a strong probability that unpaired cysteines form disulfide-linked oligomers. Therefore, we maintained a β -mercaptoethanol concentration of 5 mM in all crystallization experiments. For complex crystals, the protein was incubated with ribose (5 mM) and ADP (5 mM) for 1 h before crystallization setup.

Finally, diffraction quality crystals were obtained within 7 days at 293K in hanging drop for both apo and complex proteins where 1-3 μ l A3 of grid screen PEG 6000 (Hampton Research) (5% (w/v) PEG6000, 0.1 M Mes pH 6.0) was used as precipitant with 2-3 μ l protein. A 600 μ l solution containing 30-35% (w/v) PEG6000, 200 mM NaCl, 100 mM Tris was used as reservoir. Crystals of VcRK is shown in **figure 3.5**. Cs^+ -bound crystals were obtained by soaking ribose- and ADP-bound VcRK crystals with a cryoprotectant containing 20 mM CsCl. Detailed crystallization conditions of VcRK in different ligand bound form had been shown in **Table 3.1**.

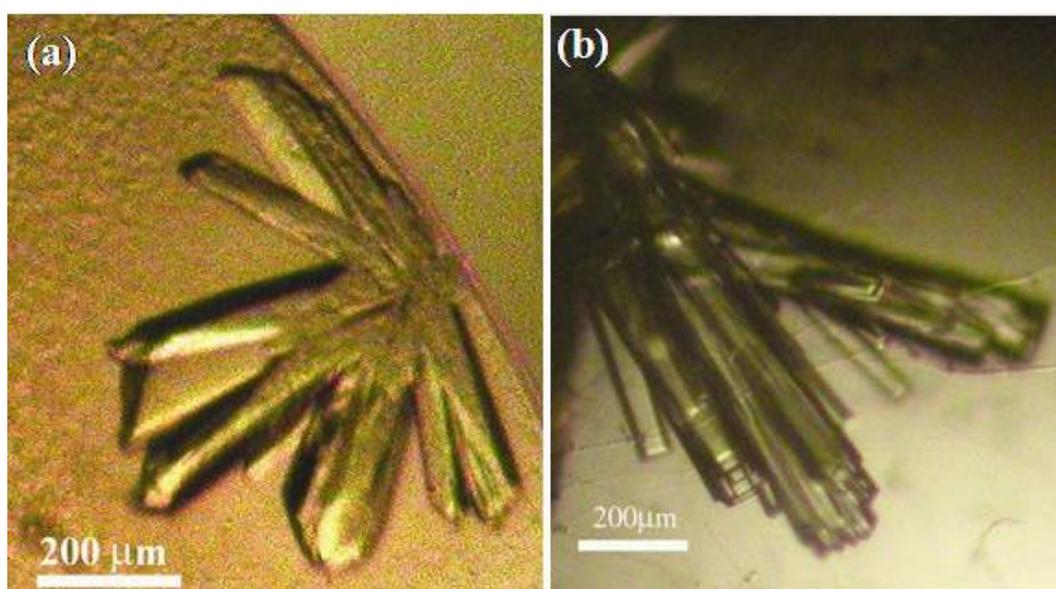


Figure 3.5: VcRK crystals in (a) apo form, (b) ribose +ADP bound form.

3.5.2 Crystallization of VcFK

Crystals were obtained for unliganded protein with A3 (5% PEG 6000, 0.1M MES pH 6.0) and A4 (5% PEG6000, 0.1M HEPES pH 7.0) of Grid Screen PEG 6000 as precipitant with 35%(w/v) PEG 6000 as reservoir after 15 d at 293 K. For co-crystallization of VcFK with ADP and fructose, VcFK was mixed with 5 mM fructose, 5 mM ADP and 5 mM CaCl₂, the mixture was incubated at room temperature for 1 hr. Co-crystals of VcFK were obtained at 293 K using conditions A3 (5% PEG 6000, 0.1M MES pH 6.0) and A4 (5% PEG 6000, 0.1M HEPES pH 7.0) of Grid Screen PEG 6000 as precipitant with 30%(w/v) PEG 6000 as reservoir after 3 days. The BeF₃ bound crystal was obtained by soaking the ADP-fructose bound crystal with 5 mM BeF₃ before data collection. The crystals of VcFK with and without different ligand bound form have been shown in **figure 3.6**. The crystallization conditions for two proteins in different ligand bound forms are shown in **Table 3.1**.

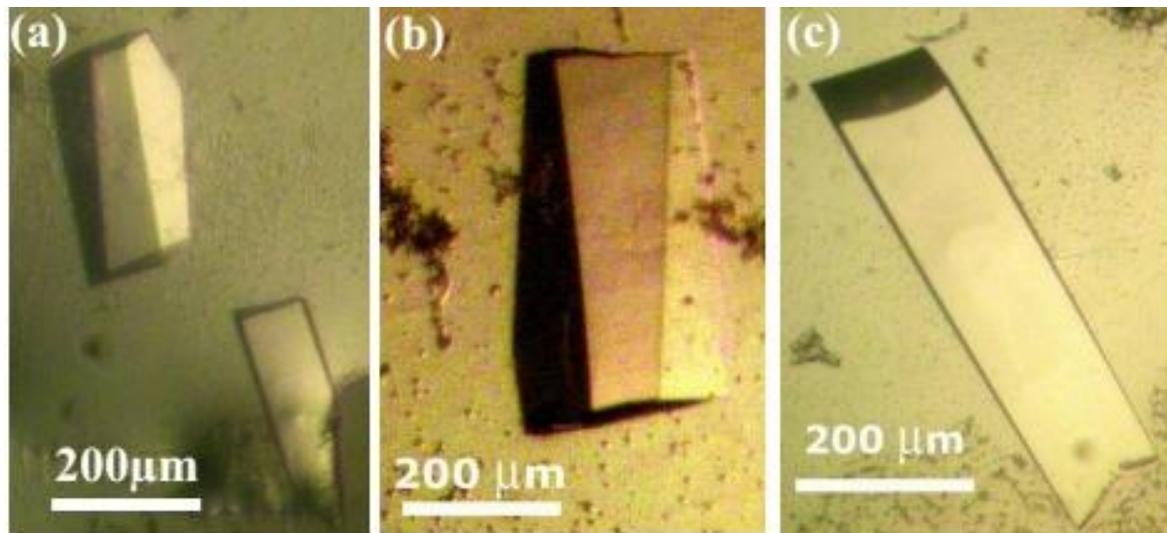


Figure 3.6: VcFK crystals in (a) apo form, (b) sugar bound form, (c) sugar +ADP +Ca²⁺ bound form.

Table 3.1: Detailed crystallization conditions of VcRK and VcFK crystals in apo and ligand bound form.

Protein name	Composition of precipitant	Composition of reservoir solution	Incubation temperature
VcRK (apo) + 2 mM β -mercaptoethanol *	PEG grid screen A3 (5% PEG 6000, 0.1M MES, pH 6.0)	30 % (w/v) PEG 6000 + 200 mM NaCl + 100 mM Tris pH 8.0	293 K
VcRK (complex with 5 mM ADP + 5 mM Ribose) + 2 mM β -mercaptoethanol *	PEG grid screen A3 (5% PEG 6000, 0.1M MES, pH 6.0)	35%(w/v) PEG 6000 + 200 mM NaCl + 100 mM Tris pH 8.0	277 K
VcRK (complex with 5 mM ADP + 5 mM Ribose + 2 mM β -mercaptoethanol *) + 50 mM CsCl (soaked)	PEG grid screen A3 (5% PEG 6000, 0.1M MES, pH 6.0)	35%(w/v) PEG 6000 + 200 mM NaCl + 100 mM Tris pH 8.0	277 K
VcFK (apo)	PEG grid screen A3 (5% PEG 6000, 0.1M MES, pH 6.0), PEG grid screen A4 (5% PEG 6000, 0.1M HEPES pH 7.0)	35%(w/v) PEG 6000 + 200 mM NaCl + 100 mM Tris pH 8.0	293 K
VcFK (complex with 5 mM Fructose)	PEG grid screen A3 (5% PEG 6000, 0.1M MES, pH 6.0)	30%(w/v) PEG 6000 + 200 mM NaCl + 100 mM Tris pH 8.0	293 K
VcFK (complex with 5 mM ADP + 5 mM Fructose)	PEG grid screen A3 (5% PEG 6000, 0.1M MES, pH 6.0)	30%(w/v) PEG 6000 + 200 mM NaCl + 100 mM Tris pH 8.0	293 K
VcFK (complex with 5 mM ADP + 5 mM Fructose) + 5 mM BeF ₃ (soaked)	PEG grid screen A3 (5% PEG 6000, 0.1M MES, pH 6.0)	30%(w/v) PEG 6000 + 200 mM NaCl + 100 mM Tris pH 8.0	293 K

* β -mercaptoethanol was added before crystallization to prevent cysteine mediated dimerization, as it was indicated by non-reducing SDS PAGE.

3.6 Data collection of apo and ligand bound VcRK and VcFK

All crystals of apo enzyme and complex with respective sugar and ADP were looped out from the crystallization drops using a 20 mm nylon loop and flash-cooled in a stream of nitrogen (Oxford Cryosystems) at 100 K.

3.6.1 Data collection of VcRK

Diffraction data were collected on a MAR 225 CCD on BM14 at the European Synchrotron Radiation Facility (ESRF), Grenoble with the incident X-ray beam wavelength of 0.97 Å. 360 frames were recorded with a crystal-to-detector distance of 270 mm as non-overlapping 1° oscillations with 12 s exposure per image. Typical images for diffraction pattern for apo and ribose-ADP bound crystals are shown in **figure 3.7**. The parameters used during data collection and the processing statistics for are given in **Table 3.2**.

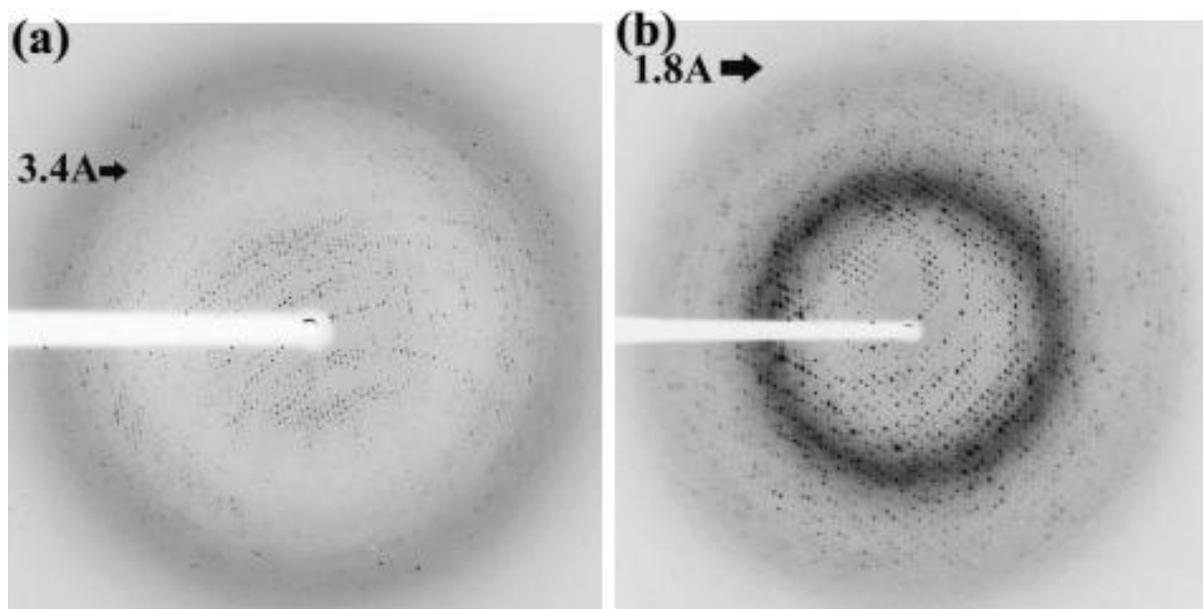


Figure 3.7: X-ray diffraction image of VcRK crystals in (a) apo form diffracted to a resolution of 3.4 Å (black arrow), (b) ADP and ribose co-crystals diffracted to a resolution of 1.75 Å (black arrow) diffraction image of unliganded VcRK.

Table 3.2 Data-collection and processing parameters of VcRK crystals.*(Values in parentheses are for the outermost resolution shell)*

	VcRK in native state	VcRK + ribose + ADP + Na⁺	VcRK + ribose + ADP + Cs⁺
Composition of cryo-protectant	30% glycerol with reservoir solution	30% glycerol with reservoir solution	30% glycerol with reservoir solution
Space group	P1	P1	P1
Unit-cell parameters (Å, °)	a=129.22, b=130.85, c=145.690, α=110.52, β=90.00, γ=119.59	a=59.42, b=70.70, c=79.91, α=106.31, β=97.66, γ=98.69	a=59.30, b=69.55, c=78.31, α=106.99, β=98.76, γ=98.46
Maximum resolution (Å)	3.40	1.75	2.37
Oscillation range (°)	1	1	1
No. of unique reflections	97045	117357	42241
Multiplicity	2.1 (2.5)	4.0 (3.9)	2.07 (2.11)
Mosaicity (°)	1.51	0.78	1.11
Completeness (%)	92.9 (85.6)	95.1 (95.1)	93.8 (93.4)
R_{merge}[†] (%)	15.1 (45.9)	9.3 (56.7)	6.4 (28.37)
Average I/σ(I)	4.7 (2.0)	8.5 (2.4)	5.1 (1.5)

3.6.2 Data collection of VcFK

For the crystals of apo protein, fructose bound and fructose-ADP bound co-crystal, diffraction data set were collected using an in-house MAR Research image-plate detector of diameter 345 mm and Cu K_α radiation (1.54 Å) generated by a Bruker-Nonius FR591 rotating-anode generator equipped with Osmic MaxFlux confocal optics and operated at 50 kV and 50 mA. Data were processed and scaled using iMOSFLM (Battye et al., 2011). Diffraction data for native VcFK crystals were collected to a resolution of 2.45 Å whereas the diffraction data for the Co-crystals of VcFK bound with ADP and Fructose were collected to

a resolution of 1.75 Å (**figure 3.8**). For native VcFK crystal 95.1% complete data set was obtained from 70 frames with 1° crystal oscillation. For VcFK co-crystal, 99.0% complete data set was obtained from 83 frames with 1° crystal oscillation.

For fructose-ADP-BeF₃ bound co-crystal, data were collected at a MAR 225 CCD on BM14 at the European Synchrotron Radiation Facility (ESRF), Grenoble. The incident X-ray beam had a wavelength of 0.97 Å. Data-collection statistics of VcFK are given in **Table 3.3**.

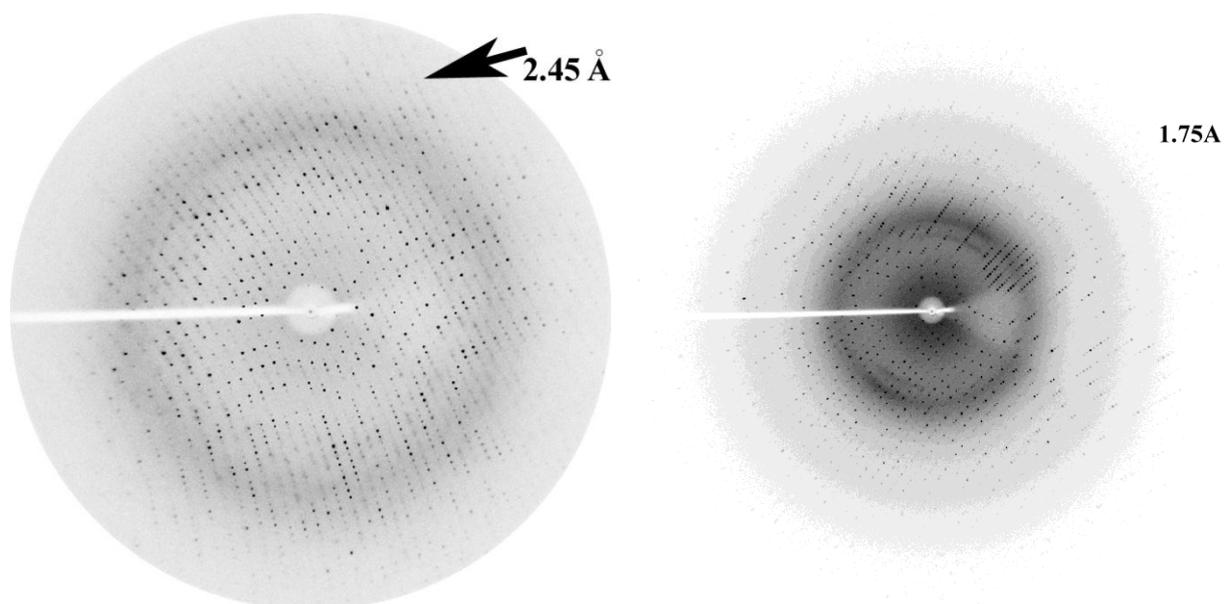


Figure 3.8: X-ray diffraction image of VcFK crystals in (a) apo form diffracted to a resolution of 2.45Å (black arrow), (b) ADP and fructose co-crystals diffracted to a resolution of 1.75 Å.

Table 3.3 Data-collection and data-processing parameters for VcFK crystals*(Values in parentheses are for the outermost resolution shell)*

	Unliganded VcFK	VcFK+ Fructose	VcFK+ADP +Fructose	VcFK+ADP+ Fructose+BeF₃
Space group	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2
Unit cell parameter(Å)	a=107.12, b=99.86, c=61.55;	a=117.24, b=64.67, c=41.89	a=107.005, b=64.291, c=41.758;	a=64.30, b=107.22, c=41.62
Max Resolution (Å)	2.45 (2.5- 2.45)	2.4 (2.44-2.4)	1.73 (1.76-1.73)	1.29 (1.31-1.29)
Oscillation range (°)	1°	1°	1°	1°
No. of unique reflections	23435	15710	30437	71986
Redundancy	2.73	3.3	3.15	3.9
Mosaicity	0.38	0.52	0.40	0.36
Completeness (%)	92.5 (94.3)	94.0 (96.6)	95.1 (92.2)	97.4 (81.1)
R _{merge} †	4.59 (21.4)	10.5 (20.2)	5.46 (26.38)	4.4 (46)
Average I/σ(I)	7.0 (1.8)	7.1(1.7)	6.7 (1.6)	31.5(1.7)

3.7 Initial structure determination of VcRK and VcFK by molecular replacement

A BLAST [22] search for a homologous structure for each protein showed that ribokinase from *E. coli* (PDB code 1RK2, [60]) has the highest sequence identity (54%) with VcRK. Similarly, Aminoimidazole riboside kinase from *S. Enterica* (PDB code 1TZ6, [61]) has maximum sequence similarity with VcFK. It is to be noted that fructokinase from *H. orenii* [62] has only 27% sequence identity with VcFK and therefore the coordinates of 1TZ6 [61] has been chosen as a better starting model for molecular replacement. A model was prepared for each case using the coordinates of maximum homologous structure where the mismatched residues were truncated to Ala. This model was subsequently used as a search model during maximum likelihood molecular-replacement calculations using *Phaser-MR* from CCP4 packages [46, 63]. The detailed result of *Phaser-MR* [47] of VcRK and VcFK are given in **Table 3.4** and **Table 3.5** respectively.

Table 3.4 Selection of initial model for molecular replacement of VcRK

	Apo VcRK	VcRK + ribose + ADP + Na⁺	VcRK + ribose + ADP + Cs⁺
Initial model used for MOLREP	1RK2 (E. coli Ribokinase)	Previously solved structure of apo VcRK	Previously solved structure of apo VcRK
Highest sequence identity with model (%)	54	100	100
Resolution (Å) range used for molecular replacement	43.57-4.0	27.47-3.5	26.64-2.5
LLG value produced in MOLREP	697	4367	4317
R-factor produced in MOLREP (%)	54.3	49.8	46.1
TFZ produced in MOLREP	12.0	35.9	35.8
Molecules / a. u.	16	4	4
Matthews coefficient $V_M (\text{Å}^3 \text{Da}^{-1})$	3.9	2.4	2.3
Solvent content (%)	68	50	48

Table 3.5: Selection of initial model for molecular replacement of VcFK

	VcFK apo	VcFK+FRU	VcFK+ADP+FRU+Ca²⁺+Na⁺	VcFK+ADP+FRU+Ca²⁺+Na⁺+BeF₃⁻
Initial model used for MOLREP	1TZ6 (AIRS kinase from <i>S. enterica</i>)	Previously solved structure of apo VcFK	Previously solved structure of apo VcFK	Previously solved structure of apo VcFK
Highest sequence identity with model (%)	100	100	42	100
Resolution (Å) range used for molecular replacement	40-4.0	40-3.0	40-2.5	40-2.5
LLG value produced in MOLREP	531.3	1540.6	1934.4	2053.3
R-factor produced in MOLREP (%)	58.5	48.2	49.7	53.8
TFZ produced in MOLREP	16.7	20.7	19.9	22.2
Molecules per ASU	2	1	1	1
Matthew's coefficient $V_m (\text{Å}^3 \text{Da}^{-1})$	2.4	2.3	2.1	2.1
Solvent content (%)	48	46	40	40

3.8 Structure refinement and interpretation of final models of VcRK and VcFK.

The structures of both proteins were refined with several cycles of model building in COOT [52] and refinement with Phenix [48]. TLS refinement was performed during the final stages of refinement [51]. Average B-factors for each residue were calculated using B average in CCP4 [64]. The refinement and structure determination statistics of VcRK and VcFK are listed in Tables 3.6 and 3.7 respectively.

Table 3.6: Refinement statistics and description of final models of VcRK

	Apo VcRK	VcRK + ribose + ADP + Na⁺	VcRK + ribose + ADP + Cs⁺
Resolution range (Å)	43.57-3.40	27.47-1.75	26.64-2.37
R-factor/R free (%)	23.86/30.60	18.46/22.20	21.13/25.48
No. of protein residues per molecule	306	306	306
ADP/ d-ribose/Na ⁺ /Cs ⁺	0/0/0/0	1/1/1/0	1/1/0/1
Solvent molecules	0	1169	248
Average B-factor/r.m.s.d. (Å²)			
Main chain (306 residues)	81.2	26.1	45.5
Side chain	82.6	33.1	48.4
Solvent (molecules)	NA	38.5	44.2
r.m.s.d. from standard geometries			
Bond length (Å)	0.012	0.011	0.007
Bond angles (°)	1.840	1.278	1.173
Ramachandran plot (%)			
Most favored	81.31	97.62	96.00
Allowed	12.73	1.80	3.59
Outliers	5.96	0.57	0.41
PDB ID	4X8F	4XDA	4XCK

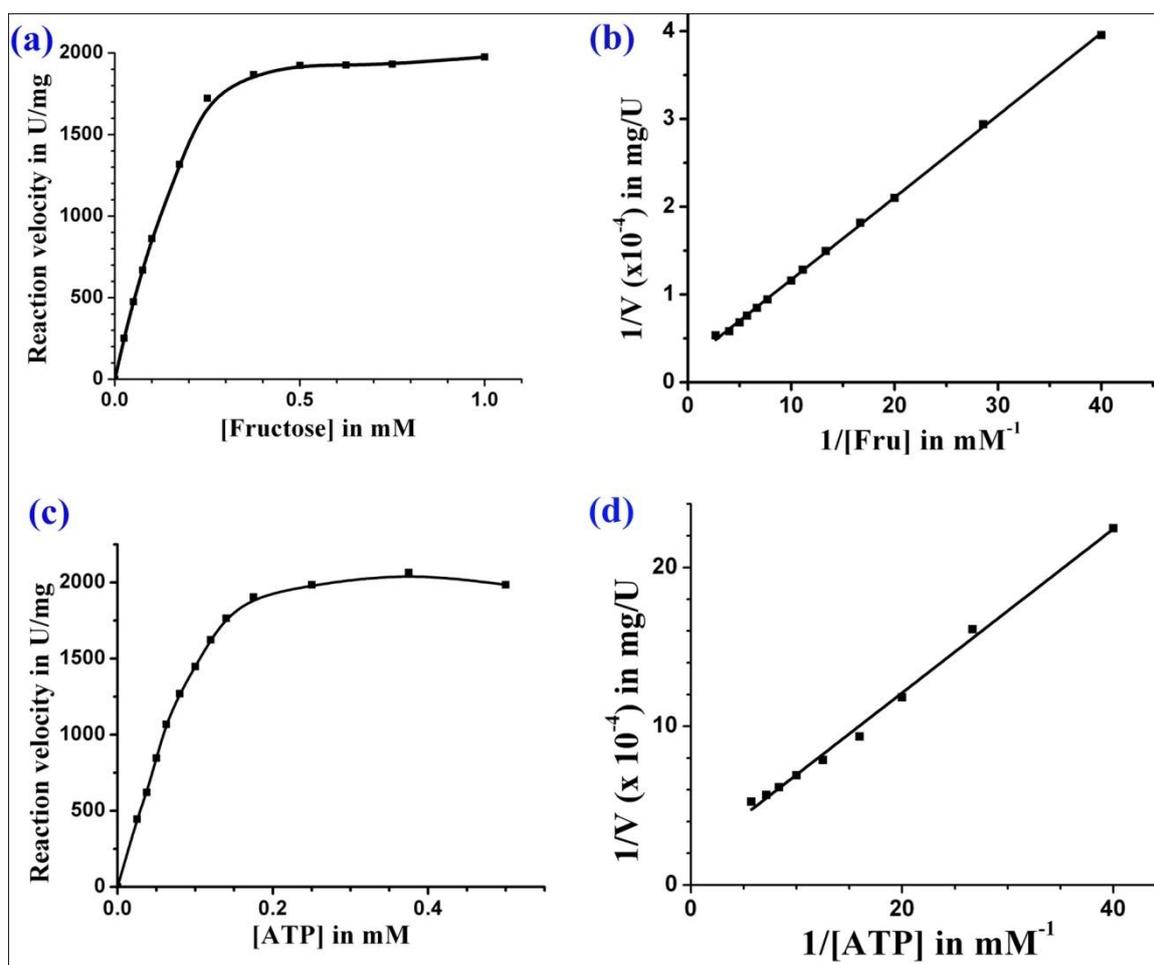
Table 3.7: Refinement statistics and description of final models of VcFK

	VcFK apo	VcFK+FRU	VcFK+ADP+ FRU+Ca ²⁺ +Na ⁺	VcFK+ADP+FRU +Ca ²⁺ +Na ⁺ +BeF ₃ ⁻
Refinement resolution range (Å)	29.41-2.46	25.66-2.30	35.00-1.75	32.15-1.29
R _{work} (%)	21.7	21.1	14.5	13.5
R _{free} (%)	27.2	26.3	17.9	15.8
No. of protein residues	323	323	323	323
ADP/Ca ²⁺ /Na ⁺ /d-fructose/BeF ₃	0/0/1/0/0	0/0/1/1/0	1/1/1/1/0	1/1/1/1/1
Solvent molecules	NA	210	356	379
r.m.s.d. from standard geometries				
Bond lengths (Å)	0.004	0.003	0.010	0.02
Bond angles (deg)	0.968	0.806	1.377	1.46
B-factors (Å²)				
Protein main chain	65.4	31.1	12.3	12.4
Protein side chain	66.1	33.3	17.8	17.6
Ca ²⁺	-	-	9.0	12
ADP/FRU	-	-/27.2	53.7/5.7	23.6/7.3
Ramachandran plot (%)				
Favored	90.7	96.1	98.7	98.7
Generously allowed	6.1	2.9	1.3	1.3
Disallowed	3.2	1.0	0.0	0.0
PDB ID	5EY7	5F11	5F0Z	5EYN

3.9 Results of enzymatic activity of VcFK

The kinetic parameters of VcFK were determined at temperature 20°C and pH 7.5. Under standard reaction conditions [58] (see Materials and Methods, **section 2.4**). The K_m values of sugar, ATP and cations were determined with the help of Michaelis Menten curve and a corresponding Lineweaver-Burk plot. VcFK catalyzes the ATP-dependent phosphorylation of d-fructose to fructose-6-phosphate with apparent K_m and V_{max} values of 0.37 mM fructose and 1975 U/mg of protein, respectively (**figure 3.9a and b**). The K_m for ATP was determined to be 0.279 mM (**figure 3.9 c and d**). The catalytic activity of VcFK is

highly dependent on divalent and monovalent cations. Both magnesium and calcium can participate in phosphotransfer with K_m values of 0.027 mM and 0.17 mM respectively (**figure 3.9e, f, g and h**). Monovalent cation potassium activates the enzyme with $K_m= 0.173$ mM (**figure 3.9i and j**). Cesium can also be used as an activator with $K_m= 9.54$ mM. No activity was observed for sodium upto 300 mM (**figure 3.9k and l**). The mutant protein E110Q (discussed in **section 5.3**) showed nominal activity towards fructose-ATP.



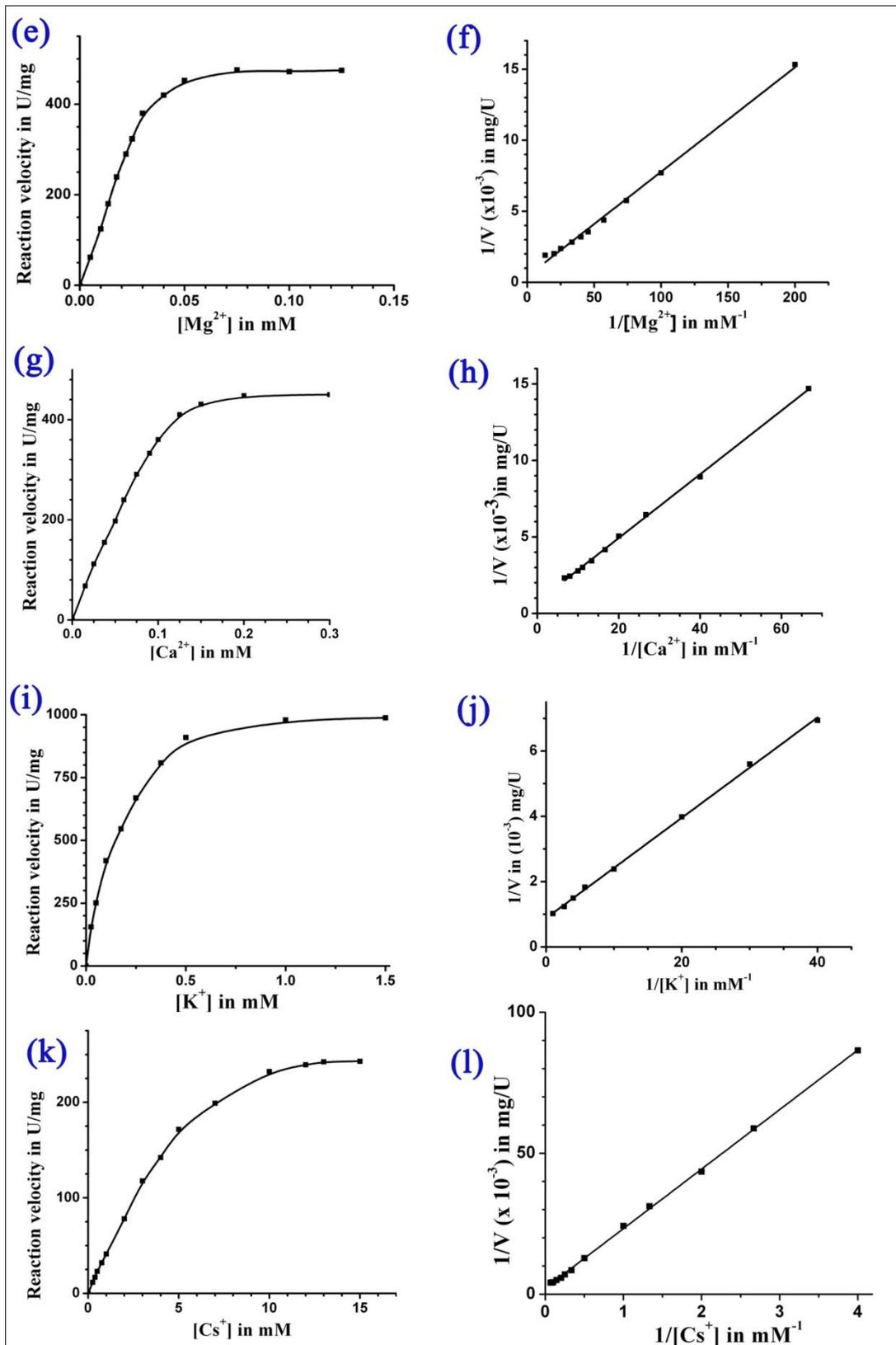


Figure 3.9: Activity assay of VcFK with respect to fructose (a and b), ATP (c and d), magnesium (e and f), calcium (g and h), potassium (i and j) and cesium (k and l). In each case, first curve shows the Michaelis-Menten curve and the second one illustrates Lineweaver-Burk double reciprocal plot.

CHAPTER- 4

ANALYSIS OF THE CRYSTAL

STRUCTURE

OF TWO SUGAR KINASES:

VcRK AND VcFK

Crystal structures of VcRK and VcFK, in apo form and in complex with different substrates/substrate analog have been analyzed thoroughly and their results are reported here. Structural knowledge of the chemical environment of the substrate binding sites facilitate to decipher the mechanism of function of the candidate enzyme against their substrates.

4.1 Structure of Ribokinase from *Vibrio cholerae* (VcRK)

Crystal structures of VcRK were determined in three different forms: unliganded or apo VcRK, VcRK in complex with ribose and ADP, VcRK in complex with ribose, ADP and Cs⁺ ion. All crystals belong to the triclinic space group P1, characterized by the complete absence of any rotation axes, rotation-inversion axes, screw axes or planes.

4.1.1 Overall monomeric structure of VcRK

The monomeric structure of VcRK contains mainly two domains: a large, globular in shape, central $\alpha\beta\alpha$ domain and a small β -sheet region that distinctly protrudes from the former (**figure 4.1.1**). The core of the central $\alpha\beta\alpha$ domain consists of a nine stranded twisted β -sheet region surrounded by eleven α helices flanked on both sides of the central β -sheet region that dominate the surface of the molecule. The pocket formed between central $\alpha\beta\alpha$ domain and the protruding β -sheet region harbors the catalytic site. The first six β -strands of this nine-stranded sheet together with the associated α helices (α 1- β 4- α 2- β 5, α 4- β 8- α 5- β 9- α 6- β 10) have the topology of a typical Rossmann fold. The ribokinase fold is extended by three additional β strands (β 11- β 13) that are connected by short reverse turns. Two helices, α 7 and α 8 shield one face of this extension and α 9 and α 10 screen the other. β 13 is connected to α 9 by a very long loop formed by residues 237-250, indicated by green color in **figure 4.1.1**. Another long loop (shown in burgundy color) connects helix α 10 and C-terminal helix α 11. The protruding β -sheet domain is formed by two insertions into the central $\alpha\beta\alpha$ fold, each of which forms two pairs of antiparallel β -strands connected by loops: β 2- β 3 (residue 13-36)

and $\beta 6$ - $\beta 7$ (residue 95-107). The first pair connects $\beta 1$ and $\alpha 1$, and the other is located between $\beta 5$ and $\alpha 3$. A long loop (residues 18-34, termed as long lid loop, colored blue) joins $\beta 2$ and $\beta 3$, which bends almost perpendicularly at Pro19 and Pro24. $\beta 6$ and $\beta 7$ are connected by a short turn (residues 100-105, colored blue). This overhanging domain covers the active site like a lid, therefore it is termed as lid or flap domain and the two loops joining $\beta 2$ - $\beta 3$ and $\beta 6$ - $\beta 7$ are termed as ‘long and short lid loops’ respectively. The loops shown in different colors in **figure 4.1.1** have diverse significance, e.g., the blue colored loops clutch with adjacent monomer and help to dimerize whereas the loop in green participates in ATP binding and burgundy colored loop embraces the monovalent cation.

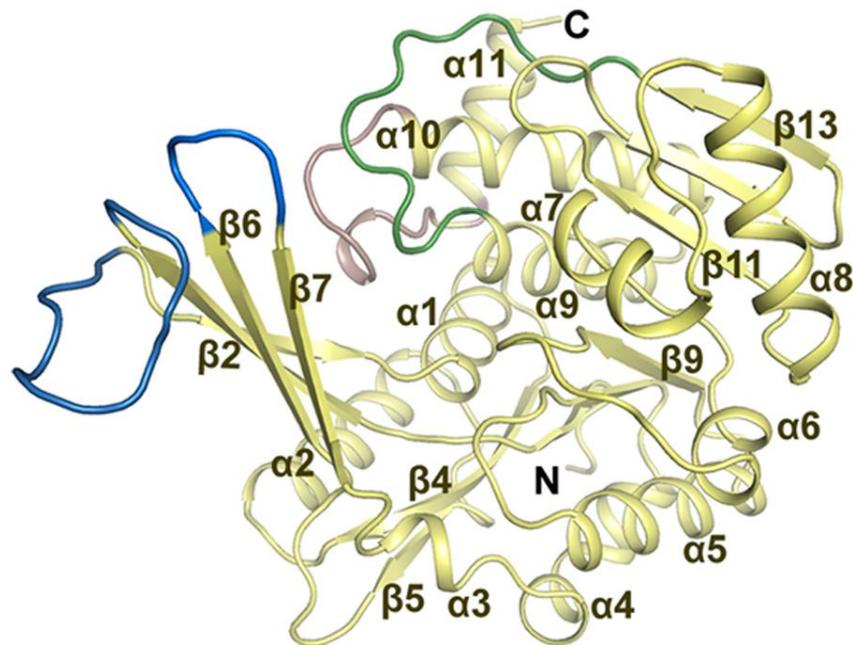


Figure 4.1.1: Cartoon representation of VcRK monomer illustrates nine-stranded central β -sheet domain with α -helices above and below. The 4-stranded β -sheet region is overhanging towards left side from the central domain. The two loops (between $\beta 2$ - $\beta 3$ and $\beta 6$ - $\beta 7$) colored ‘blue’ participate in dimerization, the green loop (between $\beta 13$ - $\alpha 9$) helps ATP binding and the burgundy colored loop (between $\alpha 10$ - $\alpha 11$) involves in monovalent cation binding.

4.1.2. Packing of molecules inside the asymmetric unit of VcRK apo and VcRK complex crystals

Apo VcRK crystallized with a large unit cell ($a=129.22$, $b=130.85$, $c=145.690$) compared to that of the enzyme-substrate bound crystals ($a=59.42$, $b=70.70$, $c=79.91$) (**table 3.2**). Matthew's coefficient calculation (program *MATTHEWS_COEF*, CCP4 program suite, [65]) predicted the presence of 26 copies of VcRK molecules in the asymmetric unit as the most value, corresponding to a solvent content of 48% and V_m of 2.3. However, molecular replacement calculation (program *PHASER* at PHENIX) placed only sixteen molecules corresponding to a solvent content of 68% and V_m of 3.9 (**table 3.4**). Analysis of the packing of the sixteen molecules inside the crystal (**figure 4.1.2a**), in the form of eight dimers (A : B, C : D, E : F, G : H, I : J, K : L, M : N, O : P), exhibit a spiral arrangement leaving a prominent cavity at its centre, accounting the high solvent content (**figure 4.1.2a**). One such spiral arrangement is depicted in **figure 4.1.2b** and its 90° rotated view in **figure 4.1.2c** depicting the double stranded helical arrangement.

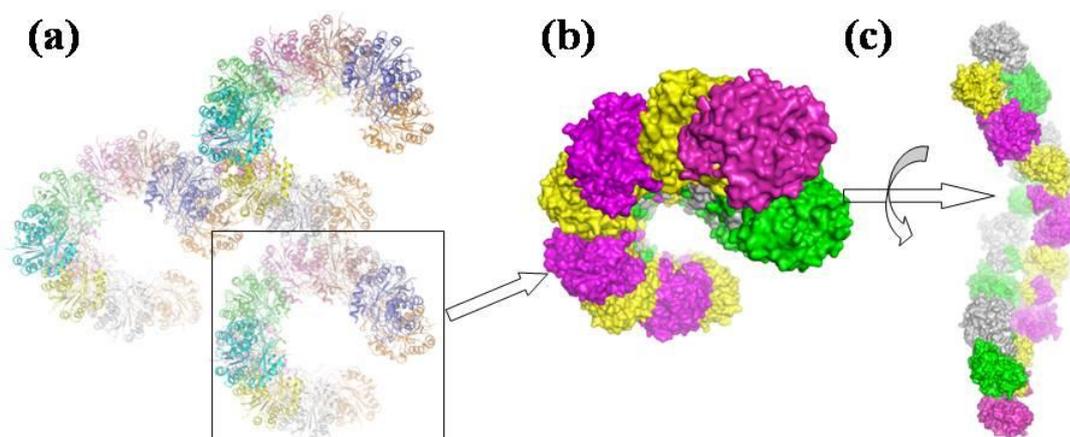


Figure 4.1.2: Spiral arrangement of apo VcRK molecules in the crystal. (a) In cartoon where the spirals encompassing large cavity is evident. (b) Surface representation elucidating eight dimers forming a double-stranded helical pattern; ‘yellow and magenta’ dimers constitute the upper strand whereas ‘grey and green’ dimers the lower, (c) 90° rotated view to make the spiral more evident.

Each asymmetric unit of VcRK complex crystal contain two dimers (**figure 4.1.3a**). These two dimers (chains A:B and C:D) are related by twofold non-crystallographic symmetry. Neighboring asymmetric units are packed one beside another to form a plane layer (**figure 4.1.3b**) leaving little empty space. The crystal structure consists of such planar layers one above another. This explains the smaller solvent content (~40%) of the structure.

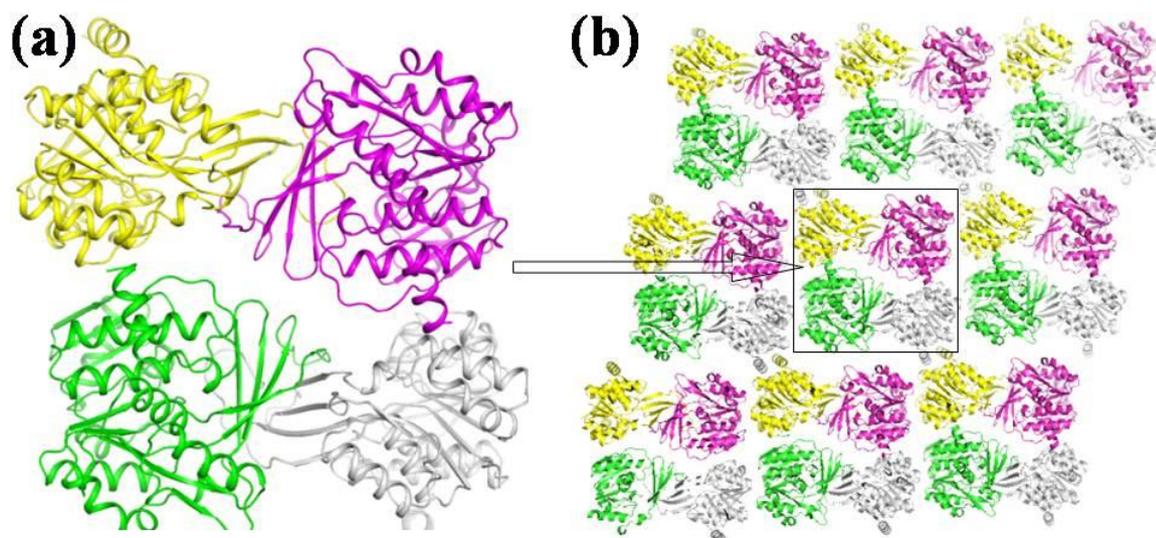


Figure 4.1.3: Crystallographic arrangement of sugar-ADP bound VcRK. (a) Arrangement of two dimers in each asymmetric unit, (b) arrangements of asymmetric units illustrate the packing of crystal.

4.1.3 Domains involved in dimerization and formation of β -clasp

VcRK, in both apo and enzyme-sugar-ADP complexed form, exist predominantly as dimeric, as predicted by PDBePISA [53]. The residues involved in dimerization were identified by PIC server [66] and the program *CONTACT* (CCP4 suite [46]) was used to calculate interatomic distances. Dimers are primarily formed through interaction between the protruding β -sheet regions of the constituent monomers. Four β -strands from each neighboring monomers approach close to each other and twist to form the dimer. These β -turns pack around a twofold symmetry axis, forming a flattened β -barrel type structure, called ‘hand-shake motif’ or ‘ β -clasp’ (**figure 4.1.4**). Longer β -turn from each monomer bends

orthogonally near Pro19 towards the $\beta 7$ of other monomer to close the clasp. Eight β -strands are arranged as $\beta 3A$ - $\beta 2A$ - $\beta 6A$ - $\beta 7A$ - $\beta 3B$ - $\beta 2B$ - $\beta 6B$ - $\beta 7B$ to form the clasp (fig 4.1.4). The core of the clasp contains hydrophobic residues like Leu16, Ile94, Phe21, Pro22, and Met96.

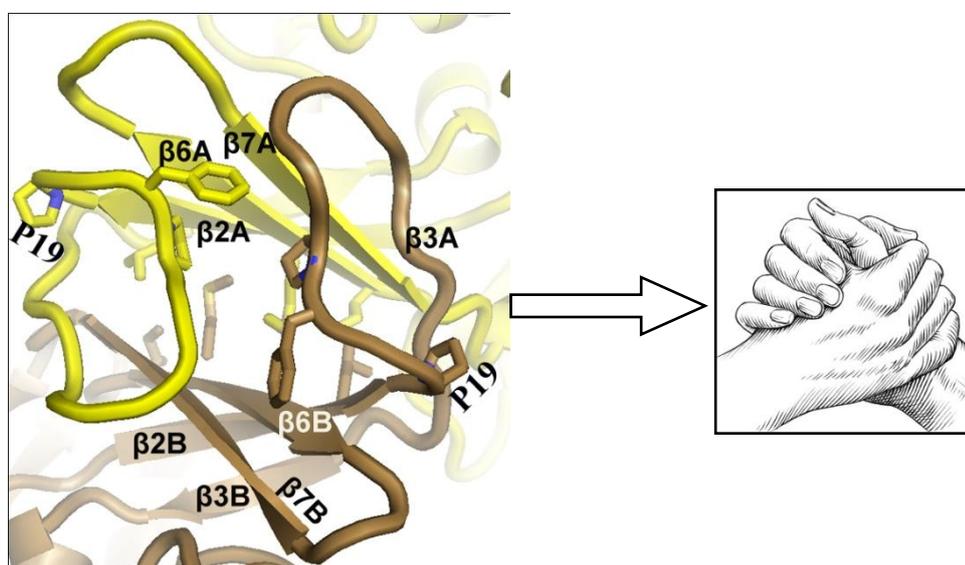


Figure 4.1.4: The ‘hand-shake motif’ or ‘ β -clasp’ that forms the dimer of VcRK. Chain A and chain B are shown in ‘sand’ and ‘yellow’, hydrophobic residues and Pro19 in sticks.

4.1.4 Difference in dimeric structure of apo and substrate-ADP bound structure

4.1.4.1 The Open homodimeric form of apo VcRK

Apo VcRK displays a symmetric homodimeric structure with dimer axis along the core of the clasp. Hydrophobic interactions, between Leu16, Phe21, Pro22, Pro24, Tyr33, Pro91, Ile94, and Met96 from each subunit, stabilize the dimeric form. The dimeric form is additionally stabilized by main-chain main-chain hydrogen bonds between Gly30, Ile107, Ile109 and Leu28 from each monomer and H-bonding interaction through side chains of His14, Leu28, Gly30, Tyr33, Gln98, Cys108, Ser110 and Glu112. According to PDBePiSA program [53], the average buried surface area per dimer is approximately 9 % of the total surface area (buried surface = 2350 \AA^2 whereas total surface = 25500 \AA^2) with ΔG_{int} value of $\sim -22 \text{ kCal/mol}$ and ΔG_{diss} value of $\sim 16 \text{ kCal/mol}$ implying that the dimer is energetically

stable. In apo form, the central $\alpha\beta\alpha$ domains remain relatively away from the clasp. Separation between the central domain of constituent monomers are about $\sim 54 \text{ \AA}$ implying a more open cavity. This structure is called an ‘open dimer’ (**figure 4.1.5a**) which favors easy entrance of sugar inside the catalytic pocket for efficient catalysis.

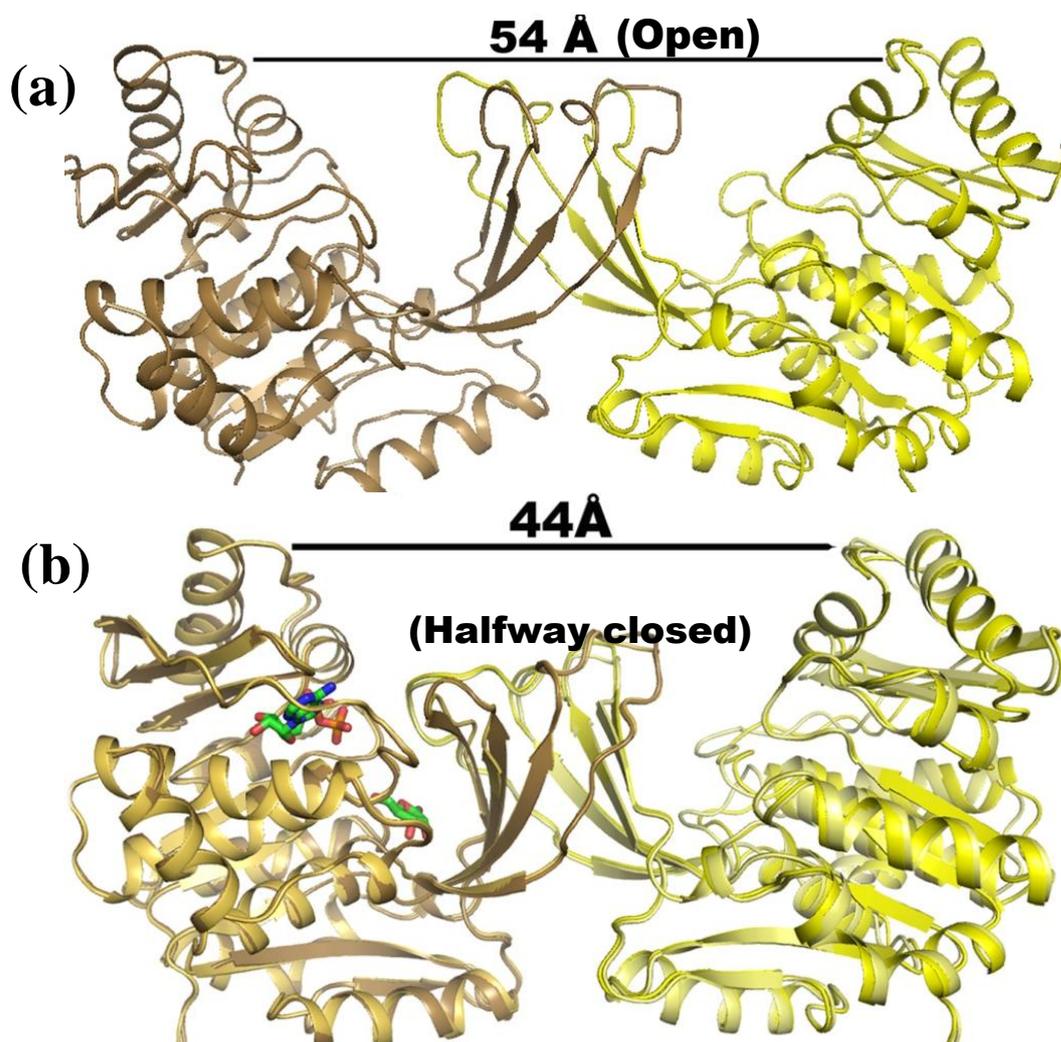


Figure 4.1.5: The (a) apo and (b) substrate bound dimeric form of Vc-RK; symmetric apo dimer about the β -clasp axis and more open ternary complex structure. Distance between the C_{α} atoms of residue Thr195 from each monomer is indicated.

Footnote:

- i) ΔG_{int} indicates the solvation free energy gain upon formation of the assembly, in kcal/M. The value is calculated as difference in total solvation energies of isolated and assembled structures. This value does not include the effect of satisfied hydrogen bonds and salt bridges across the assembly's interfaces.

- ii) ΔG_{diss} indicates the free energy of assembly dissociation, in kcal/M. The free energy of dissociation corresponds to the free energy difference between dissociated and associated states. Positive values of ΔG_{diss} indicate that an external driving force should be applied in order to dissociate the assembly, therefore assemblies with $\Delta G_{\text{diss}} > 0$ are thermodynamically stable.

4.1.4.2 Formation of halfway closed dimer upon ribose-ADP binding

The structure VcRK in complex with ADP+ribose+Na⁺(or Cs⁺) comprise of two dimers (chains A:B and chains C:D) per asymmetric unit. The two monomers of the constituent dimers show significant structural difference (RMSD=0.327). Chains A and C contain both ribose and ATP at their respective active sites, but in chains B and D no bound ribose is seen. Instead, chain B binds an extra ADP at the sugar-binding site whereas the position remains vacant in chain D. Members of the ribokinase family exhibit large scale movements of the flap domain ($\beta 2$ - $\beta 3$ and $\beta 6$ - $\beta 7$) upon trapping of sugar molecule [60]. Chains A and C display such movement and assume closed structure upon sugar binding. For chain B, non-specific ADP molecule hinders in closing the lid whereas chain D, being devoid of sugar, such lid closure is not executed. This difference in lid closure leads to heterogeneous dimer with one closed chain (A or C) and one open chain (B or D) producing ‘halfway closed dimer’ (**figure 4.1.5b**). Average buried surface area per dimer, calculated by the program PDBePISA, is approximately 12 % of the total surface area (buried surface = 5300 Å², total surface = 46000 Å²) with ΔG_{int} value of ~ -39 kCal/mol and $\Delta G_{\text{dissociation}}$ value of ~13 kCal/mol implying that the dimer is energetically stabler than the apo dimers.

4.1.5 Location of the active site in VcRK

The active site is located in the valley between the central domain and the protruding flap region (**figure 4.1.1, figure 4.1.6**). This site contains a distinct sugar binding region and a region that binds an ATP or its analog so that its γ -phosphate directs towards the accepting alcohol group of sugar. A so called ‘anion hole’ exists near the γ phosphate group of ATP. Four flexible loops (**figure 4.1.6**) have distinct roles in the catalytic process. The blue loops

(residues 18-34 and 100-105) are part of β -clasp, which are also essential for sugar binding and dimerization. The green loop (residues 237-250), is termed as large ATP loop generates framework for ATP binding site. It compartmentalizes ATP from the monovalent cation and thereby acts as allosteric effectors. The monovalent cation binds between green and burgundy loops (residues 284-295). From **figure 4.1.6**, it is clear that these loops exhibit structural flexibility in apo VcRK structure whereas the rest portion are relatively rigid.

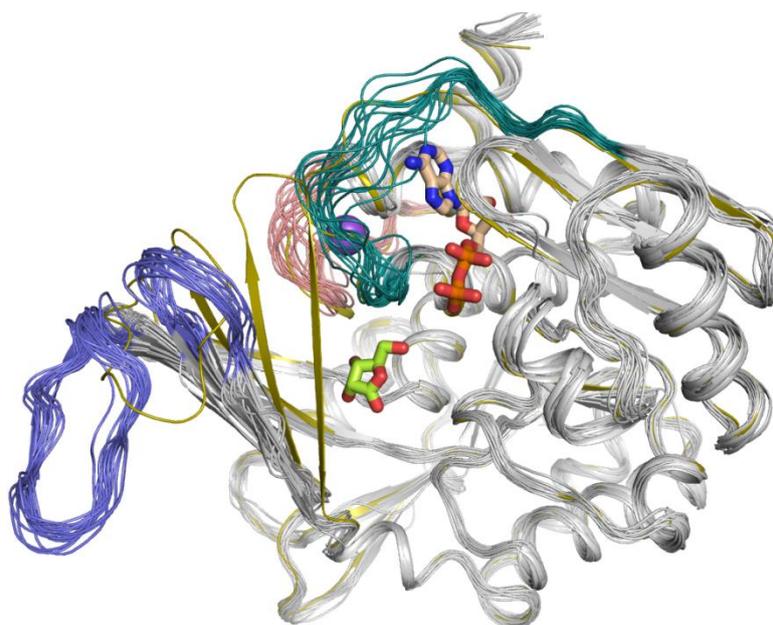


Figure 4.1.6: Structural flexibility of the active site loops in apo VcRK. Superposition of all 16 chains of apo form VcRK (grey) alongwith ribose-ADP-Cs⁺ bound form (gold).

4.1.5.1 The five-membered ring form of ribose is the substrate for VcRK

Ribose remains in both open and closed chain form in solution. Ribose exists in an equilibrium among 5 forms-the linear form $\text{H}-(\text{C}=\text{O})-(\text{CHOH})_4-\text{H}$ and either of the two ring forms: α - or β -ribofuranose ("C3'-endo"), with a five-membered ring, and α - or β -ribofuranose ("C2'-endo"), with a six-membered ring as depicted in **figure 4.1.7a** [67]. The beta-ribofuranose form predominates in aqueous solution. The active site of ribokinase is highly specific towards α -D-ribofuranose, though this form is relatively less abundant (~7%).

In VcRK, the ribose exists in O4'-endo puckering motif, where ribose projects the ring oxygen atom O4' and carbon C5' above the plane made by four carbon atoms of the ring while all the remaining non-ring oxygen atoms remain on the other side of the plane (**figure 4.1.7b**). This form allows the O5' atom of ribose to be available for phosphorylation [67].

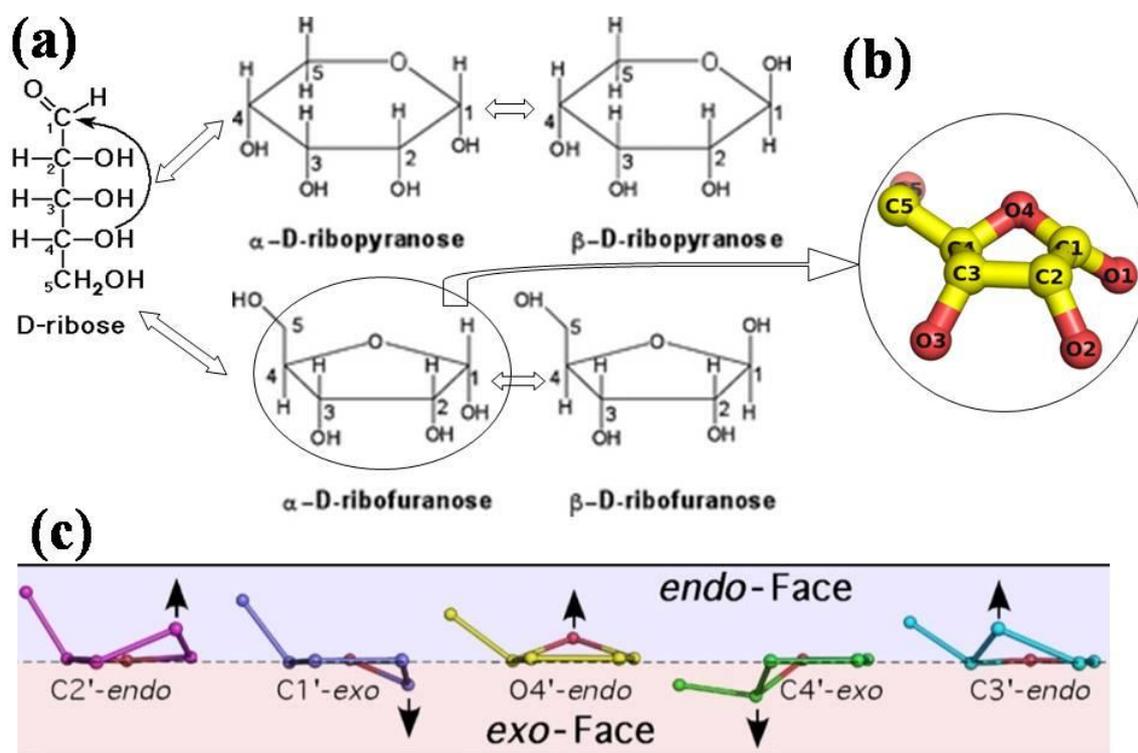


Figure 4.1.7: All five forms of D-ribose in solution. (a) Open chain form and four different closed forms: 6-membered rings are named as ribopyranose and 5-membered rings are termed as ribofuranose. In α -form, all ring oxygens are projected in same direction whereas in β -form, O1' heads opposite to others. (b) VcRK binds ribose in α -D-ribofuranose form with O4' endo pucker configuration, (c) diagram to illustrate different sugar pucker motif for ribose.

4.1.5.2 The ribose-binding pocket

VcRK binds one ribose molecule between the N-terminal end of the helix $\alpha 1$ on one face, and the lid domain on the other (**figure 4.1.8**). A conserved GGK motif (slate color) near the N-terminus of $\alpha 1$ creates the base of ribose binding pocket. The C-terminal ends of large ATP loop (green) and monovalent cation binding loop (burgundy color) together with

four β -strands from the flap partially enclose the sugar molecule. The compartmentalization becomes complete when the large β -loop from the adjacent monomer (grey color) approaches towards sugar during dimerization and totally buries it from solvent molecules.

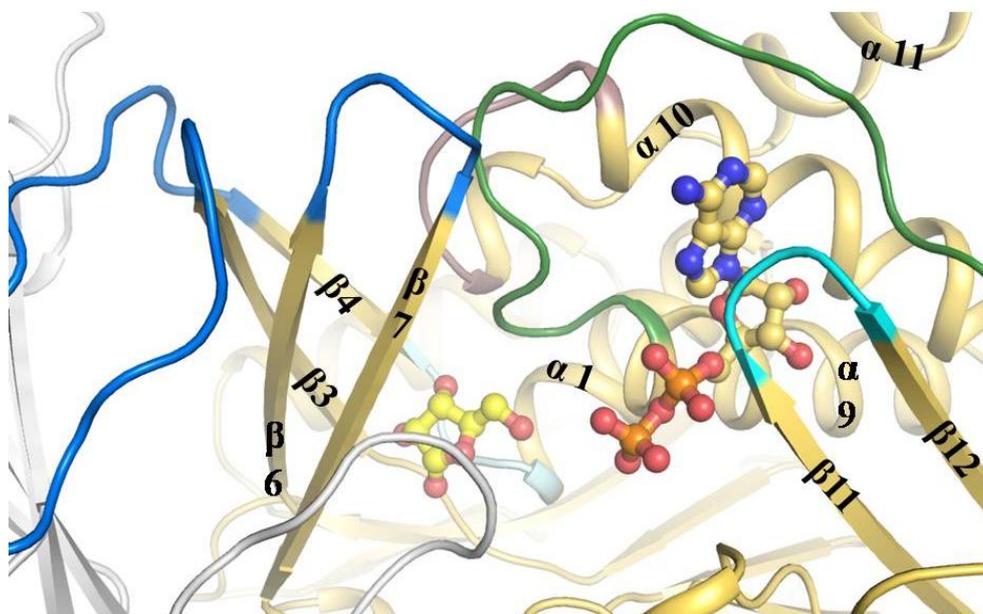


Figure 4.1.8: Location of ribose in VcRK and surrounding region. ADP and sugar are shown in sticks. The GGK motif (in slate) provides the base of ribose, large ATP loop (green) monovalent cation binding loop (burgundy) and short β -lid from self-monomer (β 6- β 7) partially cover the sugar molecule. Large lid loop from adjacent monomer (grey loop) completes the covering of the sugar.

4.1.5.3 Residues involved in ribose binding

Figure 4.1.9a shows the superposition of substrate bound VcRK structure (golden) with apo VcRK structure (grey). The electron density of sugar and its neighboring residues are clear in 2Fo-Fc map (**figure 4.1.9b**). All the non-ring 'O' atoms of the ribose molecule are projected towards the $\alpha\beta\alpha$ domain forming altogether 12 hydrogen bonds with the side chains of Asn11, Asp13, Lys40, Asn43, Glu140 and Asp252 and the main-chain amide atoms of Gly39 and Asn11. The opposite surface of ribose, constituted by four ring carbons and C5', is hydrophobic and faces hydrophobic residues from the bottom of the lid like Ile97, Ile107,

Ile109, Ala95 and Ala288 (**figure 4.1.9c**). This surface does not form any hydrogen bond, Ile107 from $\beta 7$ shifted approximately 4Å to cover the sugar via hydrophobic interaction. Most of the residues involved in sugar binding do not suffer major structural change. The side chain of Lys40 moved away $\sim 3\text{\AA}$ to avoid collision with sugar and created H-bond with O4' of ribose. These sugar-binding polar residues and hydrophobic neighbors are highly conserved among the whole RK superfamily (**figure 4.1.12**). Asp252 acts as the activator of sugar as it makes O5 a better nucleophile in attacking the γ -phosphate of ATP.

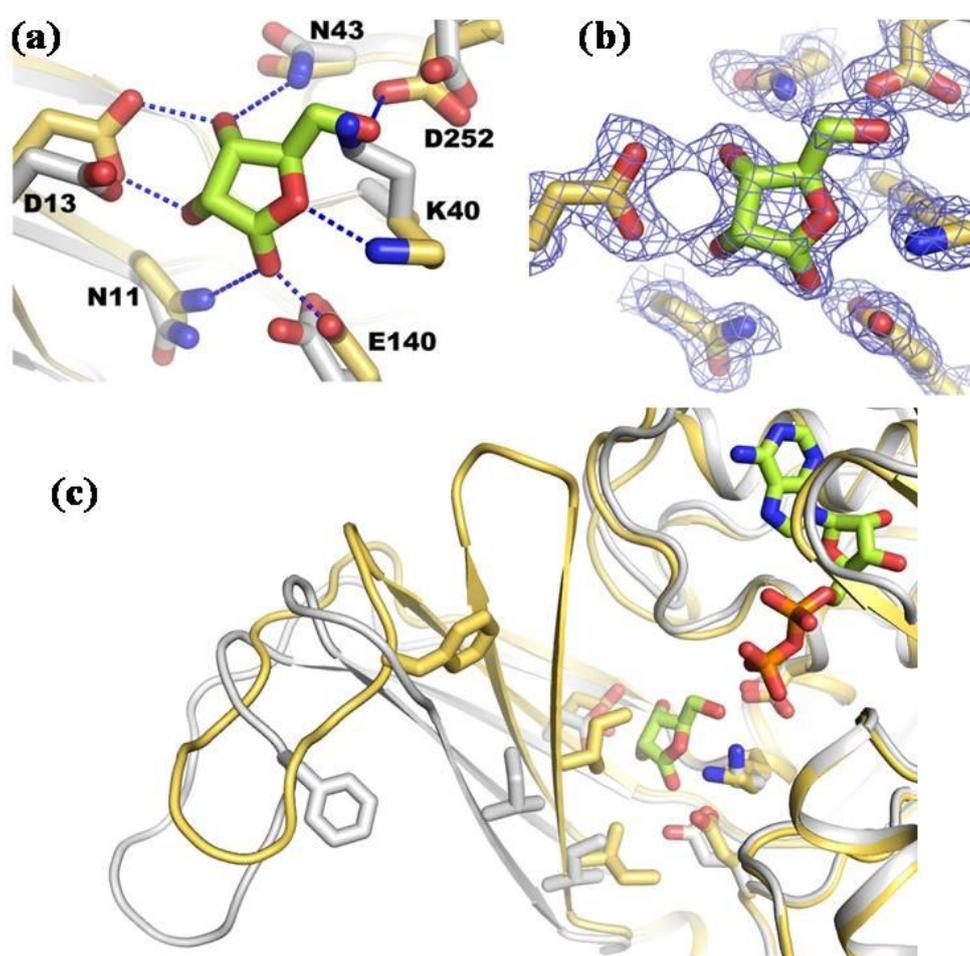


Figure 4.1.9: Details of ribose binding in VcRK. (a) Side chains of VcRK involved in ribose binding are shown in stick representation (complexed VcRK in yellow, apo VcRK in grey). (b) $2F_o - F_c$ electron density map contoured at 1σ around ribose and ribose interacting residues. (c) Lid closure upon ribose binding: hydrophobic and polar residues before (grey) and after (yellow) sugar trapping are shown in stick representation. ADP molecule is shown for clarity.

4.1.5.4 ADP binding site of VcRK

In VcRK, ADP binds in a shallow pocket with its β -phosphate heading towards O5' of ribose. **Figure 4.1.10** shows that the adenosine part is stacked between small ATP loop (cyan), $\alpha 9$ and large ATP loop (green). The phosphate groups face N-terminal of $\alpha 10$ on one side and the other side is exposed to solvent. The adenine group does not involve any bonding interaction, whereas the O2' and O3' of ribose part make hydrogen bonds with His276 and Gly225 (**figure 4.1.10**). The α - and β -phosphate groups involve in H-bonding interactions with Gly222 and side chain of Asn184 respectively. In this configuration, β -phosphorus atom of ADP remains at a distance of 7.7 Å away from O5' of ribose. Gly222 and Gly225 are highly conserved in the Ribokinase superfamily (**figure 4.1.12**).

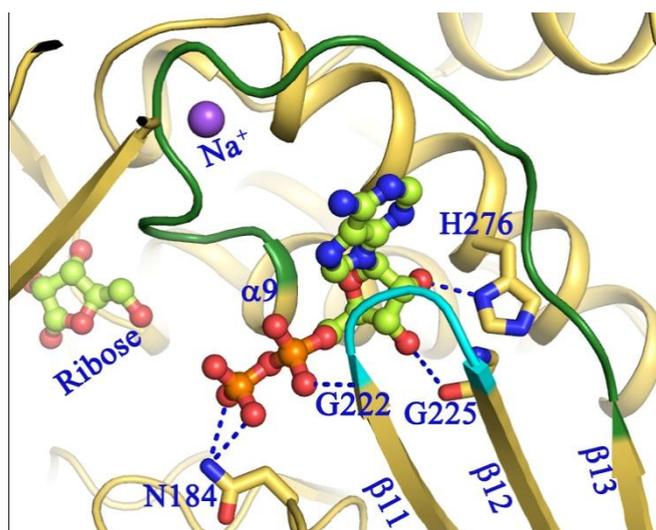


Figure 4.1.10: Zoomed view of ADP binding site. Residues involved in hydrogen bonding with ADP are shown in stick representation. The blue dashes indicate H-bonds.

4.1.6 Divalent and monovalent cation binding

Divalent and monovalent cations play significant role in the catalysis of ribokinase superfamily members. Divalent cations such as Mg^{2+} , Mn^{2+} , Co^{2+} , Ca^{2+} , Ni^{2+} and Cu^{2+} bind between the β and γ -phosphates of ATP to form a metal chelate with the nucleotide, which aids phosphoryl group transfer [68]. Divalent cations also counteract the multiple negative

charges of the nucleotide, and probably help arrange the phosphate groups in a favorable conformation for the reaction. In VcRK, divalent cation could not be detected in spite of addition at sufficient amount during crystallization. This is probably due to the absence of γ -phosphate group in ADP molecule used in crystallization.

All the complexed VcRK structures were observed to be in activated form although sodium chloride was used as the only monovalent ion during crystallization. Therefore Na^+ ion was assigned in the electron density map at the place of monovalent cation site **figure 4.1.11a**. Although electron density of Na^+ is indistinguishable from that of water (both contain 10 electrons), their coordination distance and geometry are different. Na^+ assumes hexa-coordination and is within polar contact with main chain oxygen atoms of Asp246, Ser282, Arg285 and side chain oxygen atom of Ser291 and two water molecules. Na^+ interacting residues are disposed in an octahedral geometry and the interacting residues are roughly same as *E. coli* RK (**figure 4.1.12**).

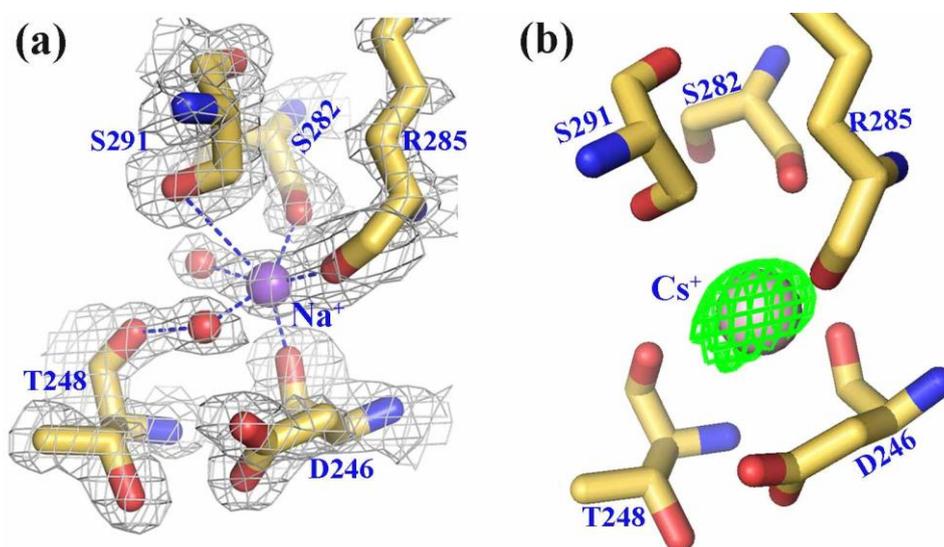


Figure 4.1.11: Monovalent cation binding site in VcRK. (a) 2Fo-Fc map contoured at 1.0σ (grey colored mesh) showing Na^+ ion (violet sphere) and the residues interacting with it (golden sticks). Waters are shown in small red spheres (b) Cs^+ ion (burgundy sphere), placed in the Fo-Fc difference map (light green mesh) at 5.0σ level, occupy roughly the same position where Na^+ ion binds.

In order to verify Na⁺ binding site precisely, ribose-ADP bound VcRK crystals were soaked with 50 mM CsCl and a data set was collected to a resolution of 2.35Å. While solving the structure without monovalent cation, larger electron density ~5.0 σ was obtained in Fo-Fc difference map (**figure 4.1.11b**) at the same position where Na⁺ was placed. Coordination of Na⁺ and Cs⁺ are found to be almost identical (**figure 4.1.11a, b**).

4.1.7 Sequence alignment VcRK with ribokinases from other sources

We have aligned the sequence of VcRK with RKs from other sources like *Escherichia coli* [15], *Staphylococcus aureus* [69], *Homo sapiens* [70], *Thermotoga maritima*, *Agrobacterium tumefaciens* (**figure 4.1.12**). Polar residues involved in ribose-binding and hydrophobic residues covering ribose are highly conserved throughout all RKs. The high conservation of sugar and ATP binding residues in different species prove that family divergence of ribose specificity occurred before species divergence [16].

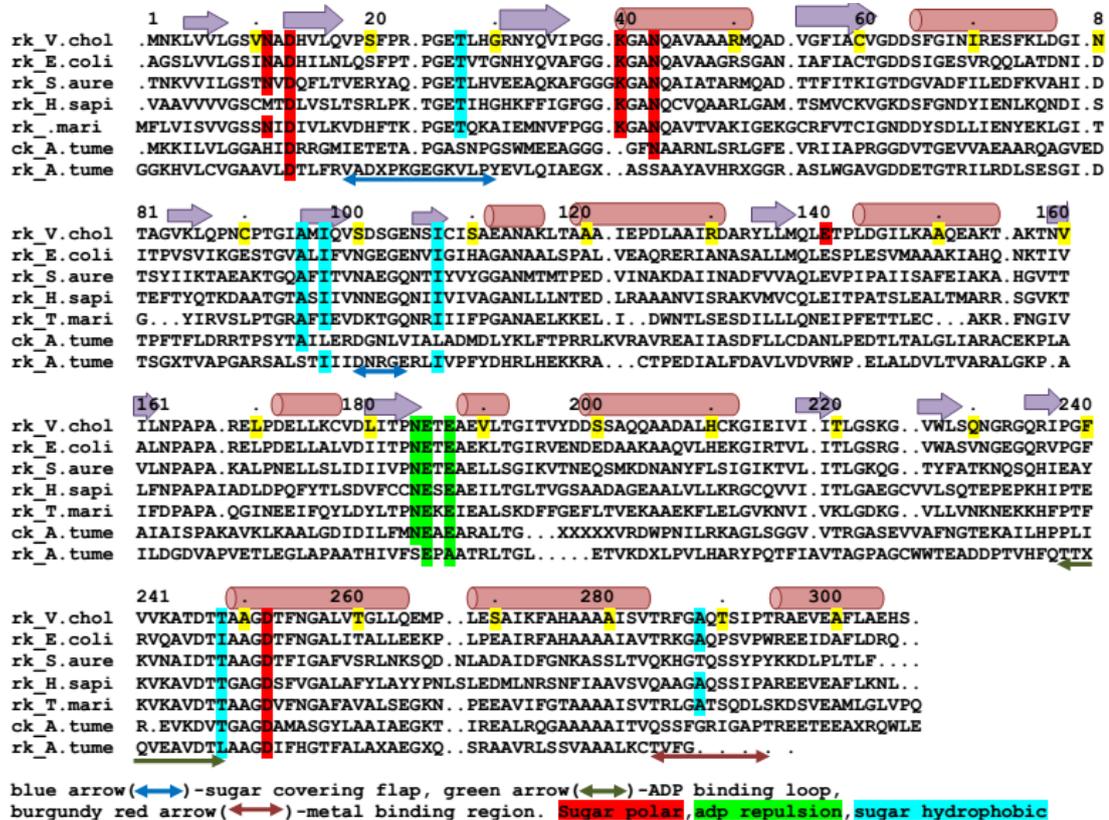


Figure 4.1.12: Alignment of VcRK with ribokinases from other sources

4.2 Structure of fructokinase from *Vibrio cholerae* O395

(VcFK)

4.2.1 Overall monomeric structure

Overall electron densities for all VcFK structures are of good quality, continuous and unambiguous for most of the regions. All the residues, except N-terminal fifteen amino acids, could be located in the electron density map. VcFK structures refined well and the stereochemical parameter obtained from the refined model is consistent with the quality of the diffraction data (**Table 3.7**). Each monomer of VcFK shows the typical ribokinase-like fold consisting of a central nine-stranded β -sheet surrounded by ten α -helices and two 3_{10} helices. **Figure 4.2.1a and b** shows the β -sheet core formed by the parallel arrangement of β -strands β 5- β 4- β 1- β 8- β 9- β 10- β 11 along with an antiparallel pair of β 12- β 13. Among the helices, α 4- α 5- α 6- α 7 and two short helices, named H1 and H2, contribute to one face of the β -sheet core while rests of the helices, containing most of the active site residues, reside on the other side. Like other members of Ribokinase superfamily, VcFK also possesses a four-stranded lid domain that extends from the central β -sheet core, as seen in **figure 4.2.1**. Two antiparallel pairs of β -strands, namely β 2- β 3 and β 6- β 7 and the connecting loops, form the lid domain. This domain plays a key role in dimerization of VcFK, like other members of RK superfamily. Moreover, structural flexibility of this loop orchestrate either in the entry of substrate or exit of product before and after the phosphotransfer reaction. β 2, β 3 and the joining loop called small lid loop, whereas β 6, β 7 and the joining loop called the large lid loop. The active site is contained between the lid domain and the central core. Like other homologous enzymes, here also ATP binds near a large loop between β 13 and α 8 (**Fig 4.2.1**) [60, 61].

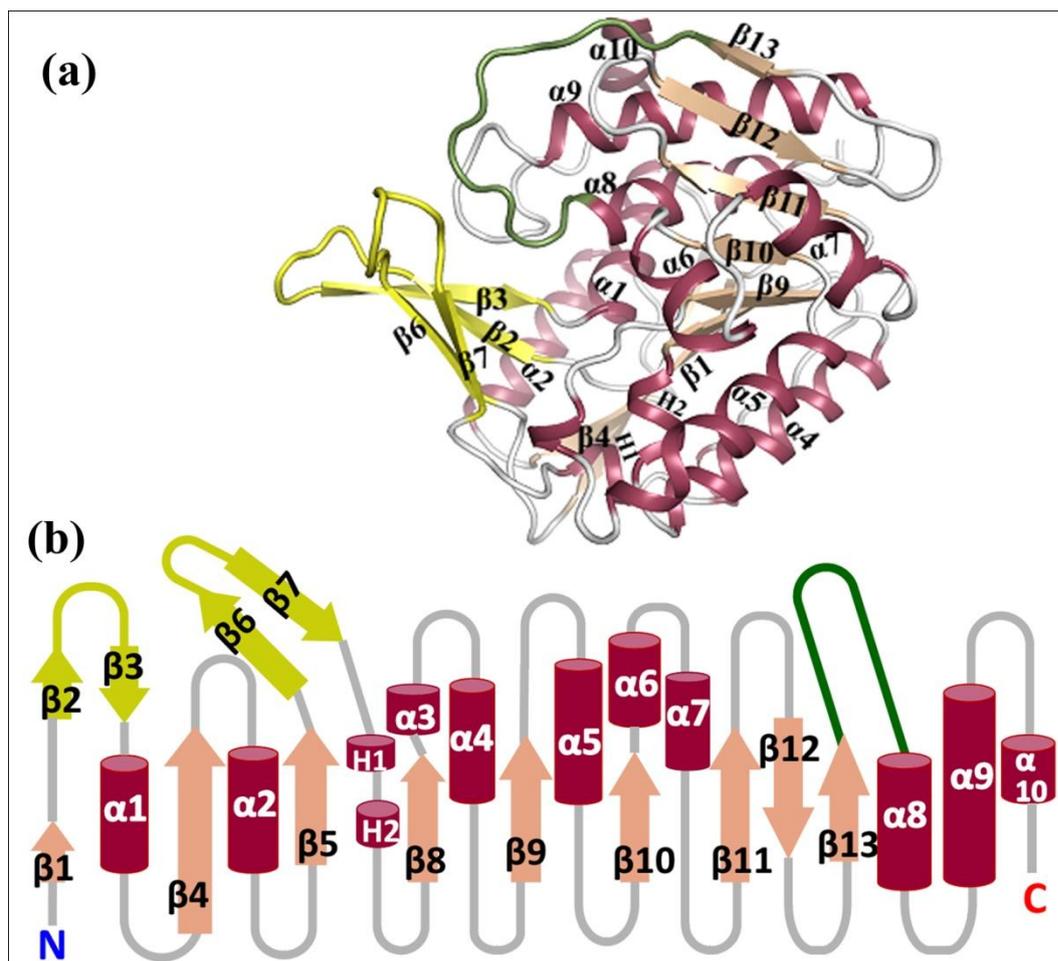


Figure 4.2.1: Secondary structure of VcFK. (a) cartoon representation and (b) topology diagram. For both diagrams, β -sheet constituting core domain is shown in coffee color, α -helices in magenta color and lid domain in pale-yellow. The ADP binding loop is highlighted in green.

4.2.2 Dimerization of VcFK

Like most of the enzymes of Ribokinase superfamily, both apo and substrate bound complexed VcFK exhibit a dimeric structure. In case of apo form, each asymmetric unit contains one dimer and the monomers are related by twofold non-crystallographic symmetry along Z-axis. The dimerization interface between two monomers is depicted in Figure 4.2.2a and its perpendicular view 4.2.2b. Unlike VcRK, where both β -turns twists and handshake to dimerize (figure 4.1.4), the β -turns of VcFK apo structure are arranged one by one to form a stretched β -sheet with an arrangement $\beta7A$ - $\beta6A$ - $\beta2A$ - $\beta3A$ - $\beta3B$ - $\beta2B$ - $\beta6B$ - $\beta7B$ (figure

4.2.1). β 2- β 3 and the connecting loop (pale-yellow color) of one monomer directly interact with other monomer at the centre of the dimerization region whereas β 6, β 7 reside on both ends and participate in dimerization. The loop joining β 6, β 7, however, stays away from the dimerization region. The stretched β -sheet is mainly stabilized through H-bonds between β 3 of each constituent monomer (blue circle in **figure 4.2.2b**). Another area of dimerization (red circle in **figure 4.2.2a**) is between the loop connecting β 5 to β 6 of one monomer with the loop connecting β 7 to H1 from the other (or vice versa). Side chain polar atoms from Asp34, His38, Thr100, Met116 and main chain N-atoms from Gln37, His38, Lys41 and Lys118 participate in H-bonding during dimerization.

The sugar-bound and sugar-ADP-Ca²⁺ bound structures of VcFK contain smaller unit cell volume than the apo form (**table 3.3**) which can only allow one monomer per asymmetric unit. Two adjacent symmetry related molecules assemble to complete the dimer (two perpendicular views are shown in **figure 4.2.2c** and **d**). Like apo structure, here also β 3 and the two loops joining β 5 to β 6 and β 7 to H1 participate mostly in interchain interactions (marked by blue and red ellipses). In addition, β 2 from each monomer slightly tilt towards β 3 of the other to optimize their extent of interaction. **Figure 4.2.2** demonstrate the dimer of VcFK in apo form (**a** and **b**) and in complexed form (**c** and **d**) keeping their orientation roughly same. Comparison of apo and complexed dimers elucidate that upon sugar binding, the region Asp106-Arg111 of the large lid loop (connecting β 6- β 7, light-blue color) folds back towards the active site but the small β -lid loop (β 2- β 3, pale-yellow) do not experience any such structural change and remain engaged in dimerization.

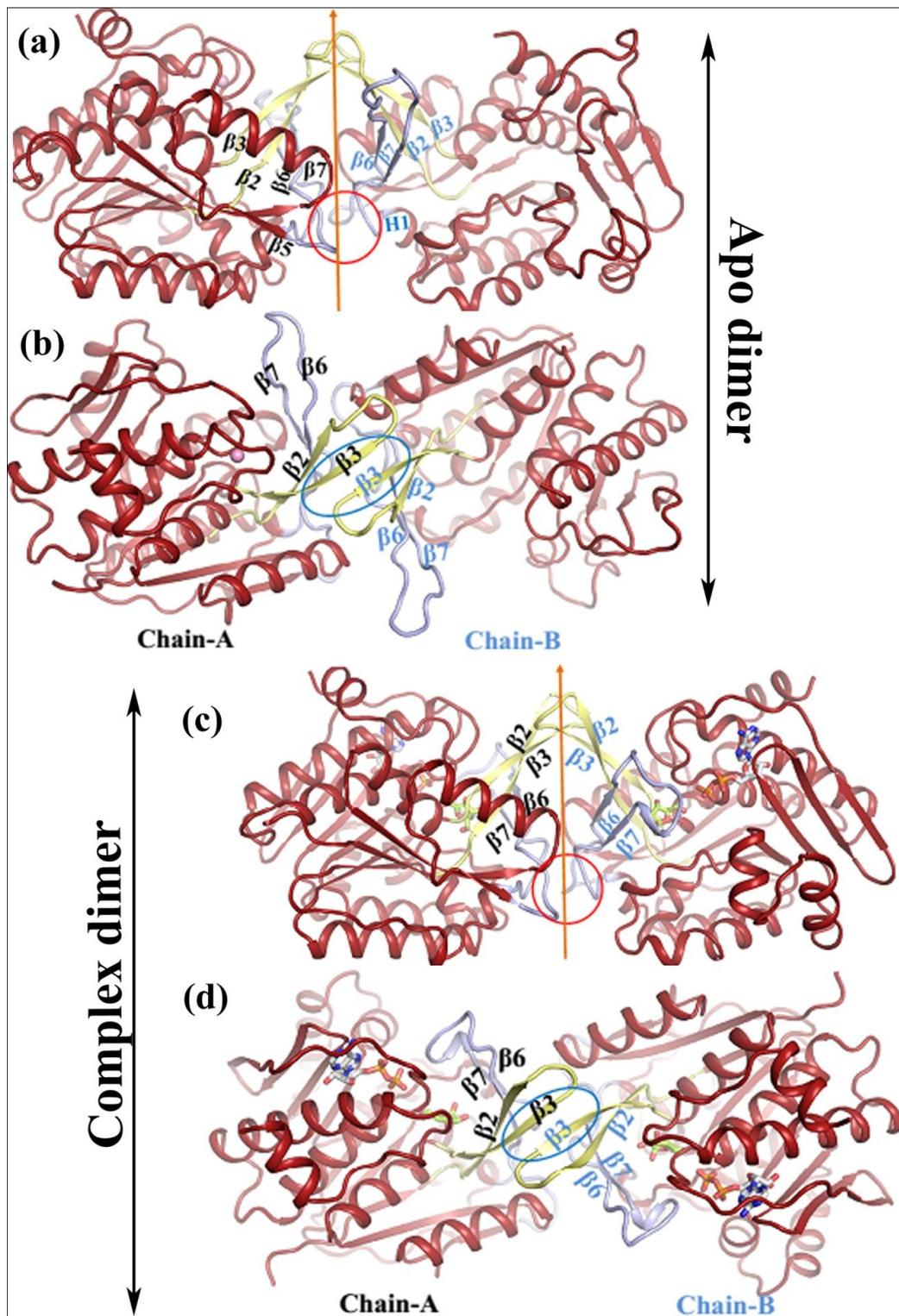


Figure 4.2.2: Cartoon representation showing the dimeric structure of VcFK in apo form (a and b) and in sugar-ADP bound form (c and d). (a) Dimerization occurs around the symmetry axis shown in orange. (b) 90° rotated view of dimer-interface. (c and d) complexed VcFK dimer shown in similar orientation of apo dimer (a) and (b). In all figures, central domain are shown by maroon color, β 2- β 3 are colored pale-yellow and β 6- β 7 were shown in light blue. In figures (a-d), labels for chain A are in black while labels for chain B are in sky-blue.

The dimerization interface buries several aromatic/hydrophobic residues like Leu31, Tyr39, Pro71, Phe72, Val102, Met116 and Val117 from each subunit contributing into considerable hydrophobic interactions between themselves (**figure 4.2.3**). PDBePISA [53] indicates that the total buried surface due to dimerization constitute about 7% of the total surface (1900 Å² out of total 25510 Å²) and the dimer is stable in solution ($\Delta G_{\text{interface}} = -13.3$ kcal/mol).

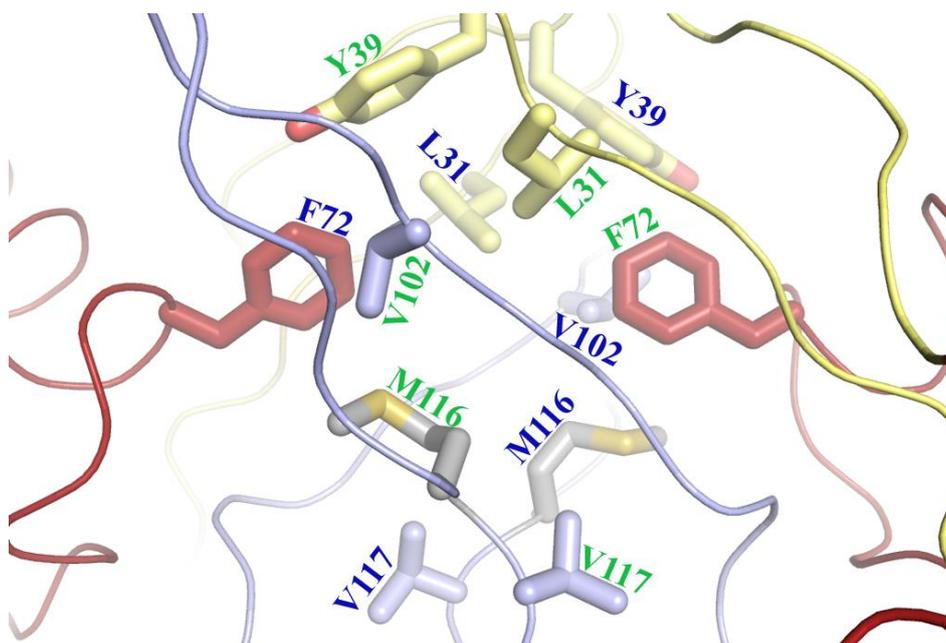


Figure 4.2.3: Hydrophobic residues involved in dimerization. Secondary structures are represented by ribbon and hydrophobic residues are shown in sticks. Residues of the two lid β turn $\beta 2$ - $\beta 3$ and $\beta 6$ - $\beta 7$ are shown in pale-yellow and light-blue color respectively whereas residues of the central domain are shown in maroon. Green and blue color label are used for chain A and B respectively.

4.2.3 Monovalent cation creates ATP binding site

Although both chains of apo VcFK superposes well, some of the loops around the active site exhibit significant structural dissimilarities. For thorough comparison, the two chains of VcFK were superposed separately with ribokinase (ribose+AMP-PCP+Cs⁺ bound structure, PDB code 1GQT, [11]) and AIRS kinase (AIRS+AMP-PCP+K⁺ bound structure, PDB code

1TZ6, [61]) structures (**figure 4.2.4a**). The binding sites for Cs⁺ and K⁺ are similar and chain A of VcFK superposes well with these two structures but chain B differs around the monovalent cation-binding site. An inspection of the electron density map indicated a positive density at 3 σ near the monovalent cation binding position for chain A only. Although the electron density looks somewhat similar to that of a water molecule, careful inspection of the disposition of the neighboring polar atoms and their distances indicated a pentavalent ligand binding geometry, which impelled us to assign this density as a cationic species. As Na⁺ was the only monovalent cation present in crystallization buffer and sodium ion has similar number of electrons to that of water molecules, this density was assigned as Na⁺.

Structural superposition of chain A and chain B (**figure 4.2.4b**) indicates few major structural differences near active site. In chain A, monovalent cation binds between green and cyan colored loops (named monovalent cation binding loop). Here, conformations of large ATP loop (green) and small ATP loop (pink) accommodate the ATP molecule without any clash. On the other hand, the large ATP loop assumes distorted shape in chain B and encroaches into the ATP binding site (loop shifts approximately 7Å measured between their Thr261 position **figure 4.2.4b**). However, absence of monovalent cation in chain B renders flexibility to the α 8- β 13 loop. Moreover, C-terminal helix could not be located in electron density map. In addition, the two lid loops in chain A (colored yellow and blue) move closer to the $\alpha\beta\alpha$ domain to produce more suitable environment for sugar binding. The ill-defined active site in absence of monovalent cation and eventual development upon sodium binding, is the key step for the monovalent cation induced activation of fructokinase.

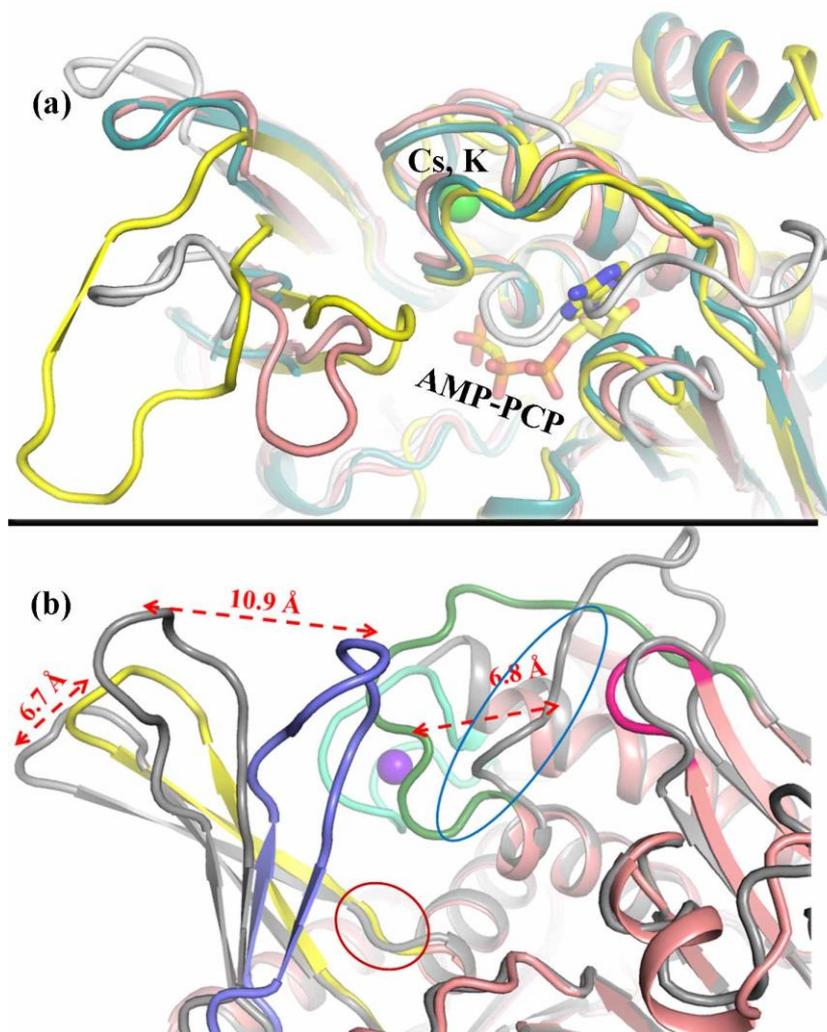


Figure 4.2.4: Structural changes due to monovalent cation binding in VcFK. (a) Alignment of chain A (burgundy and pink) and chain B (grey) of apo VcFK with *E. coli* Ribokinase (yellow) and *S. enterica* AIRS kinase (teal green). Monovalent cation is shown in green sphere. (b) Superposition of chain A and chain B of apo VcFK. Chain B is shown in grey, chain A is shown in burgandy cartoon with large ATP loop (green), small ATP loop (hot pink), monovalent cation binding loop (cyan), small lid loop (yellow) and large lid loop (blue). The fructose and ATP binding sites are indicated by red and blue ellipses. Monovalent cation is shown in violet sphere. Loop shifts are measured between Ca-atoms of Gln36 (small lid loop), His108 (large lid loop) and Thr261 (monovalent cation loop).

The main chain nitrogen atoms from the two loops (green and cyan, **figure 4.2.4b**) specifically Ala302, Gln305, Gly307 and two water molecules participate in sodium ion binding. The H-bonds around Na^+ ion assume a pentavalent geometry that is evident in

figure 4.2.5a. Electron density of the monovalent cation and its binding residues are unambiguous as seen in the 2Fo-Fc map (**figure 4.2.5b**). The distances between Na^+ and oxygen atoms vary from 2.0 Å to 2.5 Å, which are shorter than H-bonds formed by water molecules (~2.7-3.0 Å).

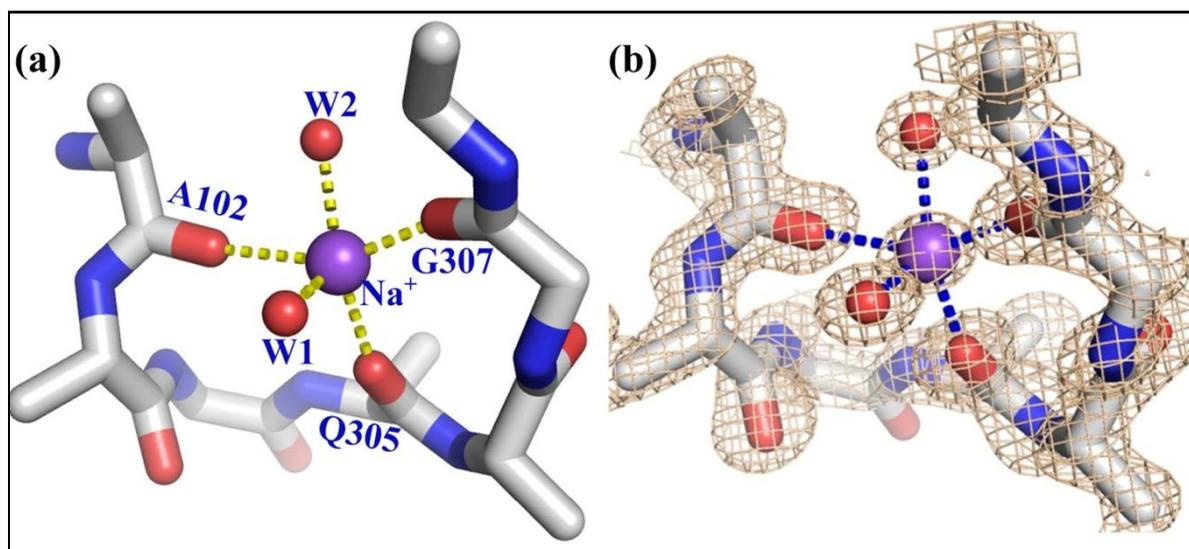


Figure 4.2.5 : Pentavalent geometry of the monovalent cation Na^+ . (a) The sodium ion is depicted in violet sphere, water molecules in small red sphere and the neighbouring residues are shown as sticks. For simplicity side chains of the neighbouring residues are shown upto $\text{C}\beta$ position. (b) 2Fo-Fc map in 1σ level showing electron density of sodium ion and the residues involved in binding Na^+ .

4.2.4 Fructose binding and associated structural changes

Like other Ribokinase superfamily members, VcFK also possesses a unique sugar-binding cleft. The fructose molecule binds in this well-formed cleft located exactly above the conserved GGA motif (G44-G45-A46) located between β_3 to α_1 (**figure 4.2.6**). The sugar molecule is surrounded by the large ATP loop (deep green), monovalent cation binding loop (cyan), the small lid loop containing β -strands β_3 - β_2 (yellow), large lid loop containing β_6 - β_7 (slate) and the long loop connecting β_8 to α_5 (brown) respectively, as shown in **figure 4.2.6**. In such a trapped condition, all the fructose atoms (except $\text{O}6'$) is totally segregated

from the solvent molecules and lose their flexibility, which are very important step before catalysis. The O6' of sugar is so oriented that it can attack the γ -phosphate of the ATP molecule and become phosphorylated.

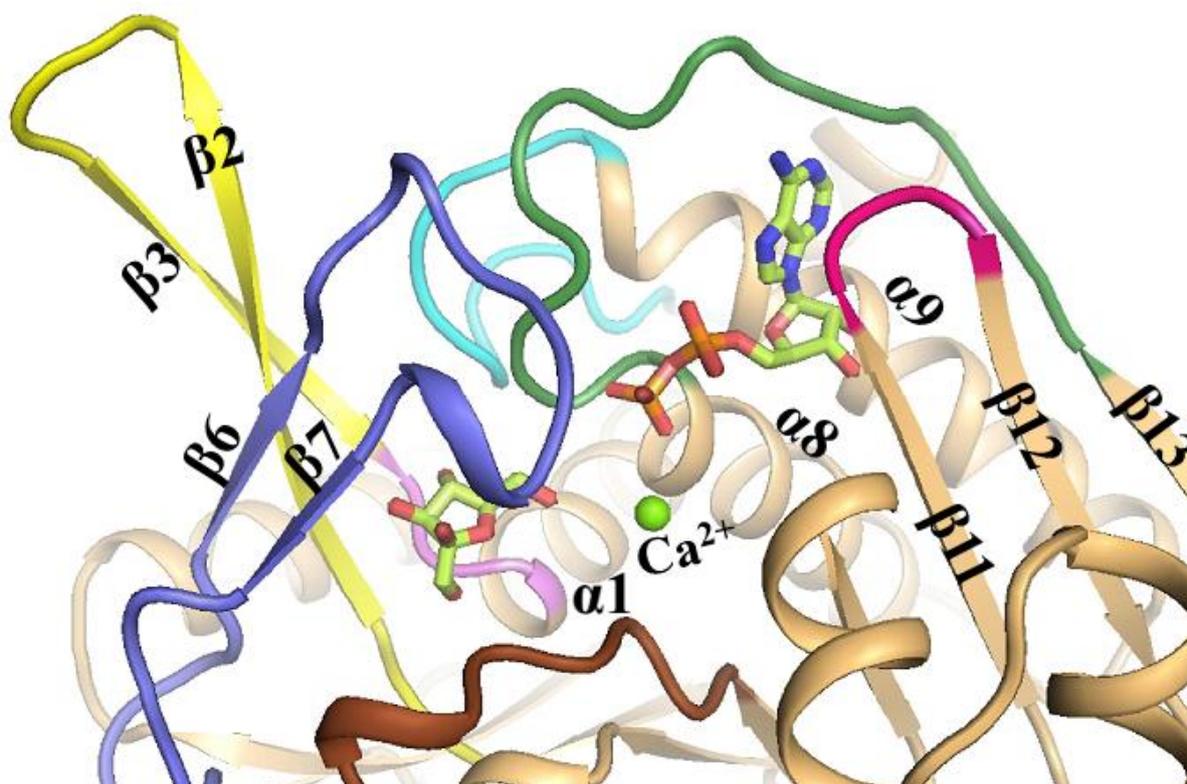


Figure 4.2.6: Entrapment of fructose molecule in VcFK, shown anti-clockwise from O6' end of fructose are large ATP loop (green), monovalent cation binding loop (cyan), yellow and blue β -sheet from lid, brown loop connecting $\beta 8$ - $\alpha 5$ respectively. The GGA motif at the base of fructose is colored pink. Rest of the VcFK molecule is shown in sand colored cartoon. The fructose and ATP molecules are represented in sticks (lemon-green). The calcium ion is shown in green sphere.

The VcFK bound fructose molecule assumes a relatively uncommon 5-membered ring form (**figure 4.2.7a**), called C4'-endo configuration of D-fructo-furanose, where the C4' atom moves upward from the plane containing O5' (**figure 4.2.7b**). Oxygen atoms of the fructose residue are involve in hydrogen bonding with the neighboring side chain polar atoms of

residues Asp26, Asp30, Asn49, Arg176, Asp266 and main chain nitrogen atoms of Gly45 and Ala46 (**figure 4.2.7c**). To make hydrogen bonds these residues changes their conformation and reorient to optimize H-bond with sugar. Altogether, nine H-bonds are made between fructose and the VcFK molecule that tightly bind the sugar with the enzyme. Moreover, Phe113 moves about 5.5Å to covers the hydrophobic parts of the sugar.

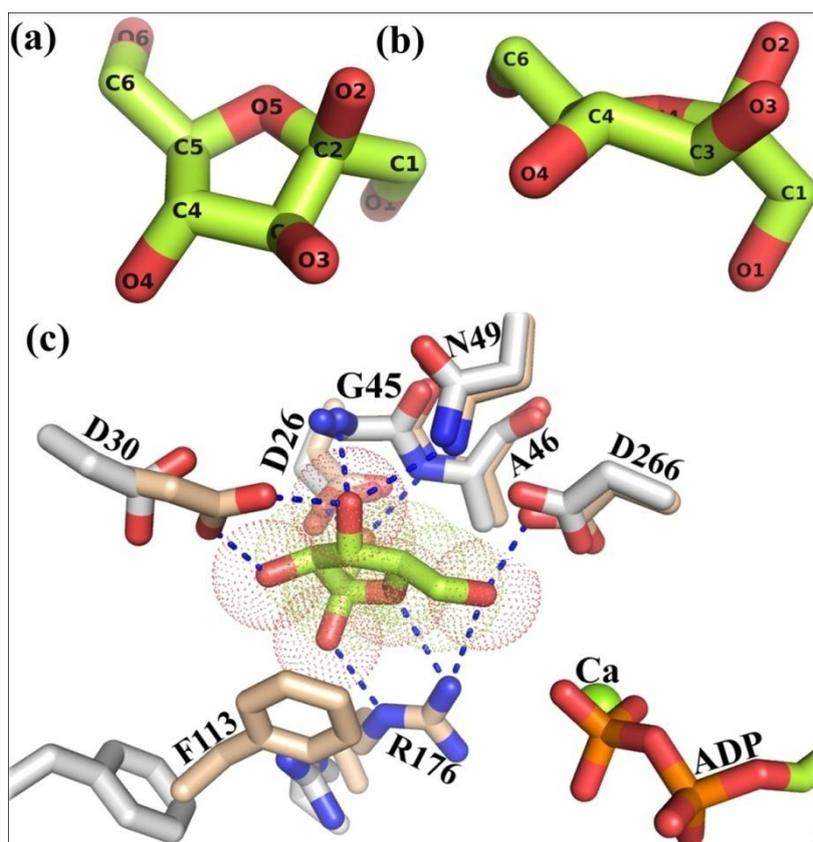


Figure 4.2.7: Detailed interaction of fructose with VcFK. (a) Furanose ring of fructose where all the atoms are distinct. (b) Rotated view of figure (a) where C4' is above the plane of O5' to form C4'-endo configuration. (c) Residues involved in fructose binding are shown in stick representation. Fructose bound VcFK molecule is shown in sand color, apo VcFK molecule is shown in grey and the fructose molecule in lemon-green sticks along with dotted surface.

4.2.5 Loop movement associated with sugar binding

In section 4.2.3, the structural changes in loop region due to monovalent cation binding was discussed in detail. To disclose the structural changes in detail, sugar bound

VcFK structure has been superposed with both chains of apo form (**figure 4.2.8**). The large lid loop including $\beta 6$ - $\beta 7$ partially closes after monovalent cation binding and the closure completes upon fructose binding. Short lid loop including $\beta 2$ - $\beta 3$ shifts only due to monovalent cation but not for sugar binding. Residues like Val103 and Phe113 from the lid reallocate themselves to cover the fructose molecule hydrophobically. Upon fructose binding, the Glu110 residue shifts 10.7\AA to make hydrogen bond with main chain N-atom of Glu177 (**figure 4.2.8**) and stabilize the sugar entrapped closed structure.

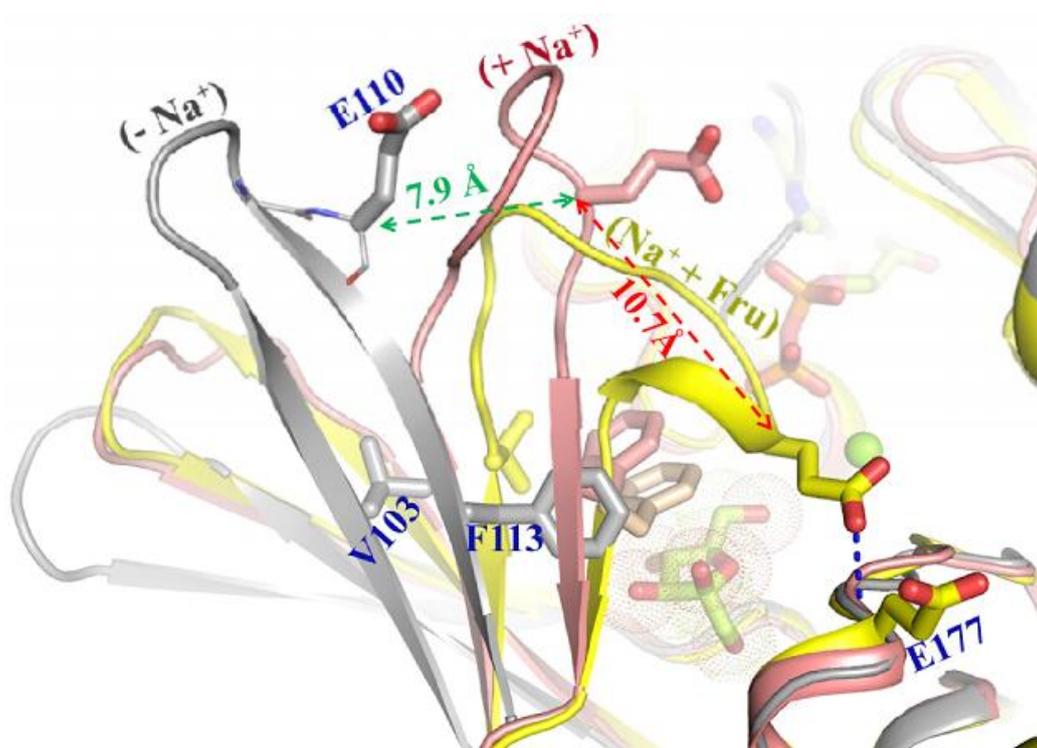


Figure 4.2.8: Movement of large lid loop on sugar binding. Fructose bound (yellow) structure has been superposed with chain A (monovalent cation bound, burgundy color) and chain B (without monovalent cation, grey color) of apo form. Hydrophobic lid residues are indicated by sticks, fructose is represented by lemon-green stick along with dotted surface.

4.2.6 ADP binds in a pocket formed upon monovalent cation binding

ADP molecule binds in a shallow groove between two surface exposed loops, a bigger loop that connects $\alpha 8$ - $\beta 13$ called large ATP loop and a smaller loop connecting $\beta 11$ - $\beta 12$

called small ATP loop (**figure 4.2.6**). The pocket remains distorted before monovalent cation binding but shapes up upon sodium binding to accommodate the ATP molecule.

The ribose sugar part of ADP binds with the oxygen atoms of Gly 239 main chain and Asn 296 side chain. The imido group binds with the main chain O-atoms of Val256 only. The β -phosphate group remains within H-bonding distance from main chain polar atoms of Thr261, Gly263, Ala264 and Gly265 and Ca^{2+} ion. However, the overall interactions of FK with ADP are weaker compared to that of fructose. This may probably due to ADP being a product of the catalytic reaction. Superposition of ADP bound structure with apo structure indicates that ATP binding do not induce much structural changes (**figure 4.2.9**).

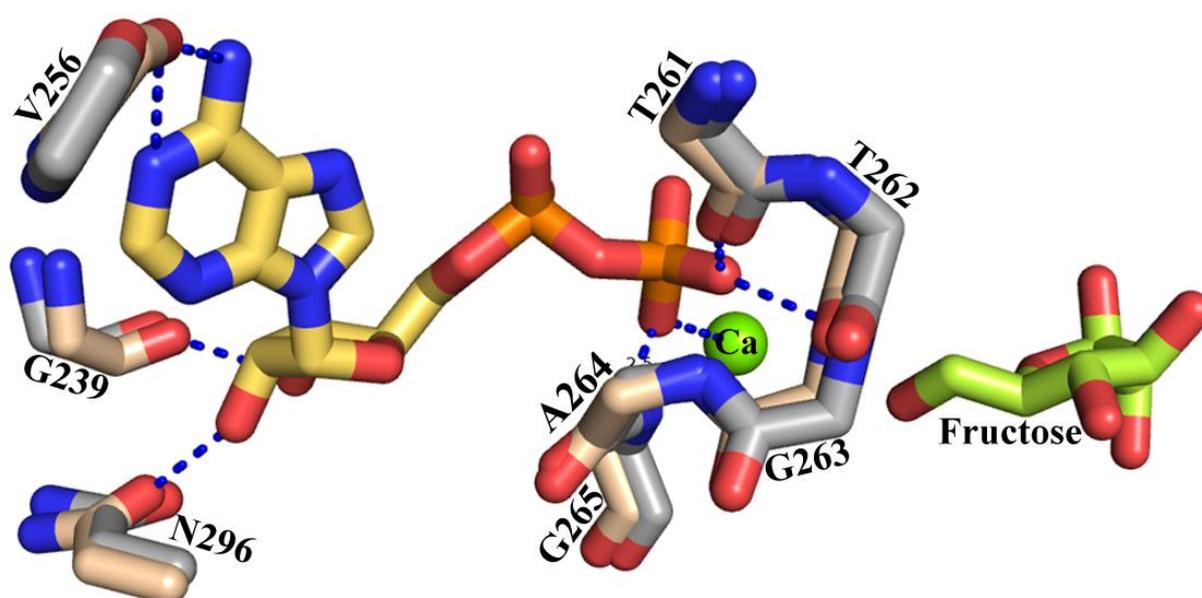


Figure 4.2.9: The nucleotide-binding site of apo and sugar-ADP bound VcFK structure. The residues involved in Ca^{2+} coordination and ADP are shown in stick mode. Sand colored sticks represent ADP bound structure whereas apo VcFK structure is shown in grey sticks.

4.2.7 Divalent cation binding site

Few members of PfKB family like AIRs kinase possess a divalent cation bound between the β - and γ -phosphates of ATP analog like AMP-PCP or AMP-PNP. These

molecules contain three phosphorus atoms but the oxygen atom connecting the β - and γ -phosphate groups is replaced by carbon and nitrogen respectively, rendering AMP-PCP or AMP-PNP unable to phosphotransfer. In VcFK, the active site binds with ADP, having only two phosphorus atoms and therefore it is unlikely to observe the divalent metal ion at the space specified, similar to VcRK. However, electron density of divalent cation has been obtained near the β -phosphate and fructose residue in VcFK. The electron density is much larger than that of water molecules (**figure 4.2.11a**); it is further confirmed to be divalent calcium ion because it has octahedral coordination geometry where all the neighboring atoms remain approximately at 2.4 Å distance (**figure 4.2.10**). Five water molecules and the side-chain oxygen atom from Asn174 complete the octahedral coordination of Ca^{2+} .

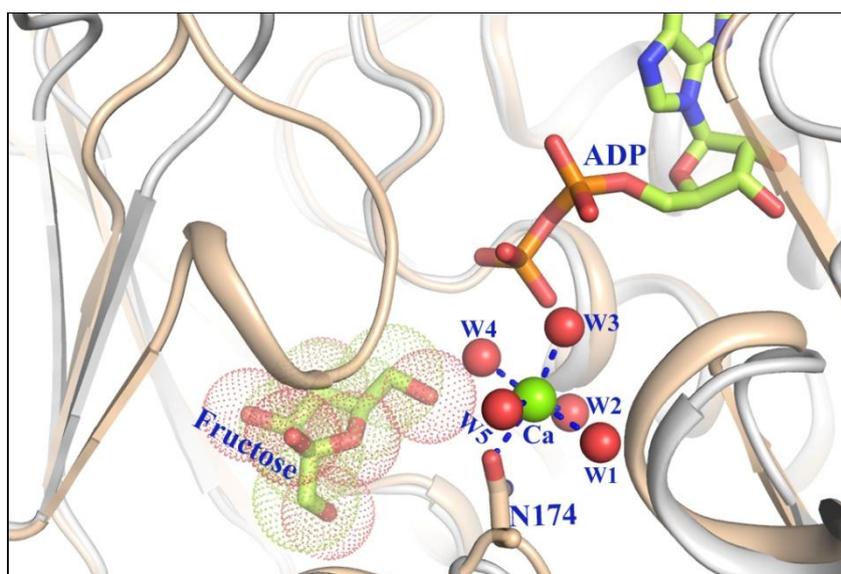


Figure 4.2.10: Octahedral geometry of calcium ion in between sugar and β -phosphate of ADP.

The Ca^{2+} ion is indicated by light green sphere; water molecules within its coordination distance are shown in red spheres, ADP and fructose molecules by lemon green sticks. Apo (grey) and sugar ADP- Ca^{2+} bound VcFK (sand) are shown in ribbon.

4.2.8 BeF₃ bound structure

ADP lacks one phosphate group and the β -phosphate of ADP remains 5.7 Å away from the O6' atom of fructose. Crystallization trial using ATP analogue AMP-PNP went in vein probably due to the hydrolysis of AMP-PNP at the pH of the crystallization condition used here. In 1997, Xu et al. [71] showed that for nucleoside diphosphate kinase, metallo-fluorides could replace the position of γ -phosphate, when used with ADP. AlF_4^- could not serve the purpose due to presence of four fluorine atoms and planar geometry, but AlF_3 or BeF_3^- could. Attempt to crystallize *E.coli* RK (PDB code 1RK2, [60]) with AlF_4^- showed that AlF_4^- binds elsewhere along with solvent molecules but not into the specific site. The neutral AlF_3 has trigonal bipyramidal geometry and BeF_3^- exhibits tetrahedral geometry; therefore, BeF_3^- proves to represent the exact bulk and charge environment of γ -phosphate.

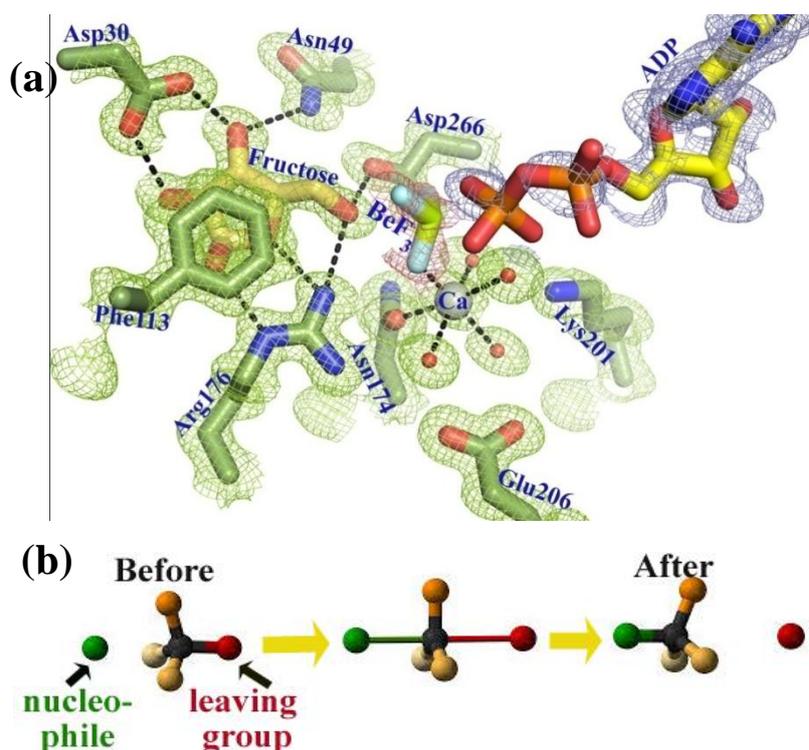


Figure 4.2.11: Overall active site and geometry of phosphotransfer. (a) 2Fo-Fc map at 1.0 σ level around the active site (including fructose, ADP, Ca²⁺, BeF₃⁻ and few catalytically important residues in stick representation) justify the acceptability of the structure. (b) Illustration of Walden inversion by ball and stick model during a S_N2 reaction.

In VcFK, fructose-ADP-Ca²⁺ bound crystals were soaked with freshly prepared (BeCl₂+NaF) solution producing BeF₃⁻. High-resolution (1.29 Å) structure shows the presence of BeF₃⁻ in electron density map (**figure 4.2.11a**) located roughly where γ -phosphate is supposed to occupy. Here BeF₃⁻ binds exactly midway between the β -phosphorus atom and the O6' of fructose residue (3 Å from each). In BeF₃⁻, fluorine atoms remain in three corners keeping beryllium atom at the vertex. The plane containing three fluorine atoms head towards the substrate, with the vertex Be pointed towards ADP. According to Walden inversion [72], in a S_N2 reaction the chiral center of tetrahedral molecule inverts from the leaving group towards nucleophile like an umbrella turns inside out (**figure 4.2.11b**). Therefore the geometry of BeF₃⁻ from **figure 4.2.11a** proves that the transition state formed in VcFK was before phosphorylation begins.

CHAPTER-5

DISCUSSION

In the preceding chapters we have reported the structures of two sugar kinases, namely ribokinase and fructokinase, belonging to ribokinase superfamily, from *Vibrio cholerae* O395. Basically these two enzymes use the γ -phosphate of ATP to phosphorylate the hydroxyl oxygen of their cognate sugar. Divalent cation is necessary for this catalytic reaction whereas monovalent cation acts as an allosteric effector. Based on their structural studies on *E. coli* ribokinase, Mowbray group [15, 60, 11] proposed the basic mechanistic details and identified several active site motifs like ATP binding loop, monovalent cation binding loop and anion hole, β -clasp etc. that are implicated to perform a definite part towards its overall catalytic function. Since then structures of sugar kinases like adenosine kinase [73], aminoimidazole riboside kinase [61], phosphofructokinase [74], 2-keto 3-deoxy gluconate kinase [75], belonging to ribokinase superfamily have been reported but detailed structural studies about the members of fructokinase is limited to the structure of *H. orenii* fructokinase in apo form [62]. Moreover, the exact mechanism of allosteric activation was unknown. In a recent work, Sa239 ribokinase structure is reported in apo form and upon comparison with *E. coli* RK it is possible to shed light on the activation mechanism by monovalent cation. Crystal structure of apo and three complexed fructokinase provide a platform to visualize the detailed substrate induced structural changes and when coupled with the kinetic data allow us to propose a reaction mechanism of this family. Moreover, two different chains of apo FK in same asymmetric unit provide the structural basis of monovalent cation induced activation.

5.1 Monomeric structures of VcRK and VcFK display some common features with other ribokinase superfamily members

Overall the structures of VcRK and VcFK are very similar. Their central β -stranded core domain and most of their helices superpose well, while their substrate-binding site and

the lid region participating in dimerization are quite different **figure 5.1a**. In particular, the lid (β 2- β 3) of VcRK is 10 residues longer compared to that of VcFK and suffers perpendicular bending for dimerization and sugar binding whereas in VcFK, the two lid β -turns (β 2- β 3 and β 6- β 7) are of nearly equal length and the lid between β 6- β 7 folds back upon sugar binding. Moreover, VcFK contains an insertion of 10 residues (connecting β 9 and α 6) near the sugar-binding site indicated by cyan color in **figure 5.1a**. Superposition of VcRK and VcFK with other members of ribokinase superfamily available in PDB (**figure 5.1b**) such as *Polaromonas sp. JS666* phosphofructokinase B (**PfKB**, PDB ID 4DU5), 2-keto-3-deoxygluconate kinase from *S. tokodaii* (**KdgK**, PDB ID 2DCN), **VcFK**, *S. enterica* Airs Kinase (**AirK**, PDB ID 1TZ6 [61]), adenosine kinase from *S. meliloti* (**AdnK**, PDB ID 3UBO), LacC from *E. Faecalis* (**LacC**, PDB ID 2F02) and *L. innocua* D-Tagatose-6-Phosphate Kinase (**TagK**, PDB ID 3JUL) display same feature, with gross superposition of their secondary structural elements, ATP binding loop and monovalent cation binding loop but varying degree of dissimilarities in their lid domain (orange ellipse) and substrate-binding region (encircled in violet). Therefore it seems that similar framework is used to bind ATP and monovalent cation while the variable regions evolved to bind different substrates. Structure based sequence alignment with these proteins provide some interesting results and certain portions of the alignments are shown in **figure 5.1c**. It shows that the base of substrate binding site (red) contains a conserved Glycine residue. In the ATP binding site, the adenosine part of ATP molecule do not make strong hydrogen bonds with the neighboring residues, rather than it binds via hydrophobic interaction. This fact is reflected in the sequence alignment (green, **figure 5.1c**), which contain mostly hydrophobic residues with no conserved pattern. The γ -phosphate of ATP is stabilized by the anion hole (N-terminal of α 8, blue), which remain strictly conserved (AGD motif, blue). Most of the RK superfamily members bind monovalent cation with the main chain carbonyl oxygen of the large ATP loop

(green) and monovalent cation binding loop (yellow). Consequently, these loops adopt similar structure but their amino acid sequences are different (yellow, **figure 5.1b and c**).

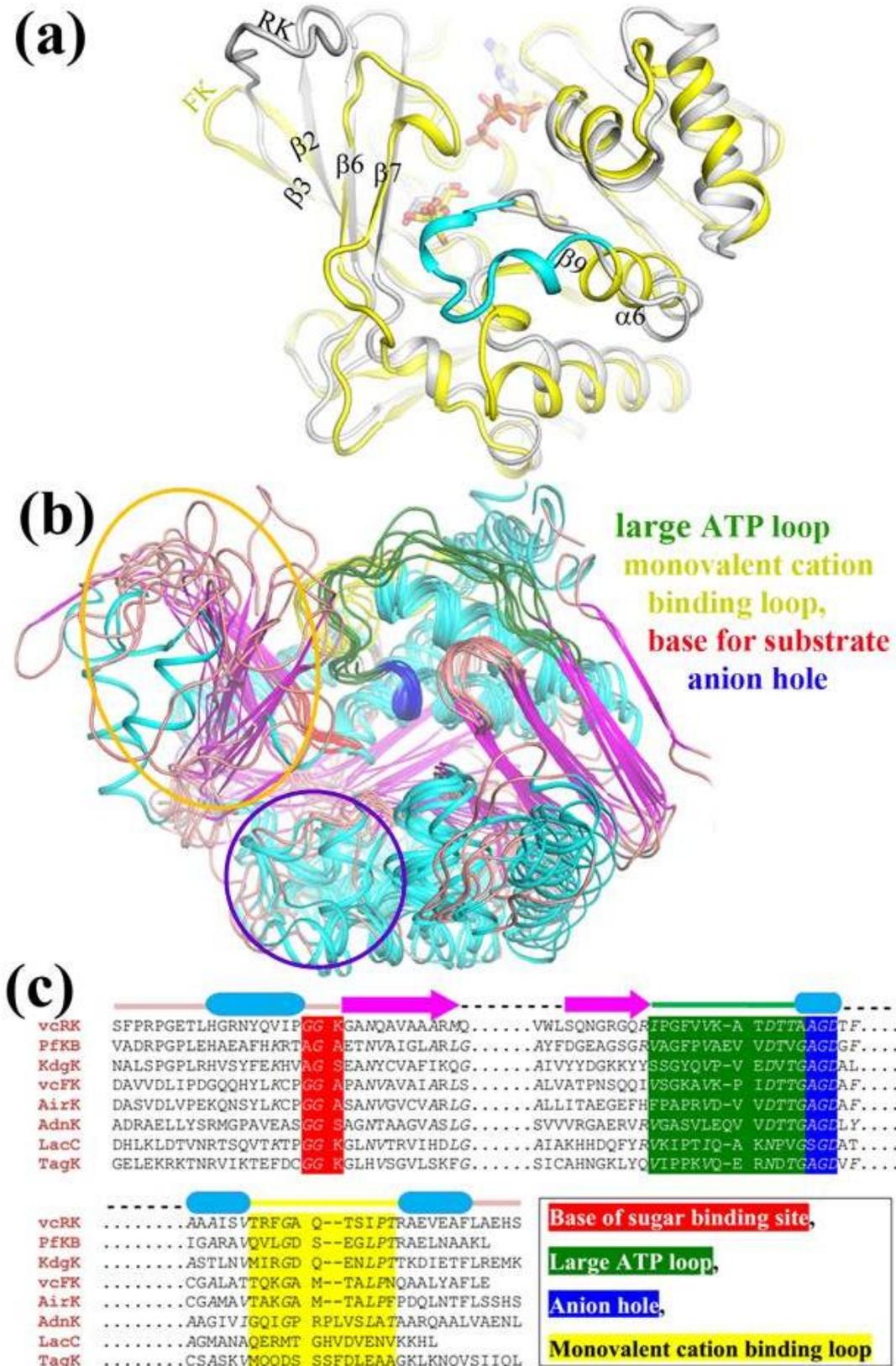


Figure 5.1: Structure and sequence alignments of ribokinase family of proteins. (a) Superposition of VcFK (yellow) and VcRK (grey) monomers in cartoon representation. The insertion region is shown in cyan. (b) Structural alignment and (c) sequence alignment (active site neighbourhood shown only) of substrate bound forms of **VcRK**, *Polaromonas sp. JS666* phosphofructokinase B (**PfKB**, PDB ID 4DU5), 2-keto-3-deoxygluconate kinase from *S. tokodaii* (**KdgK**, PDB ID 2DCN), **VcFK**, *S. enterica* Airs Kinase (**AirK**, PDB ID 1TZ6), adenosine kinase from *S. meliloti* (**AdnK**, PDB ID 3UBO), LacC from *E. Faecalis* (**LacC**, PDB ID 2F02) and *L. innocua* D-Tagatose-6-Phosphate Kinase (**TagK**, PDB ID 3JUL). In both (b) and (c), same colors are used to highlight the base of sugar binding, anion hole, large ATP loop and monovalent cation binding loop.

5.2 Tight and specific binding of sugar induces large structural changes

Both VcRK and VcFK tightly bind their respective sugars with 9-11 hydrogen bonds that restrict the conformational freedom of the sugar, a crucial step before phosphorylation. Moreover, binding of sugar is such that its hydroxyl group to be phosphorylated oriented towards the γ -phosphate of ATP. Disposition of hydrogen bonding residue of the enzyme with sugar clearly indicate that binding is highly specific, so much so that one anomeric form of sugar is selected when different forms exist in solution. In addition, sugar molecule is covered such that it remains mostly secluded from the solvent. Thus, sugar binding is associated with considerable structural changes, largely in the flap/lid domain, which vary in different members of the ribokinase superfamily and influenced by the type of oligomerization adopted by the family.

5.2.1 Ribose binding triggers the lid closure in VcRK, not ADP

Complexed VcRK contain four monomers in the form of two dimers in each asymmetric unit as discussed in section 4.1.4.2. Alignments of these four chains (**figure 5.2**) indicate complete superposition of the central $\alpha\beta\alpha$ domain but all differences are seen in their

lid region. Chains A and C binds with one ribose and one ADP molecule at their specific binding site, therefore the two structures are very similar with lid loops assuming closed structure. Chain D contains only ADP molecule at its usual binding site but no ribose molecule at its sugar binding site. Here, no loop closure is observed and the structure resembles the apo form which proves that the ribose is involved in lid closure. Chain B contains two ADP molecules in one monomer (**figure 5.2**), one at the specific position and the other at the sugar binding site which binds the ribose part of ADP. This is the unique feature of ribokinase to bind ADP at sugar binding site, in spite of its high substrate specificity. This is probably possible because the sugar part of the nucleotide is the same as the cognate sugar of the enzyme. This structure assumes an open form and resembles the ‘ADP-only’ bound structure.

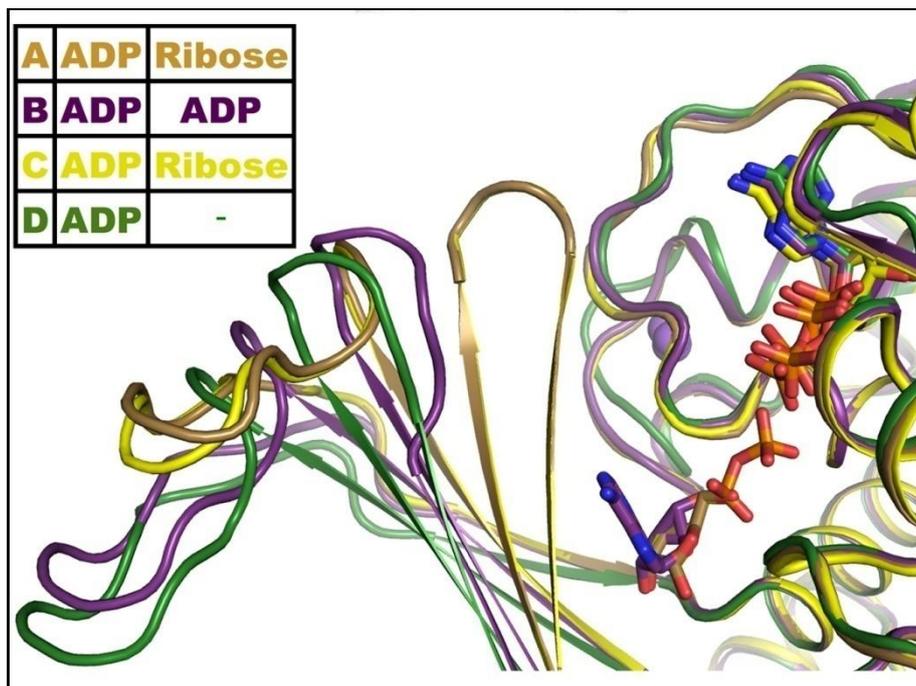


Figure 5.2: Relative movement of β -clasp region for different chains. β -clasp is closed for chain A and chain C (yellow and brown) while it is open for chain B (violet) and chain D (green). ADP and ribose is shown in sticks.

5.2.2 Sugar binding site of fructokinase is highly selective

5.2.2.1 Fructokinase phosphorylates fructose only at O6' position but not at O1'

Fructose can be phosphorylated at either O1' or O6' or both (**figure 5.3**) at different steps of metabolism. Kethexokinase (KHK) and fructokinase specifically phosphorylates fructose at the O1' and O6' positions respectively to form fructose 1-phosphate and fructose 6-phosphate. What is the basis of this specificity?

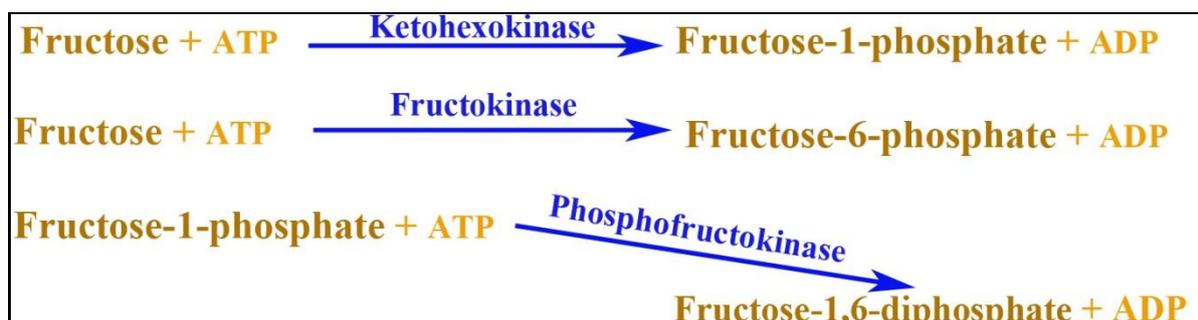


Figure 5.3: Phosphorylation at O1' and O6' positions of fructose by three different kinases.

In the superposition of VcFK and *H. sapiens* KHK (PDB ID 3NBV) [76], sugar binding site shows that O1' of fructose in KHK superposes with O6' of fructose in VcFK (**figure 5.4a**). The position of O2' of fructose is however oriented in opposite direction. In VcFK, O2' binds with polar arginine residue from a loop which is absent in KHK. The polar O2' binding residue asparagine of KHK is replaced by hydrophobic alanine in VcFK. Therefore, if VcFK has to bind fructose molecule like it was in KHK a couple of changes at its sugar binding site is needed. In other words, fructokinase cannot be a substitute enzyme for phosphorylation of fructose at O1' position.

5.2.2.2 Specificity of fructokinase towards fructose over AiRs

VcFK and aminoimidazole riboside kinase (AiRsK) from *S. enterica* [61] possess 42% sequence identity with similar overall secondary structure. The sugar AiRs contains a

ribose part along with an imido group connected through C1' of ribose. The O5' of ribose part is phosphorylated by AiRsK. When complex VcFK structure is superposed with the complex AiRs kinase, O5' of ribose part of AiRs matches with O6' of fructose (**figure 5.4b**). Most of the AiRs binding residues are same in VcFK except at two places. In AiRsK near imido group, there was a tyrosin residue which is phenylalanine in VcFK and it come closer to fructose as fructose is smaller in size than AiRs. Moreover, Ser14 of AiRs kinase is a valine in VcFK.

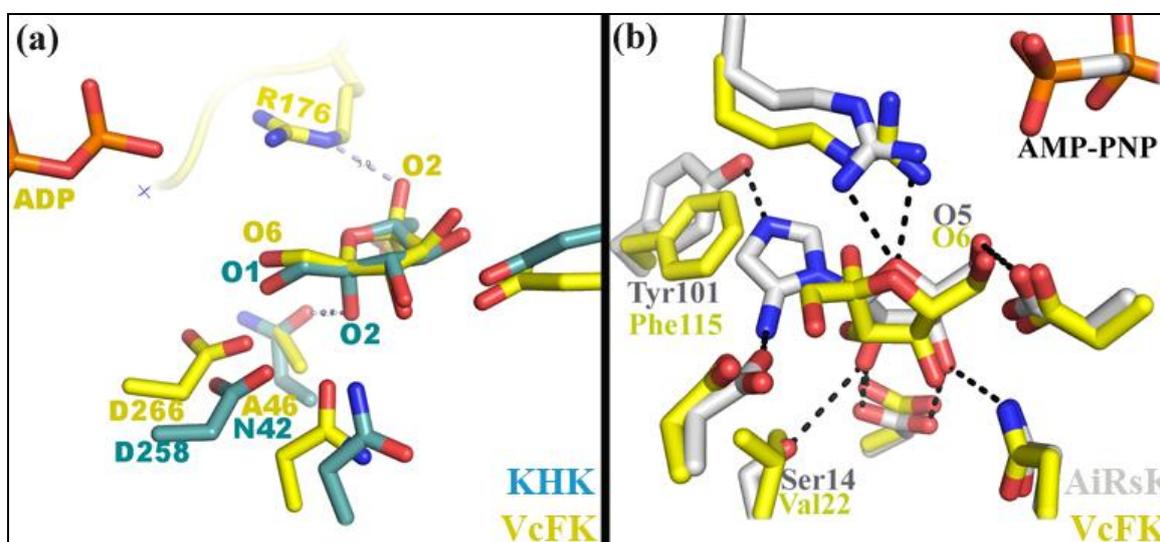


Figure 5.4: Specificity of VcFK towards fructose. Alignment of VcFK (yellow sticks) with (a) ketohexokinase (teal green) and (b) AiRsK (grey) depicting origin of specificity towards fructose.

5.3 Mode of dimerization and its role in phosphotransfer fidelity

Correct orientation and proper proximity are the two key factors for a fruitful chemical reaction to occur between two reacting groups. VcRK and VcFK accomplish these two factors, during phosphorylation of ribose or fructose by ATP, through two different methods. Moreover, to increase the fidelity of the reaction it is necessary to ensure that the γ -phosphate of ATP is not reacting with the bulk solvent but is secluded enough from the

solvent to stop abortive phosphorylation. In addition, γ -phosphate of ATP should consistently be directed towards the hydroxyl group of the sugar to be phosphorylated. Binding of ribose or fructose through several hydrogen bonds, mediated by some conserved residues in the respective kinases, not only freezes of the degree of freedom of the sugar but also orients its hydroxyl group to be phosphorylated towards ATP. Moreover, covering of sugar through several hydrophobic residues is so extensive and quantitatively precise that only the reacting hydroxyl group of sugar is made available to ATP. Extensive interactions made by the sugar triggers large scale conformational motion involving the protruding lid domain of the kinases. Lid domain of VcRK and VcFK, however, is engaged in dimerization and their involvement in dimerization is quite different. While the lid domain of VcRK form a handshake like interactions **section 4.1.3**, VcFK form a side by side arrangement of lid domain Section 4.2.2.

Dimerization of VcRK is such that its large lid loop of one subunit approach closer to the ribose-binding site of the other (**figure 5.5 a and b**). This step completely covers the ribose residue before the γ -phosphate of ATP approach towards it. This loop may also restrict the conformation of ATP through steric and/or charge repulsion and only select those conformations where γ -phosphate group of ATP directs towards sugar. The tip of this loop contain a highly conserved 'PGET' motif which is very close to ribose (closest distance 4.2 Å) and ADP (closest distance ~ 5.8 Å, which will be smaller in case of ATP) and held in place through several hydrogen bonds with residues of the other monomer. The strategic location occupied by this motif impelled us to speculate its role in the covering of sugar and ATP that reduces abortive phosphorylation and increase the reaction fidelity.

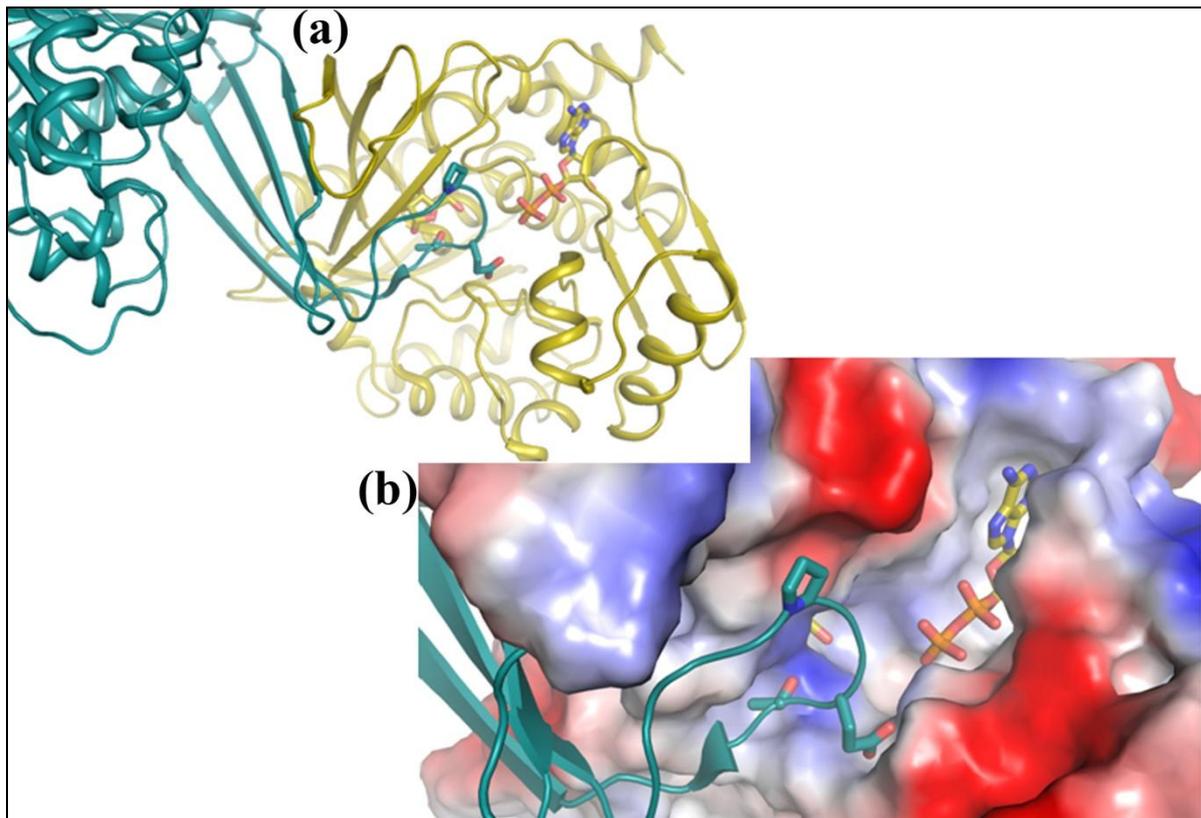


Figure 5.5: Covering ADP and ribose by the lid domain during dimerization of VcRK. (a) Sugar and ADP of one monomer (golden cartoon) is covered by the lid residues of the other monomer (teal green cartoon), (b) Same as (a) with golden cartoon is represented with surface electrostatic.

Unlike *E. coli* ribokinase or VcRK, where the protruding lid domain forms a β -clasp, dimerization of VcFK is side by side where the spatial disposition of lid domains are quite different from that of VcRK. The question is how VcFK maintain the fidelity? Superposition of VcFK structure with ribokinase structure gives a clue to it (**figure 5.6**). In ribokinase lid loop of one monomer (between $\beta 2$ - $\beta 3$, colored pink) enters the active site of the other (**figure 5.6a**). However, the same lid loop in VcFK is much shorter and differently disposed (**figure 5.6b**). Sugar covering in RK is made by the $\beta 2$ - $\beta 3$ lid loop whereas in VcFK, it is done by two other loops (colored green), one is the insertion (between $\beta 9$ - $\alpha 6$) and other is large movement of $\beta 6$ - $\beta 7$ loop due to sugar binding. These two loops spatially complement the $\beta 2$ - $\beta 3$ lid loop of RK and maintain the fidelity.

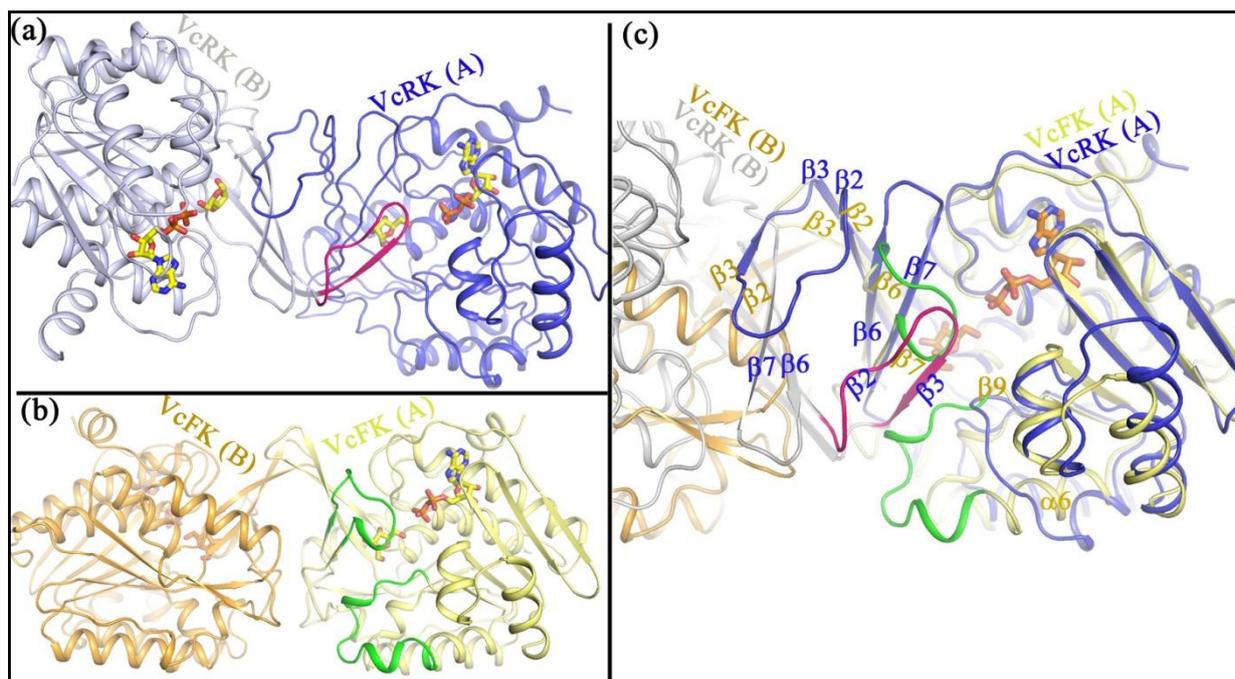


Figure 5.6: Difference in dimerization of *E.coli* RK (EcRK) and VcFK. Chain A of both structures are superposed, but chain B are oriented perpendicular to each other to make the dimer axis mutually perpendicular. (a) In EcRK, dimer axis is perpendicular to plain of paper, (b) in VcFK, dimer axis is a vertical line along the plain of paper. (c) Pink lid-loop in EcRK from neighbouring monomer covers sugar, whereas green lid loop and insertion loop in VcFK from self-monomer covers sugar residue. ‘Blue’ labels are used for VcRK and ‘sand color’ is used to label VcFK.

The insertion region between $\beta 9$ - $\alpha 6$ is a common feature of the proteins belonging to FKs and also in AiRs kinase [61] and phosphofructokinases [74]. Large movement of $\beta 6$ - $\beta 7$ loop to cover the sugar is seen only in VcFK. For AiR kinase this loop could not be located, so it is hard to speculate whether it take a close structure or not. However large size of AiR may not allow the extent of closure we see in case of VcFK. During covering of sugar, some of the $\beta 6$ - $\beta 7$ loop residues form hydrogen bonds, although weak, with the $\alpha\beta$ domain thereby ensuring the covered state some sort of stability (figure 5.6b). Closure of lid upon sugar binding introduces huge surface charge redistribution. Comparison of apo and sugar-ADP bound structures in figures 5.7a, b and c illustrate this clearly. Figure 5.7d indicates

the involvement of residues like Glu107, Glu110, Glu177 and Glu178 towards negative charge formation. The phosphate groups of ATP, due to its flexibility, can adopt different conformations. Careful inspection of **figure 5.7b** and **d** it can be speculate that those

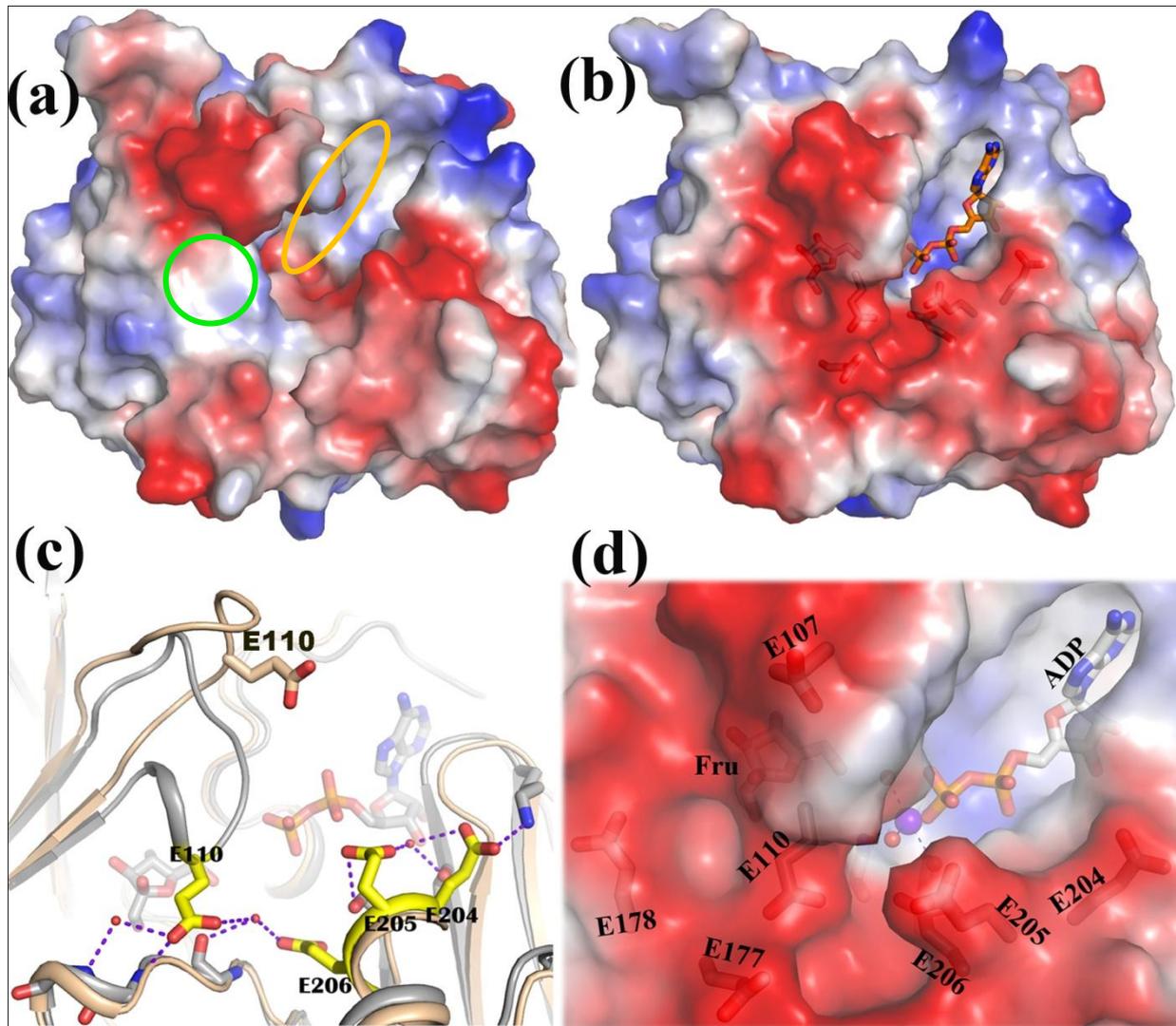


Figure 5.7: Electrostatic surface representation of VcFK in (a) apo form with fructose binding site (green sphere) and ATP binding site (yellow sphere) are marked, (b) sugar-ADP bound form in same orientation of apo form and redistribution of negatively charged surface is evident, (c) superposition of apo (wheat) and sugar-ADP bound (grey) form of VcFK around the active site where structural changes due to sugar binding and residues contributing towards negative charge are shown , (d) active site where negative charge surface and contributing glutamic acid residues are shown.

Conformations of ATP that are facing the negatively charged surface would be repelled and would not be selected, whereas those conformations that are directing towards fructose would be selected. This electrostatic entrapment of ATP plays a key role in the high fidelity of the phosphotransfer reaction. Sequence alignment of different FKs shows that these glutamic acids remain conserved in all fructokinases (highlighted by red in **figure 5.8b**). To explore the role of Glu110, we have mutated it to glutamine that has similar bulk as glutamic acid but neutral in nature, Therefore E110Q would still make hydrogen bond, although weaker, with the $\alpha\beta$ domain to close the lid but its repelling ability of the phosphate group will be absent. Kinetic data of E110Q showed more than 100 times decreased activity than the wild type.



Figure 5.8: Role of active site glutamic acid residues and their conservation throughout the superfamily. (a) Superposition of VcFK in apo form (grey) and sugar-ADP bound form (burgundy color). Side chains of Glutamic acid residues are shown with yellow sticks. (b) Sequence alignment of VcFK with other RK family members (PDB code 1TZ6- AiRs Kinase from *S. enterica*, PDB code 3LJS-fructokinase from, PDB code 3EWM- sugar kinase from *P. horikoshii*, PDB code 3Hj6-fructokinase from *H. orenii* [62]).

5.4 Formation of anion hole: stabilization of transition state

An aspartic acid residue is believed to act as a base to encourage deprotonation of the pertinent sugar hydroxyl group to be phosphorylated and thereby making it more nucleophilic to attack the γ -phosphate of ATP. So far there is no structural evidence about the position of γ -phosphate of ATP during the phosphotransfer. Atomic resolution structure of Fructose-ADP-BeF₃⁻ bound VcFK structure provides a structural evidence of this. During phosphate group transfer an anion hole formed by Gly265-Asp266-Ala267 where Asp266 acts as the activator [11]) (**figure 5.9b**). Here, an ADP + BeF₃⁻ together act as ATP analog and BeF₃⁻ mimics the γ -phosphate during phosphotransfer (discussed in **section 4.2.8**). The hydroxyl group of sugar is within 3Å from the beryllium atom of BeF₃⁻ indicating that it is poised to attack the ATP. Divalent cation Ca²⁺ binds between the β phosphate of ADP and BeF₃⁻ and forms a chelate that actually works as substrate [11]. Ca²⁺ ion helps to counteract the multiple negative charges of ADP and BeF₃⁻ from one side while the other sides of BeF₃⁻ face the anion hole favoring phosphotransfer.

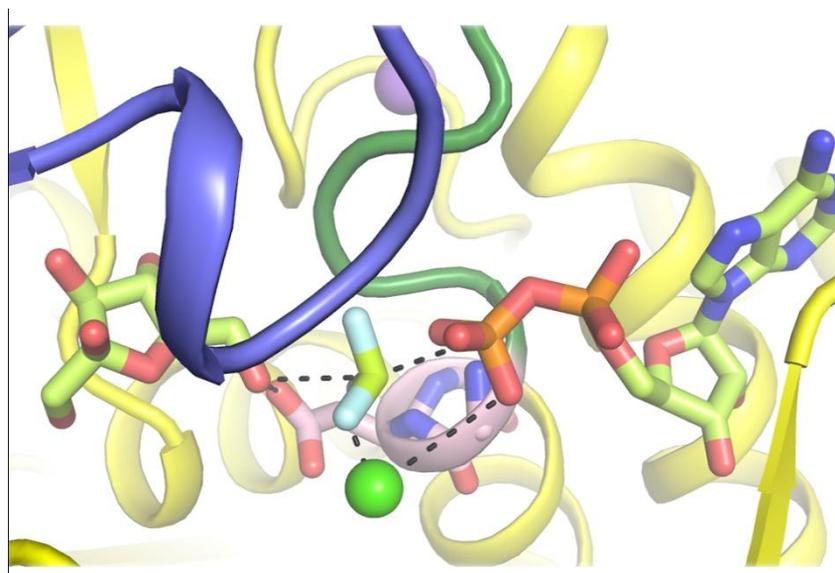


Figure 5.9: Stabilization of negative charge via anion hole. Active sites of VcFK in cartoon representation and main chains of anion hole-forming residues are shown in stick.

Therefore, structure of Vc-RK and Vc-FK provide details of substrate binding, structural alteration upon substrate binding and mechanism of monovalent cation induced activation. Kinetic data of VcFK also indicate a monovalent cation induced activation and involvements of divalent cations like Ca^{2+} and Mg^{2+} during phosphotransfer which support our structural data. The identification of key residues at the active site suggests chemical properties that must be present in candidate substrates so much so that it can distinguish between different anomers. It also suggests that the conformational changes that occur upon substrate binding in different families are concentrated in the substrate binding site and in the lid loop region and might be an advanced feature in the evolution of that superfamily. This also supports the hypothesis that the active site lid is a morphological marker for evolution within the superfamily. Comparison of the active site residues and charge atmosphere of the active site between different homologous proteins explain why a protein acts as a monomer, a dimer or a hexamer to smoothly perform the phosphorylation reaction. Further, most of the kinases of RK superfamily have high turnover number and low K_m value that characterizes the high fidelity of phosphotransfer reaction. Analysis of active site charge configuration and minimization of abortive phosphorylation through dimerization/oligomerization induced protection of ATP provides an answer to this.

REFERENCE

1. Nelson DL, Cox MM, Lehninger Principles Of Biochemistry, Fourth Edition (Whfreeman).
2. Ferrier RD, Lippincott's Illustrated Reviews: Biochemistry, Sixth Edition (Lippincott Williams & Wilkins).
3. Kim BH, Gadd GM, Bacterial Physiology and Metabolism, first Edition (2008) (Cambridge University Press) Pages: 60-61.
4. Mowbray SL. Ribose and glucose-galactose receptors. Competitors in bacterial chemotaxis. *J Mol Biol.* 1992 Sep 20; 227(2):418-40.
5. Agranoff BW and Brady RO. Purification and Properties of Calf Liver Ribokinase. *J Biol Chem.* 1956 Mar; 219(1):221-9.
6. Anderson A, Cooper RA. Biochemical and genetical studies on ribose catabolism in *Escherichia coli* K12. *J Gen Microbiol.* 1970 Aug;62(3):335-9.
7. Lopilato JE, Garwin JL, Emr SD, Silhavy TJ, Beckwith JR. D-ribose metabolism in *Escherichia coli* K-12: genetics, regulation, and transport. *J Bacteriol.* 1984 May; 158(2):665-73.
8. Chuvikovskiy DV, Esipov RS, Skoblov YS, Chupova LA, Muravyova TI, Miroshnikov AI, Lapinjoki S, Mikhailopulo IA. Ribokinase from *E. coli*: expression, purification, and substrate specificity. *Bioorg Med Chem.* 2006 Sep 15;14(18):6327-32.
9. Park J, Gupta RS. Adenosine kinase and ribokinase--the RK family of proteins. *Cell Mol Life Sci.* 2008 Sep; 65(18):2875-96.
10. Schimmel SD, Hoffee P, Horecker BL. Deoxyribokinase from *Salmonella typhimurium*. Purification and properties. *Arch Biochem Biophys.* 1974 Oct; 164(2):560-70.
11. Andersson CE, Mowbray SL. Activation of ribokinase by monovalent cations. *J Mol Biol.* 2002 Jan 18; 315(3):409-19.

12. Jecht EW, Wysocki R. Bacterial fructolysis in human seminal plasma. *Fertil Steril.* 1973 Apr; 24(4):314-8.
13. But SY, Rozova ON, Khmelenina VN, Reshetnikov AS, Trotsenko YA. Properties of recombinant ATP-dependent fructokinase from the halotolerant methanotroph *Methylomicrobium alcaliphilum* 20Z. *Biochemistry (Mosc).* 2012 Apr; 77(4):372-7.
14. Raushel FM, Cleland WW. The substrate and anomeric specificity of fructokinase. *J Biol Chem.* 1973 Dec 10; 248(23):8174-7.
15. Sigrell JA, Cameron AD, Jones TA, Mowbray SL. Structure of *Escherichia coli* ribokinase in complex with ribose and dinucleotide determined to 1.8 Å resolution: insights into a new family of kinase structures. *Structure.* 1998 Feb 15; 6(2):183-93.
16. Bork P, Sander C, Valencia A. Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. *Protein Sci.* 1993 Jan; 2(1):31-40.
17. Cheek S, Zhang H, Grishin NV. Sequence and structure classification of kinases. *J Mol Biol.* 2002 Jul 19; 320(4):855-81.
18. Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247, 536- 540.
19. Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Windells, M. B. & Thornton, J. M. (1997). CATH- a hierarchic classification of protein domain structures. *Structure*, 5, 1093- 1108.
20. Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Howe, K. L. & Sonnhammer, E. L. (2000). The Pfam protein families database. *Nucl. Acids Res.* 28, 263- 266.
21. Cheek S, Ginalski K, Zhang H, Grishin NV. A comprehensive update of the sequence and structure classification of kinases. *BMC Struct Biol.* 2005 Mar 16;5:6.
22. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* 25, 3389- 3402.

23. Rose RK, Turner SJ. Extracellular volume in streptococcal model biofilms: effects of pH, calcium and fluoride. *Biochim Biophys Acta*. 1998 Feb 2; 1379(2):185-90.
24. Houot L, Chang S, Absalon C, Watnick PI. *Vibrio cholerae* phosphoenolpyruvate-phosphotransferase system control of carbohydrate transport, biofilm formation and colonization of the germfree mouse intestine. *Infect Immun*. 2010 Apr; 78(4):1482-94.
25. Kendrew JC. The structure of globular proteins. *Comp Biochem Physiol*. 1962 Oct; 4:249-52.
26. Wlodawer A, Minor W, Dauter Z, Jaskolski M. Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *FEBS J*. 2013 Nov; 280(22):5705-36.
27. Hauptman H. Phasing methods for protein crystallography. *Curr Opin Struct Biol*. 1997 Oct;7(5):672-80.
28. Langs DA, Hauptman HA. Relaxation of the resolution requirements for direct-methods phasing. *Acta Crystallogr A*. 2011 Jul;67(Pt 4):396-401.
29. Abergel C. Molecular replacement: tricks and treats. *Acta Crystallogr D Biol Crystallogr*. 2013 Nov;69(Pt 11):2167-73.
30. Georgopoulou AN, Adam R, Raptopoulou CP, Psycharis V, Ballesteros R, Abarca B, Boudalis AK. Isomorphous replacement of M(II) ions in M(II)-Gd(III) dimers (M(II) = Cu(II), Mn(II), Ni(II), Co(II), Zn(II)): magnetic studies of the products. *Dalton Trans*. 2010 May 28;39(20):5020-7.
31. Dauter Z, Dauter M, Dodson E. Jolly SAD. *Acta Crystallogr D Biol Crystallogr*. 2002 Mar;58(Pt 3):494-506. Epub 2002 Feb 21.
32. Terwilliger TC. MAD phasing: treatment of dispersive differences as isomorphous replacement information. *Acta Crystallogr D Biol Crystallogr*. 1994 Jan 1;50(Pt 1):17-23.
33. Brünger AT. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*. 1992 Jan 30;355(6359):472-5.

34. Mullis KB. Target amplification for DNA analysis by the polymerase chain reaction. *Ann Biol Clin (Paris)*. 1990;48(8):579-82.
35. Nelson RM, Long GL. A general method of site-specific mutagenesis using a modification of the *Thermus aquaticus* polymerase chain reaction. *Anal Biochem*. 1989 Jul;180(1):147-51.
36. Birnboim HC, Doly J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res*. 1979 Nov 24;7(6):1513-23.
37. McPherson A. Introduction to protein crystallization. *Methods*. 2004 Nov; 34(3):254-65.
38. Russo Krauss I, Merlino A, Vergara A, Sica F. An overview of biological macromolecule crystallization. *Int J Mol Sci*. 2013 May 31;14(6):11643-91.
39. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr*. 2011 Apr;67(Pt 4):271-81.
40. Evans P. Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr*. 2006 Jan;62(Pt 1):72-82.
41. Rossmann MG, van Beek CG. Data processing. *Acta Crystallogr D Biol Crystallogr*. 1999 Oct;55(Pt 10):1631-40.
42. Winkler FK, Schutt CE and Harrison SC. The oscillation method for crystals with very large unit cells. *Acta Cryst*. 1979. A35, 901-911.
43. Leslie AG. The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr*. 2006 Jan;62(Pt 1):48-57.
44. Vagin AA, Isupov MN. Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr D Biol Crystallogr*. 2001 Oct;57(Pt 10):1451-6.
45. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007 Aug 1;40(Pt 4):658-674.

46. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr*. 2011 Apr;67(Pt 4):235-42.
47. Read RJ, McCoy AJ, Storoni LC. Likelihood-based molecular replacement in phaser. *Evolving Methods for Macromolecular Crystallography*. 2007 Volume 245 pp 91-100.
48. Adams PD, Afonine PV, Bunkóczy G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. 2010 Feb;66(Pt 2):213-21.
49. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*. 1997 May1;53(Pt 3):240-55.
50. Wiegels T, Lamzin VS. Use of noncrystallographic symmetry for automated model building at medium to low resolution. *Acta Crystallogr D Biol Crystallogr*. 2012Apr;68(Pt 4):446-53.
51. Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* 62:439–450.
52. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
53. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774–797.
54. Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* 38:695–699.

55. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R (2013) Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res* 41:597–600.
56. Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 1988 Nov 25;16(22):10881-90.
57. The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC.
58. Martinez-Barajas E, Krohn BM, Stark DM, Randall DD. Purification and characterization of recombinant tomato fruit (*Lycopersicon esculentum* Mill.) fructokinase expressed in *Escherichia coli*. *Protein Expr Purif.* 1997 Oct;11(1):41-6.
59. Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* 24, 409–411.
60. Sigrell JA, Cameron AD, Jones TA, Mowbray SL (1999) Induced fit on sugar binding activates ribokinase. *J Mol Biol* 290:1009–1018.
61. Zhang Y, Dougherty M, Downs DM, Ealick SE. Crystal structure of an aminoimidazole riboside kinase from *Salmonella enterica*: implications for the evolution of the ribokinase superfamily. *Structure.* 2004 Oct;12(10):1809-21.
62. Chua TK, Seetharaman J, Kasprzak JM, Ng C, Patel BK, Love C, Bujnicki JM, Sivaraman J. Crystal structure of a fructokinase homolog from *Halothermothrix orenii*. *J Struct Biol.* 2010 Sep;171(3):397-401.
63. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr.* 1994 Sep 1;50(Pt 5):760-3.
64. Potterton E, Briggs P, Turkenburg M, Dodson E (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr* 59:1131–1137.
65. Matthews BW. Solvent content of protein crystals. *J Mol Biol.* 1968 Apr;33(2):491-7.
66. Tina KG, Bhadra R, Srinivasan N. PIC: Protein Interactions Calculator. *Nucleic Acids Res.* 2007 Jul;35(Web Server issue):W473-6.

67. Quesada-Moreno MM, Azofra LM, Avilés-Moreno JR, Alkorta I, Elguero J, López-González JJ. Conformational preference and chiroptical response of carbohydrates D-ribose and 2-deoxy-D-ribose in aqueous and solid phases. *J Phys Chem B*. 2013 Nov 27;117(47):14599-614.
68. Nowak T, Suelter C. Pyruvate kinase: activation by and catalytic role of the monovalent and divalent cations. *Mol Cell Biochem*. 1981 Mar 13;35(2):65-75.
69. Li J, Wang C, Wu Y, Wu M, Wang L, Wang Y, Zang J. Crystal structure of Sa239 reveals the structural basis for the activation of ribokinase by monovalent cations. *J Struct Biol*. 2012 Feb;177(2):578-82.
70. Park J, van Koeverden P, Singh B, Gupta RS. Identification and characterization of human ribokinase and comparison of its properties with *E. coli* ribokinase and human adenosine kinase. *FEBS Lett*. 2007 Jul 10;581(17):3211-6.
71. Xu YW, Moréra S, Janin J, Cherfils J. AIF3 mimics the transition state of protein phosphorylation in the crystal structure of nucleoside diphosphate kinase and MgADP. *Proc Natl Acad Sci U S A*. 1997 Apr 15;94(8):3579-83.
72. Hughes Ed. The Walden inversion. *Sci Prog*. 1946 Jul;34:516-32.
73. Mathews II, Erion MD, Ealick SE. Structure of human adenosine kinase at 1.5 Å resolution. *Biochemistry*. 1998 Nov 10;37(45):15607-20.
74. Cabrera R, Ambrosio AL, Garratt RC, Guixé V, Babul J. Crystallographic structure of phosphofructokinase-2 from *Escherichia coli* in complex with two ATP molecules. Implications for substrate inhibition. *J Mol Biol*. 2008 Nov 14;383(3):588-602
75. Ohshima N, Inagaki E, Yasuike K, Takio K, Tahirov TH. Structure of *Thermus thermophilus* 2-Keto-3-deoxygluconate kinase: evidence for recognition of an open chain substrate. *J Mol Biol*. 2004 Jul 9;340(3):477-89.
76. Gibbs AC, Abad MC, Zhang X, Tounge BA, Lewandowski FA, Struble GT, Sun W, Sui Z, Kuo LC. Electron density guided fragment-based lead discovery of ketohexokinase inhibitors. *J Med Chem*. 2010 Nov 25;53(22):7979-91.