Metal Fluorides as Probes for Enzyme Catalysed Phosphoryl Transfer

Thesis Submitted for the Degree of

Doctor of Philosophy

Yi Jin BSc MSc

Department of Molecular Biology and Biotechnology

The University Of Sheffield

January 2012

Abstract

The work described in this thesis is based on the use of multinuclear NMR and protein crystallography as two powerful and symbiotic biological structural tools to study the catalytic mechanisms of five phosphoryl transfer enzymes: β-phosphoglucomutase, phosphoserine phosphatase, cAMP-dependent protein kinase, small G protein RhoA-RhoGAP, and UMP/CMP kinase. The central feature is the use of magnesium and aluminium fluorides as transition state analogues for phosphoryl group transfer.

β-Phosphoglucomutase is first examined using stable βG1P phosphonate analogues of which (*S*)-βCHFG1P and βCH_2G1P form MgF_3^- and AlF_4^- TSA complexes with βPGM_{wt}. Next, mutant studies show D10N and D8E form AlF_4^- complexes while D8N does not. Only D10N forms a fluoroberyllate complex on the nucleophilic Asp 8, which can also be phosphorylated by acetyl phosphate.

The general acid-base role of D13 in PSP is further established by showing the inactivity of PSP_{D13N}, a mutant that forms MgF_3^- and AlF_4^- TSA complexes with water but not serine. Fundamental functional differences between PSP and βPGM are rationalised.

cAPK forms MgF_3^- and AlF_4^- TSA complexes, disproving the 'AlF_3^0' assignment in PDB: 1L3R. The binding constant of cAPK for its MgF_3^- TSA is only ~10 fold weaker than for the AlF_4^- TSA complex. This suggests that fluoride inhibition of a wide range of signalling proteins may be dominated by MgF_3^- rather than by the more widely studied AlF_4^- with significant physiological implications.

^{19}F NMR data confirm the RhoA-GDP-MgF_3^--OH_2-RhoGAP TSA solid state structure, and identify an AlF_4^- TSA complex that obeys the Charge Balance Hypothesis. A GAP-free octahedral RhoA-GDP-$AlF_3(H_2O)^0$ complex is formed that also conforms to the CBH. NMR and crystallography show RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ and RhoA-GDP-AlF_4^--$RhoGAP_{R85A}$ TSA complexes maintain negative charge on the metal fluoride moieties *via* Tyr 34 coodination. Finally, substrate analogue GTPγF is remarkably stable to hydrolytic acitivity by RhoA alone, or with RhoGAP with or minus its arginine finger.

^{19}F NMR data disprove the earlier assignment of tbp 'AlF_3^0' in UMP/CMP kinase structures crystallised at pH 8.5 and show MgF_3^- is the actual TSA at high pH. NMR/pH titration reveals a new intermediate aluminium TSA species, AlF_3Z, not hitherto identified. NMR analysis endorses earlier solid state observations on a UMPK-ADP-BeF_2^0-UDP TSA complex and demonstrates the catalytic promiscuity of $UMPK_{dicty}$. AMPCP is a stable substitute for ADP in TSA complex formation while $AMPCF_2P$ or ADPβF fail unexpectedly. Similarly, UMP cannot be replaced by 5FdUMP.

This thesis is a prime example of the symbiotic use of ^{19}F NMR and protein crystallography to study enzyme mechanisms. ^{19}F NMR can identify TSA complex formation in solution and hence can direct crystal trials. Crystal structures are more accurately assigned using NMR data while, in turn, they give further support to detailed interpretation of NMR signals.

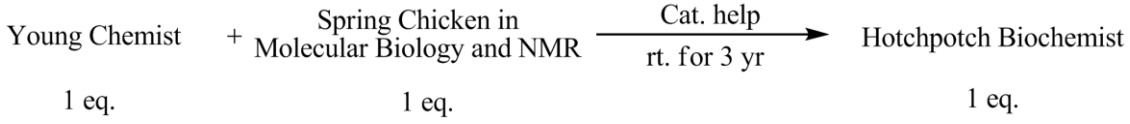
Acknowledgements

I want to first express my gratitude to my supervisor Prof. Jon Waltho who gave me the opportunity to come to Sheffield and guided me through unknown territory. I am indebted to Prof. David Hornby, Head of the Molecular Biology, for underpinning all my work with the resources of the Department. I have been the grateful recipient of an Overseas Research Students Awards Scheme (ORSAS) and University of Sheffield Studentship. The process of being initiated in molecular biology would not have been possible without the support and endless efforts of Dr. Qaiser Sheikh, Dr. James Marston and Dr. Joanna Griffin. Dr. Liu Xiaoxia's company made all my out-of-hour working time not lonely. I would also like to thank Dr. Andrea Hounslow for assisting with the running of the NMR experiments and Dr. Matthew Cliff and Dr. Nicky Baxter for invaluable discussion and suggestions towards my research projects. The collaboration with Dr. Matthew Bowler and Erika Pellegrini in ESRF, Grenoble has provided invaluable information on crystal structures complementary to my NMR work over the last three years. I am also grateful to colleagues in Prof. Chris Hunter's lab (Department of Chemistry) who gave me all their support for my synthetic work in four happy months.

Moreover, I could not have finished my PhD study without the moral support from my dearly parents in China. Their courage and positiveness always encourages me to face my ups and downs. My thanks are also due to my boyfriend Tudor for his enthusiasm and patience in the past whole year while I completed my studies.

I also must express my appreciation to all the people who have been answering HOW, WHAT and WHY, not only in my work but also in my life.

Lastly, but by no means the least, my respect and appreciation goes to Prof. Mike Blackburn, who has shared the last three and a half years of my life, inspired my exploration, taught me to speak English and dragged me through hell and high water.



Yi Jin
Sale Hill, Sheffield
January 2012

Abbreviations

APS	ammonium persulfate
acetyl-P	acetyl phosphate
βPGM	β-phosphoglucomutase
cAPK	cyclic-AMP-dependent protein kinase
CBH	charge balance hypothesis
1D	one dimension(al)
2D	two dimensions(al)
D10N	Asp to Asn single point mutation at amino acid-10 in βPGM
D13N	Asp to Asn single point mutation at amino acid-13 in PSP
DFO	deferoxamine
ddH_2O	deionised and distilled water
dNTs	deoxynucleotides
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
G6P	D-glucose 6-phosphate
GA	general acid
GAB	General acid base
GABC	general acid base catalysis
GAP	GTPase activating protein
GB	general base
GdnHCl	guanidine hydrochloride
GEF	guanidine nucleotide exchange factor
GS	ground state
GSA	ground state analogue
HAD	haloacid dehalogenase
HsPGK	human phosphoglycerate kinase
HSQC	Heteronuclear Single Quantum Coherence
IPTG	isopropyl β-D-1-thiogalactopyranoside
MWCO	molecular weight cut off
NMR	nuclear magnetic resonance
OD_{600}	optical density at 600 nm
pAsp	aspartyl phosphate/phosphorylated aspartate
PCR	Polymerase Chain Reaction
P_i	inorganic phosphate
PPi	pyrophosphate
PIXE	proton-induced X-ray emission spectroscopy
3PG	3-phosphoglycerate
PKA	protein kinase A (also known as cAPK, *v.s.*)
PLS	O^2-phospho-L-serine
PSA	product state analogue
PSP	phosphoserine phosphatase
rpm	revolutions per min
rt	room temperature
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
TAE	Tris-Acetate-EDTA buffer
TBE	Tris-Boric Acid-EDTA buffer
tbp	trigonal bipyramid(al)
TEMED	*N*,*N*,*N'*,*N'*-tetramethylethylenediamine
TS	transition state
TSA	transition state analogue
UMPK	UMP/CMP kinase

Contents

List of Figures

List of Tables

Chapter 1. Introduction

From a slow reaction like phosphate hydrolysis to fast radical reactions like redox, the rates of such uncatalysed reactions have a range of at least nineteen orders of magnitude. By contrast, enzyme-catalysed reaction rates fall into a relatively narrow range.(Lad *et al.*, 2003) Understanding what makes these chemical reactions take place in such an efficient way in a biocatalytic system is a fundamental goal of life science. These biological catalysts are called enzymes which are mostly proteins and can also be ribozymes.

1.1 Theories of Enzyme Mechanisms

The first notable theory which reflects the interaction between substrate and enzyme is the "The Lock and Key" model, proposed by Fisher in 1894.(Fischer, 1894) In the 1930s, Haldane established the theory of strain: "Using Fischer's lock and key simile, the key does not fit the lock perfectly but exercises a certain strain on it".(Haldane, 1930) This strain theory is thought to be inadequate to describe enzyme catalysis.(Britt, 1997) In the late 1940s, Pauling suggested that enzymes are molecules that are complementary in structure to the activated complex of the reactions that they catalyse.(Pauling, 1948) In order for the reaction to occur, reactants (which are usually named substrates) have to gain sufficient energy to cross a potential energy barrier, the activation energy $\Delta G^{\ddagger}$. The lower the potential energy barrier to reaction and the more energy reactants have, the faster the reaction occurs.(**Figure 1.1**) Pauling's description used the term " activated complex" which was only later described as the "transition state (TS)".

One view of how enzymes lower the activation energy is that when the substrate binds to the enzyme, the enzyme adapts its conformation to accommodate the substrate and this alters the local surroundings and orientates catalytic residues into alignment to make the reaction easier.(Johnson, 2008) In the subsequent catalytic reaction, the transition state, a concept accepted by scientists for many years, should be formed.(Laidler, 1983)

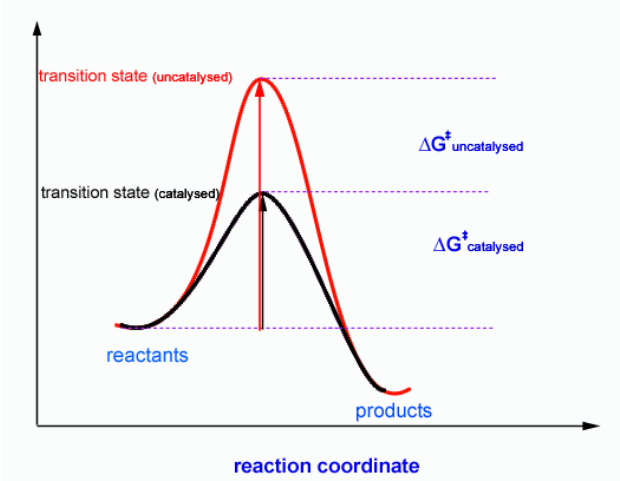


Figure 1.1 Energy profile diagram of a reaction involving the transition state.

The transition state (TS) is a high energy, transient state formed in the conversion of reactants into products. The TS in **Figure 1.2** is the maximum energy point along the reaction path in a one-step chemical reaction and the activation energy ΔG_a^* is the difference between the energy of TS and reactants. (**Figure 1.2a**) The lower the ΔG_a^*, the faster the reaction. The larger is the net free energy change ΔG_n^*, the more irreversible is the reaction. In a two-step reaction, an intermediate is formed in conversion of reactants into products. The TS remains the point of maximum energy along the reaction path (**Figure 1.2b**).

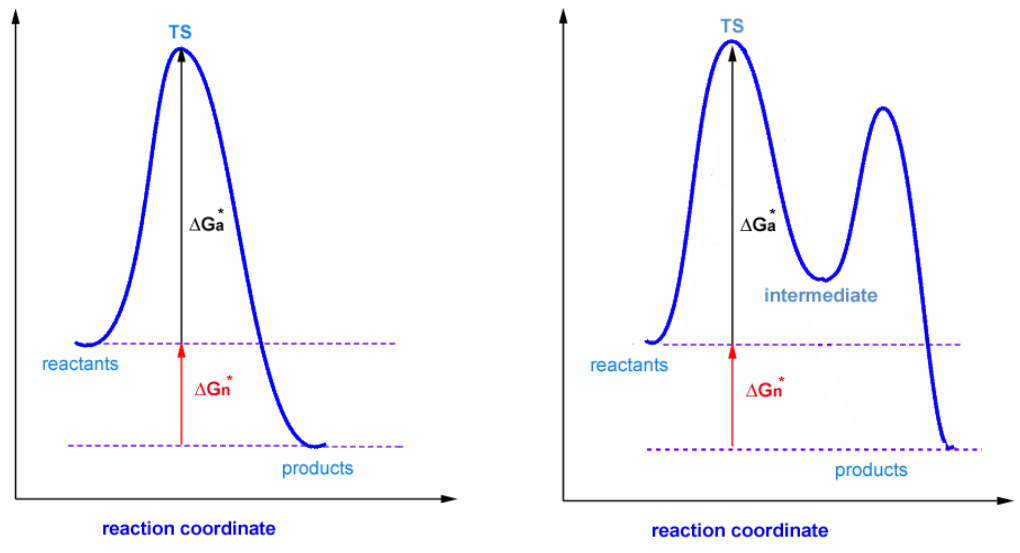


Figure 1.2 Energy profile diagrams for (**a**) one-step and (**b**) two-step reactions.

Transition state theory defines catalysis as preferentially stabilising the TS relative to the ground state (GS), because an enzyme cannot change the free energy of the unbound

reactants and the products. After the whole catalytic process, the enzymes are not consumed and the equilibrium of the reaction stays the same. Catalysis can be achieved by a combination of interrelated effects, such as electrostatic effect, entropy effect (including desolvation), concentration effect, orbital steering, strain effect, *etc.* The contribution of each effect cannot be quantified and is subject to continuous debate.(Fersht, 1985) Therefore, to understand how a proficient enzyme is capable of recognising and stabilising the TS, it is essential to characterise the charge distribution, ionic bonding and the geometry of the transition state.(Lassila *et al.*, 2010)

1.2 Phosphotransfer

Phosphotransfer reactions are central reactions in intermediary metabolism, signal communication, nucleic acid biosynthesis and processing, regulation of many cellular processes including replication, cellular development and apoptosis, and efficient high-energy transformation.(Dzeja *et al.*, 2003) Many enzymes catalyse phosphotransfer between phosphate diesters and phosphate monoesters.(Knowles, 1980) The mechanisms of these biological reactions can be grouped into two classes: the phosphoryl transfer (PO_3^-) and nucleotidyl or alkylphosphoryl group transfer ($ROPO_2^-$) as shown in **Figure 1.3**.(Frey *et al.*, 2007)

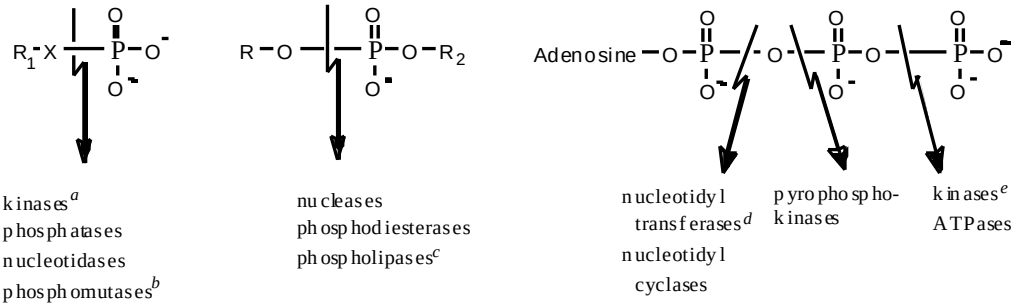


Figure 1.3 Classes of enzyme that catalyse phosphotransfer reactions. ***a***, R_1=H, COOH, PO_4^{3-} when X is O, or RX is guanidine group; ***b***, the phosphoryl group is transferred intramolecularly between two hydroxyl groups or between a hydroxyl and a carboxyl group; ***c***, phospholipase C and D; ***d***, includes ligases such as aminoacyl tRNA synthetases; ***e***, includes ligases such as glutamine synthetase and acetyl-CoA carboxylase.

Enzymes like GTPases, ATPases, kinases and phosphatases are doing a fantastic job in catalysing reactions of this nature.(Thompson *et al.*, 2001) For instance, without the catalysis of phosphatase, the half-life of the hydrolysis of a phosphate monoester dianion may be 1.1×10^{12} years (**Figure 1.4**)! However, the catalysed transfer reaction can be so efficient that k_{cat} is 10^{24} times faster ($k_{cat} = 1.5 \times 10^4$ s^{-1}) in the presence of protein tyrosine phosphatase (PTPase; EC 3.1.3.48).(Lad *et al.*, 2003)

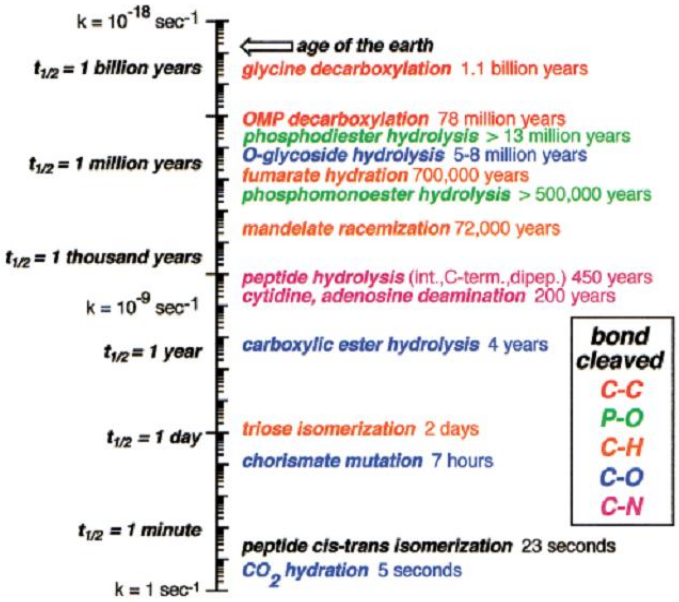


Figure 1.4 Natural half-lives of some biologically relevant reactions in neutral solution at 25°C.(Wolfenden *et al.*, 2001)

A recent review of biological phosphoryl transfer mechanisms has placed strong emphasis on four central questions needing mechanistic answers.(Lassila *et al.* 2011)

1. Are phosphoryl transfer reactions concerted or do they proceed through stable intermediates?
2. How can the structure and properties of the TS be evaluated?
3. Do enzymes manipulate the TS to make it different from those in solution?
4. How is phosphoryl transfer achieved by enzymatic catalysis?

These matters will be looked at in the sections that follow.

1.2.1 Model studies in solution

Much progress has been made by chemists on understanding the details of the mechanism of phosphoryl transfer for small organic molecules. Two extreme routes have been established to describe nucleophilic substitution at phosphorus (**Figure 1.5**). On one hand, in a purely dissociative mechanism the bond between the leaving group and phosphorus breaks first with the negative charge moving from the leaving group onto the phosphoryl group and a trigonal metaphosphate intermediate is formed in an S_N1_P (elimination-addition, D_N+A_N in the IUPAC nomenclature) process. At the other extreme, in a wholly associative mechanism, a five-coordinate, trigonal bipyramidal (tbp) pentaoxyphosphorane-like species is formed as an intermediate with the negative

charge accumulating on the leaving group in an S_N2_P (addition-elimination, A_N+D_N in the IUPAC nomenclature) process.

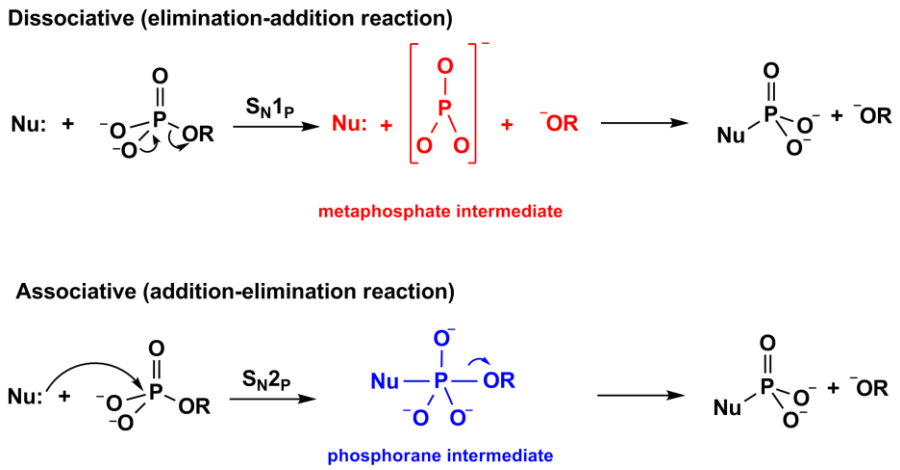
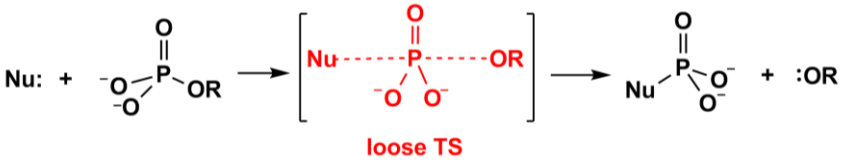


Figure 1.5 Proposed extreme mechanisms of phosphoryl transfer.

1.2.2 Phosphomonoesters

Phosphomonoesters, phosphoanhydrides, and phosphoramidates undergo phosphoryl transfer by P–O and P–N bond cleavage.



There is a general consensus that uncatalysed reactions involving phosphomonoesters with good leaving groups undergo an S_N1_P elimination mechanism *via* a trigonal metaphosphate loose TS as shown above.(Mildvan, 1997; Admiraal *et al.*, 2000; Thompson *et al.*, 2001; Webster, 2004; Cleland *et al.*, 2006; Lassila *et al.*, 2010) A typical phosphoryl group transfer mechanism is one in which the phosphoryl donor and acceptor interact weakly with the phosphoryl group in flight in a TS in which the total bonding to phosphorus is decreased relative to the ground state. Detailed studies indicate that the bonding highlighted by the dash lines in the above diagram represents partial covalency on the order of 10% to 20% of the strength of a full covalent bond, or a bond order of 0.1 to 0.2.(Frey *et al.*, 2007)

Phosphate monoester reactions play important roles in nucleotide synthesis and energy regulation, as well as transport of ions and small molecules. Most significantly, the

work of Krebs and Fisher (Nobel Prize in Physiology or Medicine, 1992) established the role of protein kinases in the activation and regulation of a majority of cellular enzymes, closely linked to the action of protein phosphatases in hydrolysing phosphate monoesters of the amino acids serine, threonine, and tyrosine.(Graves, 1999) The range and variety of such kinases is now organised in the human kinome, that is now recognised as having over 500 members (**Figure 1.6**).(Manning *et al.*, 2002) The vast majority of these kinases involve the transfer of the γ-phosphate of ATP to the hydroxyl group of an amino acid. So these kinases have become grouped into the Ser/Thr kinases that transfer phosphate to the hydroxyl group of serine or threonine (Edelman *et al.*, 1987; Hanks *et al.*, 1988) and the tyrosine kinases, that transfer phosphate to the phenolic hydroxyl group of tyrosine.(Hunter *et al.*, 1985) Some kinases show dual specificity.(Manning *et al.*, 2002)

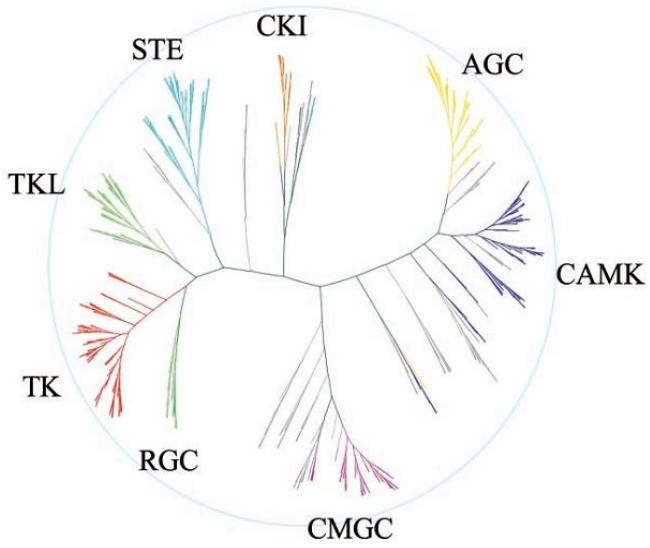
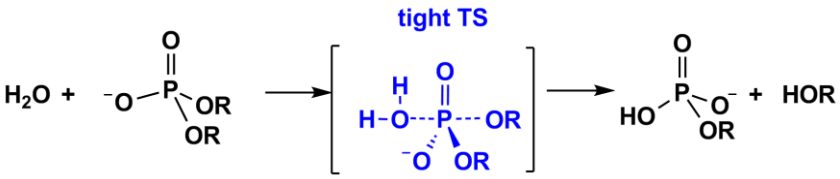


Figure 1.6 Diagram of 491 eukaryotic protein kinase domains from 478 genes.(Manning *et al.*, 2002)

1.2.3 Phosphodiesters

The mechanistic properties of P–O bond cleavage in reactions of phosphodiesters are unlike those in reactions of phosphomonoesters. The hydrolysis of phosphodiester monoanions is very much slower ($k_{monoester}/k_{diester} \approx 10^4$) for the same R groups. The rate of hydrolysis is more sensitive to the stability of the leaving group than in phosphomonoesters. The more stable the leaving group the faster it hydrolyses. The rate is also sensitive to the reactivity of the nucleophile. The kinetic isotope effect (see **Section 1.7**) of the hydrolysis shows the reaction undergoes a tight TS with proton transfer which means the bond orders between phosphorus and both the leaving group

and acceptor are about 0.5, and the net bond order to phosphorus in the TS is equal to or larger than in the GS.(Frey *et al.*, 2007)



Phosphodiester substitution reactions that are catalysed by nucleases and polymerases are essential to DNA replication and transcription.

1.2.4 Phosphotriesters

Phosphotriesters are far more reactive toward nucleophiles than phosphodiester monoanions due to the absence of negative charge. They are especially reactive toward negatively-charged nucleophiles such as OH^- so that it is 2000 times as reactive as the phosphodiester monoanions. The TS for the reactions of phosphotriesters are very tight.(Benkovic, *et al.*, 1973) However, reactions of phosphotriesters are of lesser importance in biology, apart from clearance of insecticides and some DNA damage repair,(Caldwell *et al.*, 1991) and are not described further in this thesis.

1.3 Stereochemistry of Phosphoryl Transfer

The early understanding of the mechanism of the phosphoryl transfer in kinases was based on kinetic studies showed that phosphoryl transfer takes place only when both substrates are bound. But to know the details of the interconversion of the ternary complex, the stereochemistry of phosphoryl transfer needs to be known.

1.3.1 Early work of using ATPγSγ^{18}O

In the early stages, valuable investigations on the nature of phosphoryl transfer came from stereochemical studies. Adenosine 5'-*O*-([γ-^{18}*O*],γ-thio)triphosphate (ATPγSγ^{18}O) was first synthesised enzymically using 2-[^{18}*O*]-D-glycerate phosphorothioate and ADP by enolase and pyruvate kinase. This 'double' labelled ATP was then used to thiophosphorylate glycerol catalysed by glycerol kinase (EC 2.7.1.30) and hexose catalysed by hexokinase (EC 2.7.1.1). The relative configuration of the product glycerol 3-[^{18}*O*]-phosphorothioate was identified by chemical treatment to produce diastereoisomers followed by separation by chromatography. The ^{18}O content was then

measured by GC/MS. The results showed that the stereochemistry course of the reaction at phosphorus undergoes an inversion without losing stereospecificity and this indicated the direct phosphoryl transfer from ATP to acceptor substrate.(Orr *et al.*, 1978)

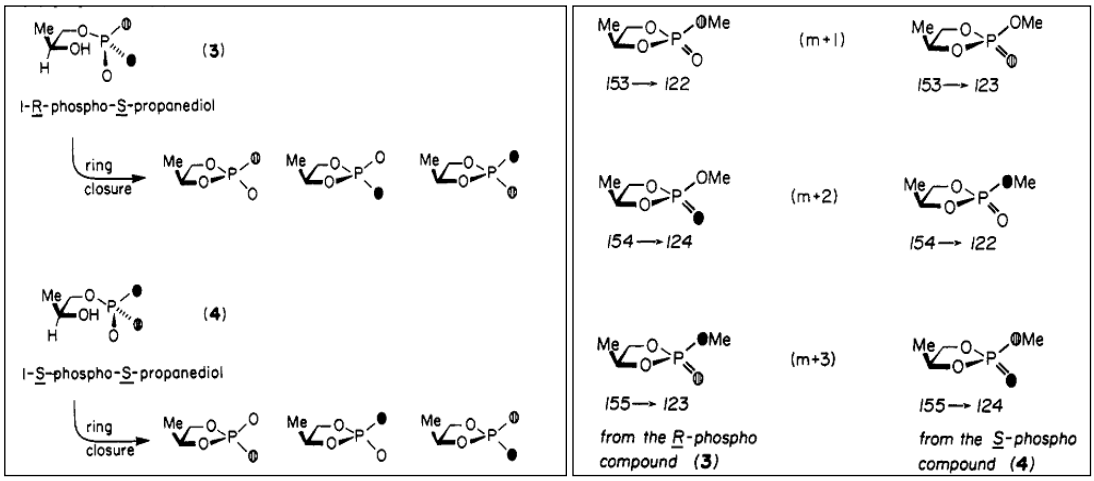
1.3.2 Later work using [γ-(*S*)-^{16}O,^{17}O,^{18}O]ATP

Later, the stereochemistry at phosphorus was proved independently of the use of sulfur as a chirality-creating substituent by using chemically synthesised adenosine [γ-^{16}*O*,^{17}*O*,^{18}*O*]-triphosphate which was able to define the transfer of the phosphoryl group rather than of the thiophosphoryl group.(Jones *et al.*, 1978; Blattler *et al.*, 1979) Two proofing methods of the configuration of the γ-phosphate were developed independently almost at the same time were developed side-by-side: one built around mass spectrometry and the second around NMR.(Jones *et al.*, 1978; Blattler *et al.*, 1979; Lowe *et al.*, 1981)

Mass spectrometry method

For mass spectrometry, the analysis was based on the knowledge that γ-phosphoryl group of ATP can be transferred to (*S*)-propane-1,2-diol without inversion of configuration using alkaline phosphatase. The mixture of phosphopropanediol (**3** or **4**) produced can then be cyclised to the 1,2-cyclic phosphate which was methylated and then subjected to stereochemical analysis developed for mass spec purposes (**Figure 1.7**). For example, the *m/z* 155 (m+3, with ^{17}O and ^{18}O on phosphorus) metastable mother ion of cyclic triester of 1(*R*)-phospho-(*S*)-propanediol will lose a ^{18}O-formaldehyde to give a daughter ion at *m/z* 123 in linked scan MS (**Figure 1.7b**, left column). But the metastable mother ion of cyclic triester of 1(*S*)-phospho-(*S*)-propanediol with opposite steoreochemistry on phosphorus will instead lose a ^{17}*O*-formaldehyde to give a daughter ion at *m/z* 124 (**Figure 1.7b**, right column).(Abbott *et al.*, 1978) The results using this method confirmed that phosphoryl transfer reactions for glycerol kinase, hexokinase, pyruvate kinase, and especially acetate kinase, for which various transfer mechanisms had been present, all go with inversion at phosphorus. Therefore the phosphoryl group is transferred in-line directly from ATP to

the other enzyme-bound substrate without going through phosphoenzyme intermediates.(Blattler *et al.*, 1979)



a **b**

Figure 1.7 (**a**) two sets of epimers derived by 'in-line' ring closure of diastereoisomeric phosphopropanediols. (**b**) After methylation, metastable ion (linked scan) species that comprise the 'syn' cyclic triesters show different daughter ions for 1-(*R*)-phospho-(*S*)-propanediol (left) or for 1-(*S*)-phospho-(*S*)-propanediol (right).(Abbott *et al.*, 1978)

^{31}P NMR method

The NMR method is based on ^{31}P NMR analysis on the special nuclear magnetic property of ^{17}O and ^{18}O on phosphate. First, the advantage of ^{17}O is, as a relaxation factor with 5/2 spin, when it is directly bonded to phosphorus because of its nuclear electric quadrupole moment that phosphorus does not show a signal in ^{31}P NMR spectra.(Lowe *et al.*, 1979) Second, the isotope shift of phosphorus caused by the direct attachment to a ^{18}O is shifted upfield by about 2 ppm (at 250 MHz). It also depends on the nature of the bond to phosphorus, involving bond order and orientation. The isotope shift of a P=O is bigger than of a P–O. Based on these facts, after phosphorylation using [γ-(*S*)-^{16}O,^{17}O,^{18}O]ATP, the chiral [^{16}O,^{17}O,^{18}O]-phosphate monoester product was chemically converted into two diastereoisomeric, conformationally-locked phosphate triesters as fused six-membered rings (**Figure 1.8**, first step).(Lowe *et al.*, 1981) During this procedure, there is equal probability for any one of the three peripheral oxygen isotopes to be lost (ignoring any possible kinetic isotope effects). Similar to the method described above for mass spectrometry, methylation of the cyclic phosphate diester was carried out to give the diastereoisomeric axial and equatorial triesters (**Figure 1.8**). The chemical shifts and peak intensities in ^{31}P NMR spectra of these triesters **1** and **4** (or **7** and **10)** were compared with the control spectra of D-

glucose 6-[(*S*)-^{16}O,^{17}O,^{18}O]-phosphatetriester in which the absolute configuration at phosphorus is known.(Lowe *et al.*, 1981)

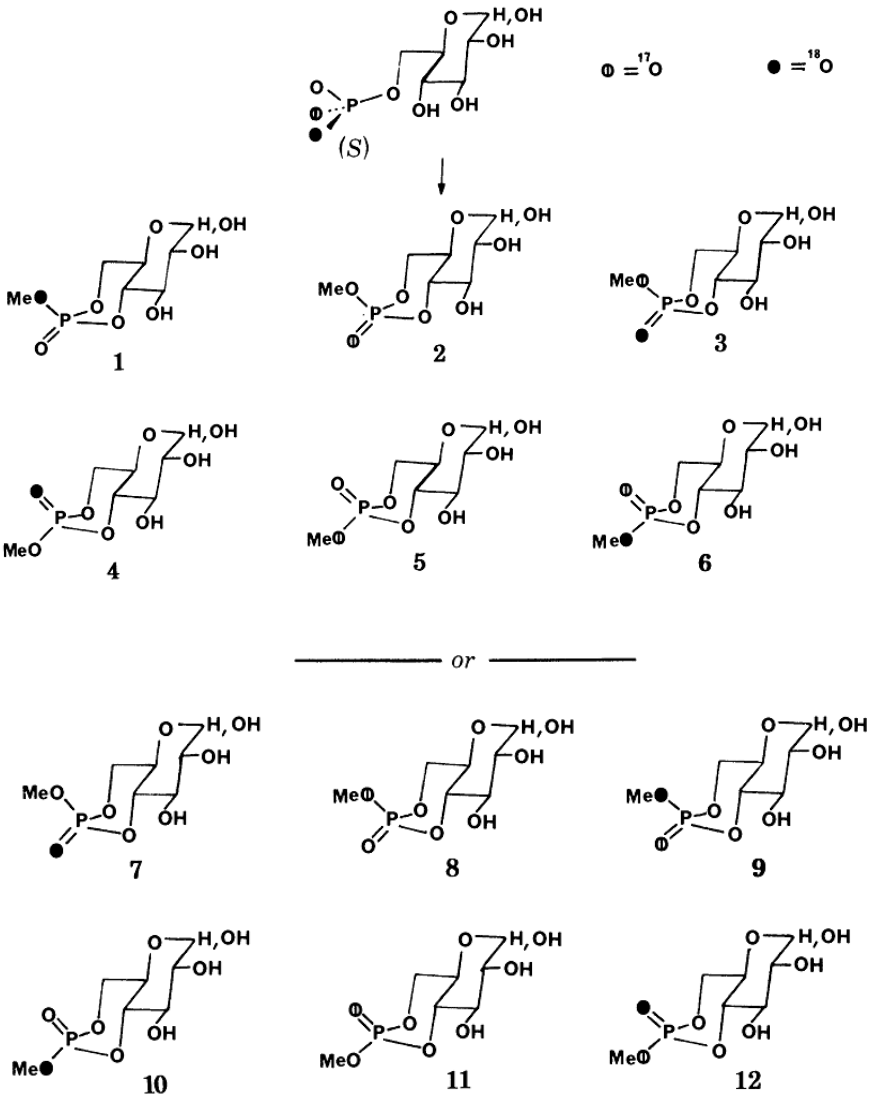


Figure 1.8 If the reaction undergoes the retention of configuration at phosphorus, after cyclisation and methylation of D-glucose 6-[(*S*)-^{16}O,^{17}O,^{18}O]-phosphate, only signals from **1** and **4** can be observed in ^{31}P NMR. If stereochemistry is inversion, then only signals from **7** and **10** can be seen.(Lowe *et al.*, 1981)

1.3.3 Results of stereochemical analysis of enzyme-catalysed phosphoryl transfers

A survey of enzyme-catalysed phosphoryl transfers studied early on for the stereochemical analysis of kinases, phosphomutases and phosphatases using one of the above two methods showed that kinases mostly go for "inversion" but phosphomutases all go for 'retention' (**Table 1.1**). The discrimination between retention and inversion was not obvious in phosphatases and these were found later to use two different

mechanisms. It is not difficult to imagine that an odd number of transfers with inversion at each step will lead to inversion as the final configuration, while an even number will lead to net retention.

Table 1.1 Stereochemistry of enzyme-catalysed phosphoryl transfer reactions (adapted from Fersht, 1999).

Kinases	
Retention	**Inversion**
Nucleoside diphosphate kinase	Acetate kinase
	Adenosine kinase
	Adenylate kinase
	Creatine kinase
	Glucokinase
	Glycerol kinase
	Hexokinase
	Phosphoenolpyruvate carboxykinase
	Phosphofructokinase
	Phosphoglycerate kinase
	Polynucleotide kinase
	Pyruvate kinase
	Ribulose-phosphate kinase
	Thymidine kinase
Phosphatases	
Acid phosphatase	ATPase (mitochondrial)
Alkaline phosphatase	ATPase (thermophilic bacterium PS3)
ATPase (sarcoplasmicreticulum)	GTPase (elongation factor G)
Glucose-6-phosphatase	Myosin ATPase
Nucleoside phosphotransferase	Purple acid phosphatase
	P21 Ras
	Pyrophosphatase
Mutases	
Phosphoenolpyruvate phosphomutase	
Phosphoglucomutase	
Phosphoglycerate mutase (co-factor independent and dependent)	

In practice, inversion indicates the direct transfer from one substrate to the other (one step) and this is what kinases nearly always use as the transfer mechanism. Later, "retention" stereochemistry was identified as involving double inversions through the formation of a phosphoenzyme intermediate as shown in **Figure 1.9d**. Phosphoserine phosphatase and phosphoglucomutase studied in this thesis fall into this second category.

Phosphoryl transfer could use any of the four mechanisms shown in **Figure 1.9**: (**a**) associative in-line phosphoryl transfer from the donor substrate to nucleophilic substrate; (**b**) dissociative in-line phosphoryl transfer in which a metaphosphate intermediate forms first then attacked by the nucleophilic substrate; (**c**) phosphorus is attacked sideways on (adjacent mechanism) followed by pseudorotation before the leaving group departs; (**d**) substrate is first attacked by a nucleophilic residue in enzyme to form a phosphoenzyme intermediate, then the phosphoryl group on enzyme is transferred to another substrate through nucleophilic attack.

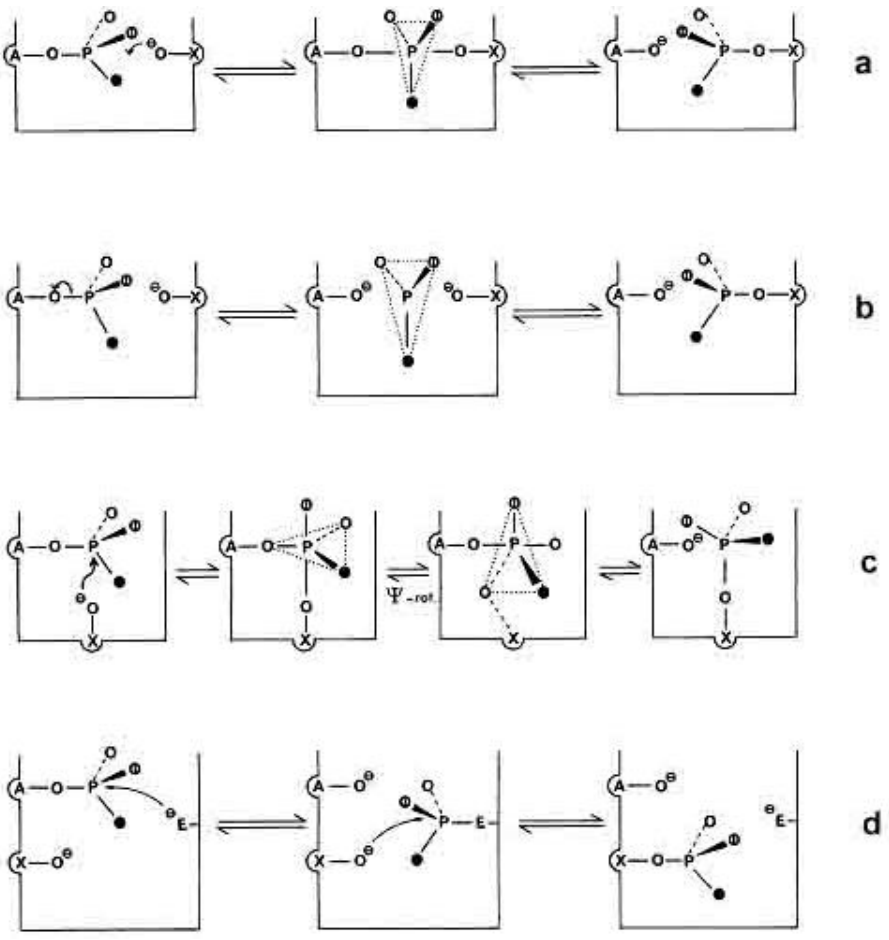


Figure 1.9 Four possible mechanisms that phosphoryl transfer enzymes could follow in the early stage of study. (**a**) "in-line" associative; (**b**) "in-line" dissociative; (**c**) "adjacent" with stable intermediate; (**d**) "in-line" double displacement.

"Adjacent" stereochemistry, rather than "in-line" stereochemistry, is demonstrated in **Figure 1.9c**. The entering nucleophile has to be apical whereas the ultimate leaving group is equatorial for the adjacent mechanism. So, for the leaving group to be expelled from an axial position (as required by the principle of microscopic reversibility), it must undergo a movement called 'pseudorotation' to transpose it into an apical position.

Such pseudorotation is not strongly supported by experimental evidence for any enzyme reaction. The only one example where X-ray structure analysis ever suggests any possibility of an adjacent mechanism is for one of the 2-phosphoglycerate mutases.(Fothergill-Gilmore *et al.*, 2006)

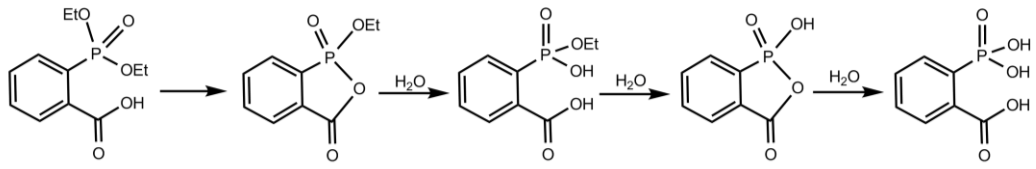
The characterisation of phosphoenzymes is less difficult today than in the past, partly because of the development of biochemical and physical methods and partly because of any increased awareness of the existence of artefacts resulting from adventitious phosphotransfer processes. Since the stereochemical configuration can be identified in phosphoryl transfer for both substrate and product, it led on to a strong focus on how nucleophilic catalysis can occur to explain the stereochemistry in each step in a two-step process.

1.4 Nucleophilic Catalysis of Phosphoryl Transfer

1.4.1 Mechanistic evidence for nucleophilic catalysis

1.4.1.1 Model studies on carboxylate catalysis

The first example of a carboxylate being a nucleophile to phosphoroxy esters was from model studies on the hydrolysis of diethyl 2-carboxyphenylphosphonate. Through intramolecular carboxylate catalysis *via* three intermediates shown as below, the hydrolysis of phosphate ester was estimated to proceed 8×10^7 times faster than the hydrolysis of its *para*-isomer. The ring-closure and ring-opening process was proposed to be significant as this route avoids any process which would violate the rules of pseudorotation.(Blackburn *et al.*, 1969)



1.4.1.2 Based on mutational and X-ray work on ground state

All the phosphoryl transfers that give retention of stereochemistry involve nucleophilic attack catalysis. It is now known to involve the negatively charged side chain from Asp and polar uncharged side chains from amino acids such as Ser, Thr, Tyr, Cys, and His. They have been identified by single point mutation in the early stages of this work and

most of them have been confirmed by X-ray crystallography on the structure of either the GS or product state form.

Enzymes that use serine as nucleophile include such examples as α-phosphoglucomutase and acid and alkaline phosphatases. For instance, *Escherichia. coli.* alkaline phosphatase uses activated serine 102 as a nucleophile to attack dianions of any phosphomonoesters to form a phosphoenzyme intermediate followed by the hydrolysis of this phosphoryl group.(Coleman, 1992) α-Phosphoglucomutase, which uses a different mechanism from the β-phosphoglucomutase (βPGM) studied in this thesis, acts on the α-C(1) anomer of D-glucose 1-phosphate (αG1P) and converts it into D-glucose 6-phosphate (G6P). The reaction proceeds with overall retention of configuration indicating two inversion steps *via* a phosphoenzyme intermediate involving reaction with a nucleophilic serine in the active site.(Shankar *et al.*, 1995)

Aspartate is widely used in ATPases including K^+,Na^+ and Ca^{2+}-ATPases to form a phosphoenzyme intermediate which is then hydrolysed. Asp 369 of K^+,Na^+-ATPases is phosphorylated by ATP when sodium is present.(Glynn, 1985; Rice *et al.*, 2001; Kaplan, 2002) Many phosphatases and phosphomutases use aspartates too, such as the βPGM and phosphoserine phosphatase (PSP) studied in this thesis.

Phosphoglycerate mutase, nucleoside diphosphokinase (NDPK), and histone phosphokinase use the imidazole of the catalytic histidine to form a phosphoryl histidine intermediate.(Zwaig *et al.*, 1966; Smith *et al.*, 1973; Xu *et al.*, 1997) In nucleoside diphosphokinase that is supported by an X-ray structure of a complex of NDPK-ADP-AlF_3 TSA while a similar complex with tetrahedral BeF_3^- mimics NDPK-ATP or phosphoNDPK-ADP complexes with imidazole at an in-line position to the BeF_3^-.(Xu *et al.*, 1997) Histone phosphokinase was proved to involve a phosphohistidine by phosphoaminoacid analysis and the acid lability of the phosphorylated histone H4.(Smith *et al.*, 1973; Chen *et al.*, 1974)

All three types of topoisomerases use catalytic tyrosine to attack the scissile phosphoryl group of the phosphodiester bond in DNA backbone to form a phosphotyrosine intermediate. For example, in human topoisomerase I (PDB: 1A36), tyrosine 723 forms a covalent O3'-phosphotyrosine intermediate (**Figure 1.10**).(Stewart *et al.*, 1998)

Cysteine features as a nucleophilic catalyst for several protein tyrosine phosphorylases.(Denu *et al.*, 1996) A thiolate anion can be formed by specific base catalysis from the nearby aspartate base and then it attacks the phosphorus atom of the substrate to form a phosphoryl-cysteine intermediate.

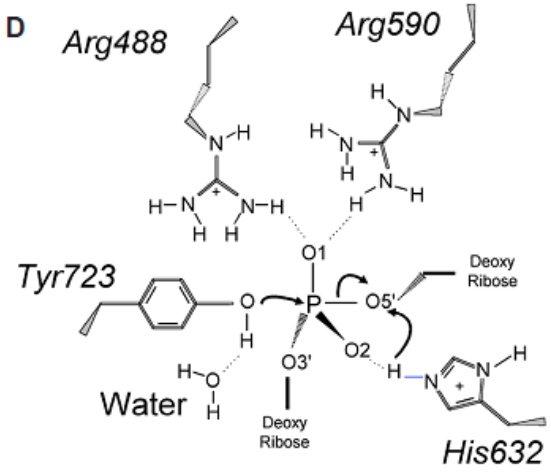


Figure 1.10 Diagram depicting a tbp TS in the cleavage step of topoisomerase I catalysis. The catalytic Tyr 723 attacks the scissile phosphate assisted by the other conserved active site residues Arg 488, Arg 590, and His 632 to give a O3'-phosphotyrosine intermediate.(Stewart *et al.*, 1998)

1.4.2 Structural Evidence for nucleophilic catalysis

The evidence for nucleophilic catalysis is most strongly provided by transition state analogues like aluminium tetrafluoride and magnesium trifluoride in X-ray structures because of the clear steric positions of the nucleophile and leaving group in the reaction and geometric arrangement of the phosphoryl group in transfer. Many crystal structures have been deposited in the PDB focused on TSA complexes of G proteins, phosphatases, phosphomutases, and kinases. For instance, the TSA structure of β-phosphoglucomutase βPGM-MgF_3^--G6P complex (PDB: 2WF5) and phosphoserine

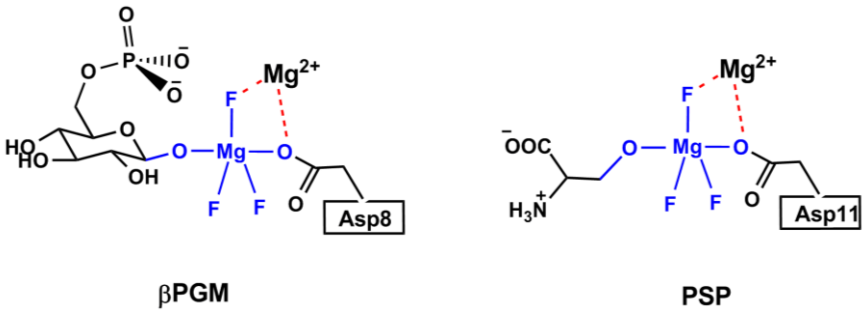


Figure 1.11 Diagrams of the structures of βPGM-MgF_3^--G6P TSA complex and PSP-MgF_3^--Ser TSA complex showing Asp 8 and Asp 11 are the nucleophilic catalysts.

phosphatase PSP-MgF_3^--Ser complex (unpublished) at high resolution both give very clear structural evidence of nucleophilic catalysis from the catalytic Asp 8 (in βPGM)

and Asp 11 (in PSP) for the phosphorylation step of the enzyme (**Figure 1.11**). This subject will be developed further in **Chapters 3** and **4** for these specific examples.

1.5 Metal Ion Catalysis of Phosphoryl Transfer

The common feature in the active site of phosphoryl transfer enzymes is the presence of positive charge, in the form either of a divalent metal centre and/or positively charged amino acid side chains.(Cleland *et al.*, 2006) Almost all the phosphoryl transferases are metal ion-dependent, such as polymerases (Steitz, 1993), kinases (Manning *et al.*, 2002), phosphatases (Hengge, 1998), ribozymes (Grasby, 1998), and nucleases (Williams, 1998). Therefore, a metal ion is virtually essential for catalysis. Zn^{2+}, Mg^{2+}, Cd^{2+}, Co^{3+}, Fe^{3+}, Mn^{2+}, Ni^{2+} have all been found as cofactors in phosphoryl transferases and occasionally Na^{+} and K^{+} also have functions in the process.

Even though the identity of some of the metal ions *in vivo* is unknown, they are still thought to play an important role for electrophilic catalysis and a source of hydroxyl ions at neutral pH.(Fersht, 1998)

1.5.1 Electrophilic catalysis

Electrophilic catalysis works mainly through the stabilisation of the negative charges developed during the reaction together with direct participation of other amino acids as general acid/base to enhance the rate.(Fersht, 1998; Tock *et al.*, 2003) In this thesis, the function of Mg^{2+} in βPGM and PSP falls into this category. But there have been other opinions that metal ions alone can be responsible for rate enhancements without catalytic amino acids participating directly.(Tock *et al.*, 2003) For example, the presence of Mg^{2+} alone increases the hydrolysis and methanolysis of ATP by a factor of 11 at 25°C, pH 7 and also narrows the specificity of the nonenzymatic phosphorylation of methanol by straining the substrate for the γ-phosphate as the site of methanol attack.(Stockbridge *et al.*, 2009)

Polarisation of the electrophile can lower the TS energy through charge reduction, but it depends on the nature of the phosphotransfer. In this part of the work, the conflict on the role of metal in the transition state is completely opposite for dissociative and associative mechanisms. In the associative phosphotransfer, the charge is more localised on the three negatively charged oxygens in the pentaoxyphosphorane-like TS (**Figure 1.12**), therefore the interaction of metal ions or positively-charged amino acid

side chains with the phosphoryl oxygen stabilises the TS. In the dissociative mechanism, there are more negative charges on the substrate than the metaphosphate-like TS (tbp PO_3^-, **Figure 1.12**). Thus the TS is only destabilised by these interactions contributed by the metal ions or positively-charged residues relative to the GS. Studies on the nonenzymatic hydrolysis of phosphomonoesters, phosphodiesters, and ATP (Admiraal *et al.*, 1995) show the reactions more or less have dissociative character *via* a metaphosphate-like pathway. Some people believe that because metal ions have no impact on changing the TS character of the nonenzymatic reaction, using metal in the enzymatic reaction has much smaller effect and does not change this dissociative character in the catalysis (Admiraal *et al.*, 1995) because many metals in the enzyme are coordinated by the carboxylate residues from aspartic or glutamic acids which neutralise the overall positive charge on metal.

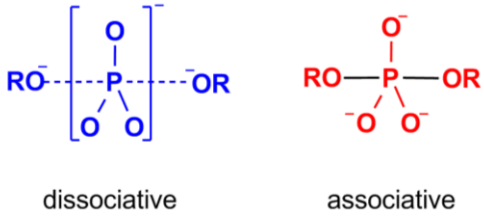


Figure 1.12 Charge distribution in dissociative and associative mechanisms.

1.5.2 Activation of water as nucleophile

Metal-bound hydroxyl ions (hydroxide) are potent nucleophiles. For example, the cobalt-bound water molecule in the following equation ionizes with a pK_a of 6.6, some 8 units below the pK_a of free H_2O.

$$(NH_3)_5Co^{3+}OH_2 \rightarrow (NH_3)_5Co^{3+}(OH)^- + H^+_{aq.}$$

Yet the cobalt-bound hydroxyl group is only 40 times less reactive than the free OH^- ion in catalysing the hydration of CO_2. Metal-bound water molecules of pK_a 7 provide the most efficient nucleophilic hydroxyl groups at neutral pH.(Fersht, 1998)

It is also common to see an activated water as a nucleophile by divalent metal ions in phosphate diester hydrolysis (**Figure 1.13**).

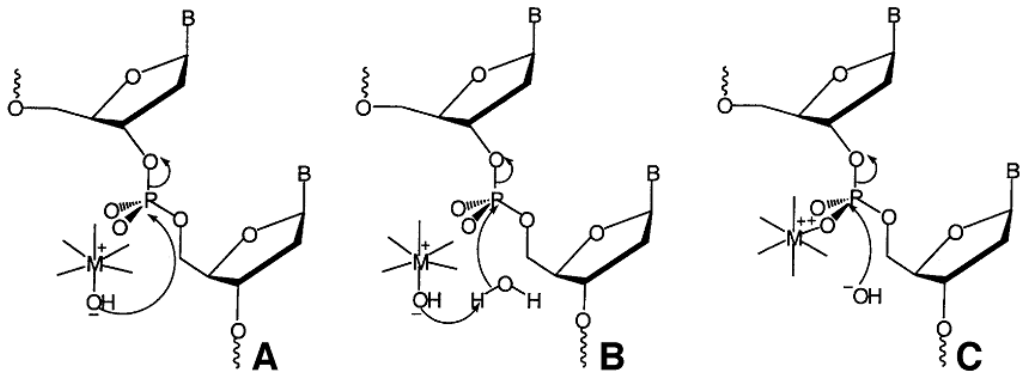


Figure 1.13 Possible roles for divalent metal ions in hydrolysis of phosphate diester. (**A**) Metal-bound hydroxyl group is the nucleophile in the reaction. (**B**) Metal-bound hydroxyl group acts as a general base to active another water as nucleophile. (**C**) A metal ion acts as a Lewis acid catalyst by binding to one of the non-bridging oxygens of the backbone phosphate.(Tock *et al.*, 2003)

1.5.3 Binding of substrate and stabilisation of tertiary structure of enzymes

The role of metal such as Ca^{2+}/Mg^{2+} in striated muscle proteins is always associated with stabilising the tertiary structure of enzymes although it is not clear whether Ca^{2+} or Mg^{2+} is the structural metal ion *in vivo*. For example, Ca^{2+} has the function of maintaining the structure of G-actin upon ATP binding which is then hydrolysed when G-actin filament is polymerised. ATP is bridging between the two domains and its β and γ-phosphates interact with Ca^{2+} (**Figure 1.14**).(Kabsch, *et al*, 1990)

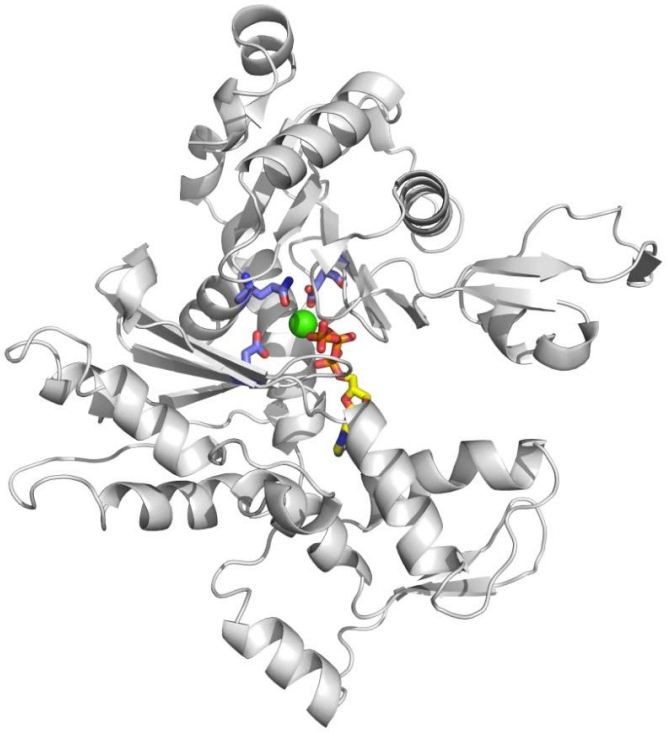


Figure 1.14 Structure of G-actin with Ca^{2+} (green sphere) and ATP (yellow sticks) bound. Residues Asp 11, Gln 137 and Asp 154 are highlighted in purple sticks and coordinate to the catalytic metal. (Adapted from PDB: 1ATN)

In most kinases and nucleotide hydrolases, nucleotide binding is always associated with metal binding. The binding of the metal ions not only enhances the binding of the nucleotide, but also aids orientation of the functional group (phosphoryl group in this case) in the substrate in the ground state. UMP/CMP kinase, cAPK, and the small GTPase RhoA studied in this thesis fall into this category and the detailed function of the metal in each of these enzymes will be described individually.

1.5.4 Multiplicity of metals in enzymes

The best studied phosphotransferases are mono- or di-nuclear metalloenzymes. The nature and number of metal ions *in vitro* are identified by the measurement of the metal ion dependence of the turnover number (k_{cat}) (Tock *et al.*, 2003), isothermal titration calorimetry (ITC) (Zheng *et al*., 2002), and X-ray crystallography. In some cases, the catalytic role for metal ions in the reaction may also be suggested from the limited available structures that no other candidate amino acid residues within the active site could account for the pH rate profile log(k_{cat}/K_M) of the enzyme, as in T5 FEN.(Tock *et al.*, 2003)

Inositol monophosphatase binds two metal ions of which the presence of Mg^{2+} is essential for activity, although manganese, zinc, and cobalt are also found to promote catalysis. X-ray crystal structures and metal titration have shown the two adjacent metal binding sites within each active site undergo a cooperative binding. A wealth of data has suggested a direct displacement mechanism, because Wat 2 coordinated to M_2 is nucleophilic. M_1 possibly binds after the substrate is bound due to the close coordination between them.(Sträter *et al.*, 1996)

Protein kinase A (cAPK) studied in this thesis is a good example of the use of two magnesium ions to carry out the function. cAPK uses a primary, high-affinity metal site to chelate the β- and γ-phosphates and a secondary site to chelate the α- and γ-phosphates of ATP in the structures. The previous study also found the nature and concentration of the divalent metal directly affects k_{cat}, or k_{cat}/K_M by attenuations in either the accessibility of the active site for substrate, the rates of phosphoryl transfer, or the release rate of the products.(Adams *et al.*, 1993)

Some phosphotransfer enzymes use more than two metals. Zinc-dependent prokaryotic phospholipase C from *Bacillus cereus* binds three zinc ions which are coordinated by

nine side-chain residues, including five histidines, two Asp, one Glu and one Trp. Zinc ions were also proved to be protective of the susceptible histidine residues.(Titball *et al.*, 1990) Soluble inorganic pyrophosphatase from yeast (PPases, EC 3.6.1.1) uses four catalytic Mg^{2+}. M1 and M2 are identified as the activating metal ions while M3 and M4 functionalise binding with the substrate (**Figure 1.15**). After the reaction, P2 is the phosphate which leaves first. It is found that the number of metal ions involved in the crystal structures depends on the pH. At low pH (6.0), PPase binds four catalytic metals, but when the pH goes higher, the number of metals can drop down to three, as supported by both crystallography and ^{19}F NMR (Y. Jin, unpublished). Therefore it is proposed that the function of M4 is to assist in generating the general acid which would be important at low pH.(Oksanen *et al.*, 2007)

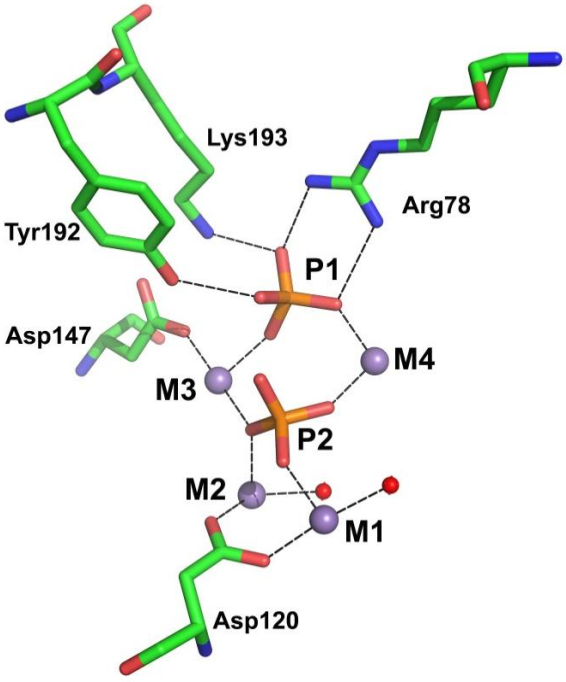


Figure 1.15 Structure of inorganic pyrophosphatases from yeast (PDB: 1E9G). The metal (in purple spheres) positions in the active site are termed M1-M4 in the proposed order of metal binding, and P2 (in orange and red sticks) is the leaving phosphate.

In the most recent publication in which the MgF_3^- was identified, an inositol pyrophosphate kinase uses four/five catalytic/structural magnesiums to meet the catalytic challenge.(Wang *et al.*, 2011)

1.6 General Acid-Base Catalysis (GABC)

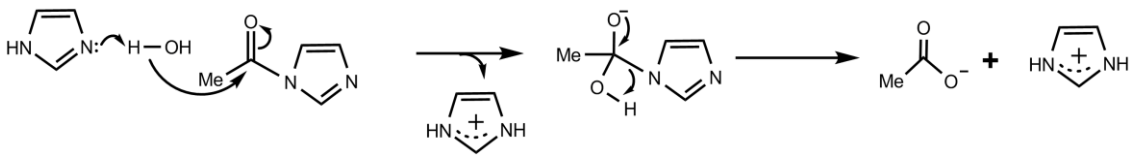
In both nonenzymatic and enzymatic reactions, there are as many or even more steps involving proton transfers than there are covalent steps. Acid-base catalysis is key to such proton transfers and is divided into two distinct types.

Specific Acid-Base Catalysis is where the proton is fully transferred before covalent bonds make or break. In water, the specific acid is always H_3O^+ and the specific base is OH^-. Therefore, the reaction rate is only determined by pH and independent of the concentration of other components. The hydrolysis of a peptide bond is a good example of this case. It is catalysed by hydrogen ion (a proton) or hydroxide in two different ways simply by changing the pH. Because it is not relevant to the phosphoryl transfer mechanisms to be discussed in this thesis, details will not be described further. The nucleophilic activity of cysteine in some tyrosine phosphorylases is a further example (**Section 1.4.1.1**).

General Acid-Base Catalysis (GABC) is where the proton is transferred at the same time as covalent bonds are being made or broken. Brønsted acids and bases are catalysts in addition to protons and hydroxide ions. Enzymes mostly use this means to carry out acid-base catalysis simply because the concentration of protons and hydroxide ions at neutral pH is too low to support the reaction rates needed for life.

1.6.1 General acid-base catalysis in physical organic chemistry

A classic example of GABC is the catalysis of the hydrolysis of acetylimidazole by imidazole buffer at constant pH. It was found that the hydrolysis rate at constant pH increases when the concentration of imidazole buffer increases. Since the hydroxide ion concentration does not change at the same pH, this means the buffer itself, as a basic component, is catalysing the reaction. Because imidazole cannot act as a nucleophilic catalyst for the hydrolysis of acetylimidazole, the following mechanism was proposed:



In a nonenzymatic reaction, the relative reactivities of acids and bases are correlated by the Brønsted catalysis law, which is a linear free energy correlation. For GBC, the law is expressed as

$$\log k_{gb} = \beta p K_a + C,$$

where k_{gb} is the rate constant for catalysis by a series of bases, pK_a is the negative logarithm of the dissociation constant for the conjugate acids of the bases, C is a constant for the reaction, and β is the coefficient of pK_a and is a measure of the sensitivity of the particular reaction to GBC. Therefore, for a certain GBC reaction, the stronger is the base, the more reactive it is and the faster proton transfer takes place as long as proton abstraction is a key part of the transition state. For GAC, the law is expressed as

$$\log k_{ga} = -\alpha p K_a + C$$

and the values of α and β are always between 0 and 1 for GABC. When the transfer of a proton is complete before the TS, α and β values are 1. At the other extreme, no proton transfer (or diffusion-controlled proton transfer) gives an α or β value of 0.(Jencks, 1987; Fersht, 1998; Frey *et al.,* 2007) Of course, the efficiency of GABC depends on the ionisation state of the catalyst. An acid has to be in acidic form to be an acid catalyst, and a base has to be in its basic form. For example, although an acid of pK_a 5 is a better GA catalyst than one of pK_a 7, at pH 7 the acid of pK_a 5 is only 1% deionised to carry out the function. By contrast, an acid of pK_a 7 is 50% in its acid form. Thus, it is not hard to conclude that the most effective GAB catalysts at neutral pH are generally those whose pK_a are about 7.

1.6.2 GABC in enzyme catalysis

Physical organic studies on reactions have often become the models for more complex enzymatic reaction discovered later. That is the reason why histidine (pK_a 6 ~7) has been so well-studied in the enzyme catalysis.(Fersht, 1998) A good example of this sort is the serine protease, chymotrypsin. The main players in the catalytic mechanism of chymotrypsin are called the *catalytic triad.* It uses histidine as a GAB catalyst along with serine as the nucleophile, while the aspartate anion participates in a charge relay. The mechanism is shown in **Figure 1.16**:

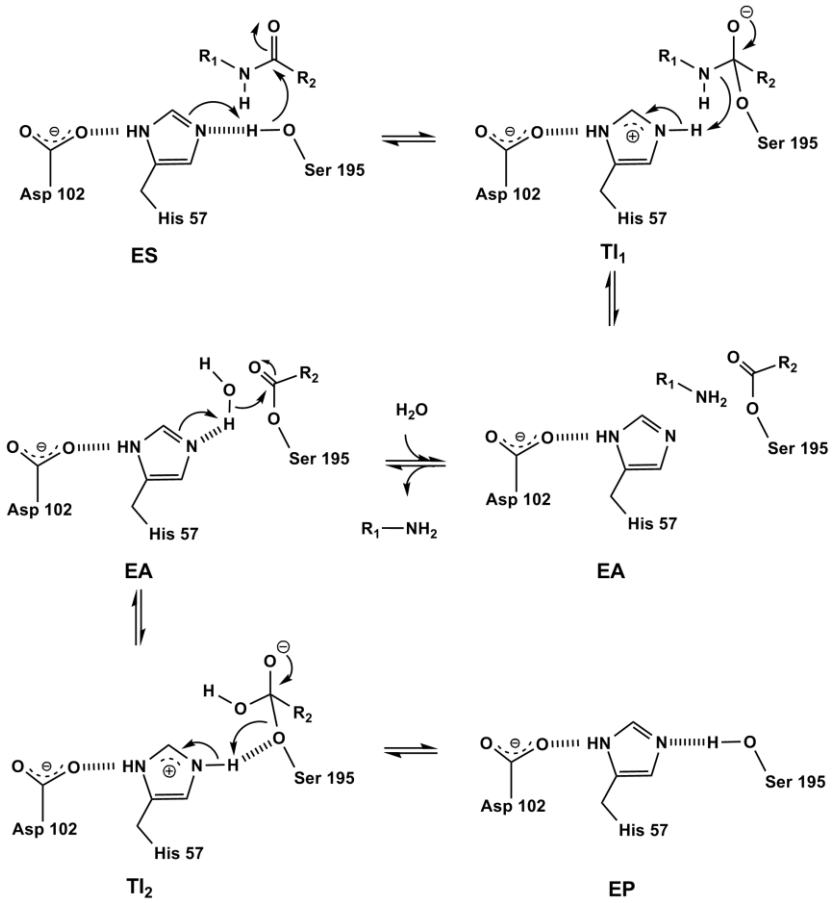


Figure 1.16 Reaction mechanism of chymotrypsin.

The first step in the hydrolysis is driven by the physical forces of attraction to form a noncovalent enzyme-substrate complex (ES). This is then followed by the nucleophilic attack of the Ser 195 hydroxyl group to the substrate peptide carbonyl catalysed by His 57 as a GB to give the first tetrahedral intermediate (TI_1). This intermediate then collapses to give the acyl enzyme (EA) using the now protonated His 57 as a GA. The acyl enzyme is hydrolysed by the same mechanism using His 57 as a GB catalyst *via* another tetrahedral intermediate (TI_2) followed by the release of the product (EP).(Fastrez *et al.*, 1973)

In addition to imidazole, there are a number of groups on the side chains of amino acids such as Asp, Glu, Cys, Lys, and Tyr hydroxyl groups which can all act as GAB catalysts in the active sites of enzymes. A particularly successful identification of GABC is for the role of Lys 208 as a GB in aspartate aminotransferase. It was demonstrated that its GB function can be partly restored by adding back primary amines to the solution of the K208A mutant of this enzyme.(Toney *et al.*, 1989; Toney *et al.*, 1992; Frey *et al.*, 2007)

1.6.3 GABC in enzyme-catalysed phosphotransfer

GABC is often observed in phosphoryl group transfer. Jencks has argued that this is because it involves attack on an electrophilic phosphorus centre that is unsaturated (Jencks, 1987) though this remains a rather intuitive opinion. Many phosphotransfer enzymes use both GABC and metal ions to carry out the function. Ribonuclease A (RNase A) is the classic example but uses only GABC and approximation effects without any participation of metal ion.

The accepted chemical mechanism for the hydrolysis of RNA by RNase was deduced by an inspired piece of chemical intuition before its crystal structure was solved.(Findlay *et al.*, 1961) It was first found to have a bell-shaped pH-activity profile, with optimal rates around neutrality. The pH dependence of k_{cat}/K_M shows the ionisation of a base of pK_a 5.22 and an acid of pK_a 6.78 in the apo-enzyme to be rate limiting, while the pH dependence of k_{cat} shows they are perturbed to pK_a .6.3 and 8.1

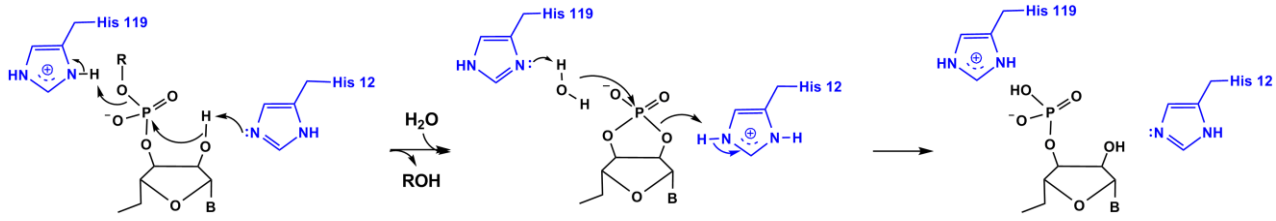


Figure 1.17 Reaction mechanism of RNase A.

in the substrate-bound complex. Therefore it was proposed that two histidines catalyse the reaction by concerted GABC (**Figure 1.17**). In the cyclisation step, His 119 acts as a GA to protonate the leaving group and His 12 acts as a GB. These catalytic roles are reversed in the hydrolysis step where His 119 acts as a GB and His 12 acts as a GA to protonate the leaving group. This is sensible from the principle of microscopic reversibility that requires that a group acting as a GA in one direction will act as a GB in the reverse direction.

The transition states for GABC include proton transfer in the rate-limiting steps. Because bonds to the proton in flight are being broken in the transition states, the GABC can be reflected by solvent deuterium isotope effects in D_2O, which indicates the extent of proton transfer. The largest deuterium kinetic isotope effect (KIE) can be expected when the proton in flight is equally bonded to the GAB and reactant in the linear and symmetric transition state.(Rigas *et al.*, 2001; Frey *et al.*, 2007)

In the reactions of PSP and βPGM, an aspartic acid is suggested to be GA and GB in the consecutive enzyme phosphorylation and dephosphorylation steps. The details about how the aspartic acid carries out its function in the GABC in these two enzymes will be covered later in the corresponding chapter.

For those kinases equipped with GAB residues in the active site, GABC is an important part of the enthalpic component of the thermodynamic basis during catalysis. It can decrease the enthalpy of activation for substrate phosphorylation. For example, the active site of homoserine kinase (HSK) uses the positively charged end of a helix dipole for stabilising an associative transition state and hexokinase is equipped with an aspartate residue that may serve as a catalytic base. In the structure of human GalNAc kinase (PDB: 2A2D), Asp 190 is a potential GAB catalyst for the C(1)-OH on galactose ring (**Figure 1.18**), although given the proximity of Arg 43 which forms two hydrogen bonds with this aspartate, some researchers think this Asp 190 is impossible to be a GAB catalyst.(Thoden *et al.*, 2005; Stockbridge *et al.*, 2009) In cAPK as studied in this thesis, because Asp 166 is calculated consistently to have a very low pK_a value when cAPK adapts its closed conformation (Bjarnadottir *et al.*, 2010), the GABC is suggested to be from Asp 166 whose negative charged oxygen removes the proton on the accepting hydroxyl of serine or threonine.

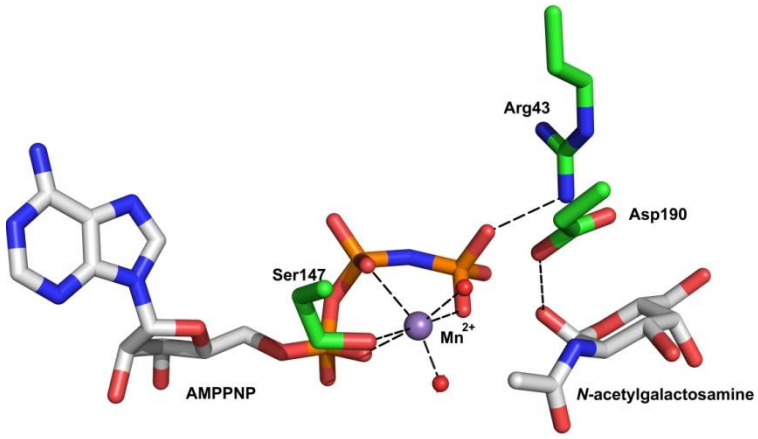


Figure 1.18 Structure of human GalNAc kinase complex with AMPPNP and *N*-acetylgalactosamine (PDB: 2A2D). Asp 190 is a potential GAB catalyst for the C(1)-OH on the galactose ring, although given the proximity of Arg 43 that can form two hydrogen bonds with this aspartate, some researchers think it is not possible for this Asp 190 to be a GAB catalyst.(Thoden *et al.*, 2005; Stockbridge *et al.*, 2009)

1.7 Transition States for Phosphoryl Transfer

1.7.1 Bond length and bond order

Chemists have known for a long time that the hydrolysis of phosphate monoester dianions proceeds by a dissociative mechanism. For the stereochemistry of phosphoryl transfer in an enzymatic-catalysed reaction to be inverted in the case of direct transfer, the reaction must be concerted. On the basis of both theoretical calculation and experimental observation that enzyme catalysed phosphoryl transfer reaction is more S_N2_P-like.(Webster, 2004) Therefore, great interest has developed on whether the mechanism of phosphoryl transfer is dissociative (loose) or associative (tight).

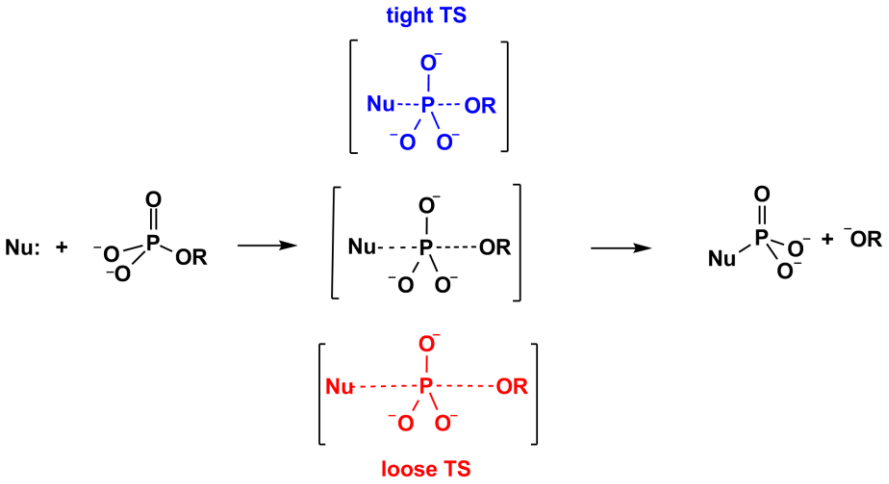


Figure 1.19 Possibilities for concerted transition state.

Whether a TS is tight or loose depends on the nature of the bonding, qualitatively defined by the distance between the leaving oxygen and the nucleophilic oxygen. This has been put in the overall range 3.6 ~ 6.5 Å.(Mildvan, 1997; Cleland *et al.*, 2006; Bowler *et al.*, 2010)

Bond order, B.O., is also used as a measure for association or dissociation. In molecular orbital theory

$$\text{B.O.} = \frac{\text{number of bonding electrons} - \text{number of antibonding electrons}}{2}$$

or it can be described experimentally as

$$\text{B.O.} = \exp\left[\frac{d_1 - d_{ij}}{b}\right]$$

where d_1 is the single bond length, d_{ij} is the bond length experimentally measured, and *b* is a constant, depending on the atoms.

The general opinion on the mechanism of uncatalysed monoester dianion cleavage has been that the geometry of the phosphoryl group involved in monoester transfer changes from tetrahedral to tbp then back to tetrahedral involving a 'dissociative' TS. If there is no compression of the reaction coordinate, the distance between the attacking and leaving group oxygens stays constant at the Van der Waals contact distance, 5.07 Å. Such a situation gives a bond order between phosphorus and the attacking or leaving oxygen (O–P–O) around 0.08 in a symmetric TS.(Cleland *et al.*, 1995) While a "true dissociative" reaction will have a monomeric metaphosphate intermediate, there is no evidence for its existence. Rather, reactions of phosphate monoesters are deemed to be concerted and not stepwise. That has caused some confusion in the adoption of the term "dissociative" for the mechanisms of such reactions and has led Cleland to make the clarification: "The word dissociative implies that the molecule expands during the reaction, but this is not the case".(Cleland et al., 1995)

1.7.2 Kinetic isotope effects in enzymatic phosphoryl transfer

An isotope effect is the ratio of the rate or equilibrium constant for the light isotope divided by that for the heavy atom when an isotopic substitution is made in one of the reactants or solvent. A kinetic isotope effect (KIE) is the isotope effect in terms of the reaction rate in a chemical process and reflects difference in bonding to the labelled light and heavy atoms in the GS compared to the TS for the rate-limiting step.

$$KIE = \frac{k_{\text{Light atom}}}{k_{\text{Heavy atom}}}, \quad \text{where } k \text{ is a rate constant.}$$

Although the techniques used in this thesis are primarily crystallography and NMR, KIEs are reviewed here as a technique because of their value in determining the presence of GABC and the dissociative or associative nature of phosphoryl transfer.

The best studied cases using KIE in phosphate chemistry are the catalysis by protein tyrosine phosphatases (PTPases) on the artificial substrate *p*-nitrophenyl phosphate (*p*NPP).(Hengge, 2002) The approach is by measuring primary ^{18}O isotope effect $^{18}k_{\text{bridging}}$ in the leaving group, secondary ^{18}O isotope effect $^{18}k_{\text{nonbridging}}$ in the nonbridging oxygens, and secondary ^{15}N isotope effect ^{15}k in the nitro group.

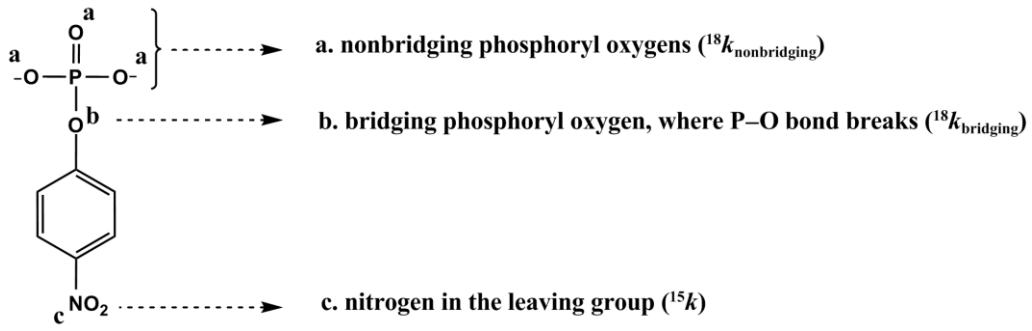


Figure 1.20 Diagram of *p*NPP showing positions where KIEs have been determined.

As shown in **Figure 1.20,** an observation of $^{18}k_{\text{bridging}}$ shows the P–O bond cleavage is partly rate-determining and a measure of the degree of cleavage of the P–O bond in the TS. Its value is usually in the range of 0 ~ 3%.(Hengge, 2002) In **Table 1.2**, for both noncatalysed and catalysed reactions, the results are indicative of TS that is loose in nature and with minimal nucleophilic participation.

Table 1.2 Isotope effects on % for noncatalysed and catalysed reactions of *p*-nitrophenyl phosphate.

Reaction	**^{15}k**	**Primary ^{18}k**	**Secondary ^{18}k**
Monoanion (95˚C)	0.04 ± 0.02	0.87 ± 0.02	1.84 ± 0.02
Dianion (95˚C)	0.28 ± 0.02	1.89 ± 0.02	–0.06 ± 0.02
Dianion in *t*-BuOH (30˚C)	0.39 ± 0.02	2.02 ± 0.02	–0.03 ± 0.02
PTPase	–0.01 ~ 0.01	1.18~ 1.52	–0.02~ 0.03
Alkaline Pase			
pH 8 (22˚C)	0.03 ± 0.02	0.03 ± 0.02	–0.18 ± 0.02
pH 6 (22˚C)	0.03 ± 0.02	0.05 ± 0.02	–0.21 ± 0.02

^{15}k and ^{18}k are ^{15}N or ^{18}O isotope effects on the rate constants. For the enzyme reactions, these are *V*/*K* data.(Hengge *et al.*, 1994; Hengge, 2002)

The secondary isotope effect $^{18}k_{\text{nonbridging}}$ reveals whether the TS is dissociative or associative. This is because isotope effects are determined by bond order especially an atom bonded to a site undergoing a hybridisation change. Therefore, $^{18}k_{\text{nonbridging}}$ should be inverse (< 1) for dissociative transition states (sp^3 to sp^2) because the bond order will approach 5/3 in metaphosphate-like TS and normal (≥1) for associative mechanisms due to the bond order would be decrease from 4/3 to almost 1 in TS. It also gives information as to whether the leaving group departs as the anion (KIE < 1), or if protonation of the leaving group has neutralised all or part of the negative charge (KIE ≥ 1) resulting from P–O bond fission. Its value is usually in the range 0 ~

0.3%.(Hengge *et al.*, 1994; Hengge, 2002) The big value for monoanion of pNPP is caused by the deprotonation of phosphate which is no longer a secondary but rather a primary isotope effect.

Secondary isotope effect ^{15}k indicates the degree to which the electrons are delocalised into the $-NO_2$ group in TS.(Hengge *et al.*, 1994) Very often, the $^{18}k_{bridging}$ and ^{15}k are found to be correlated mostly, but occasionally the delocalisation of negative charge arising from P–O bond splitting measured by $^{18}k_{bridging}$ may lag behind charge development in the TS.(Bernasconi, 1992) If the P–O bond is fully broken in the TS and the resulting charge is delocalised, maximum values of $^{18}k_{bridging}$ and ^{15}k should be expected. Conversely, a TS with a smaller degree of P–O bond cleavage and less negative charge on the leaving group will exhibit smaller values for both effects. In a catalysed reaction, GABC leads to the protonation of the leaving group in the TS which also reduces the magnitudes of $^{18}k_{bridging}$ and ^{15}k (**Figure 1.21** and **Table 1.2**).(Hengge *et al.*, 1994; Hengge, 2002)

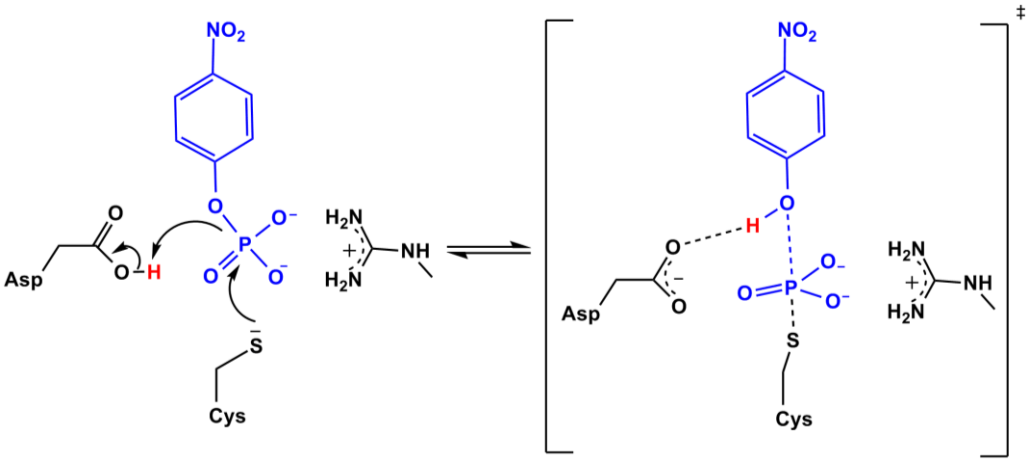


Figure 1.21 Transition state for the PTPase-catalysed reaction inferred from the KIE data.

The KIE data for alkaline phosphatase in **Table 1.2** show very small value on $^{18}k_{bridging}$ indicating the P–O bond cleavage is not rate-determining and the slight inverse ^{15}k maybe from the isotope effect for strong hydrogen bonding to arginine and/or binding the zinc.

1.8 Transition State Analogues for Phosphoryl Transfer

According to transition state theory, although the TS is bound to the enzyme more tightly than the GS by the factor k_{cat}/k_{uncat},(Wolfenden, 1969) the TS is almost

impossible to observe because the high-energy states are so short-lived that no physical or spectroscopic method can observe them directly. Some researchers believe a description of an enzymatic mechanism is coherent based on a nonenzymatic transition state as a model.(Admiraal *et al.*, 1995; Admiraal *et al.*, 2000). The history of some other scientists trying intensively to "snapshoot" the structural details of phosphoryl transfer in-flight and the discovery of the metal fluoride TSAs is a interesting story. The existence of metal fluoride TSAs can be tracked all the way back to early 1950s, Hans Krebs indicated that fluoride (0.05 M) can inhibit the reaction of monophosphotransfer between AMP and ITP catalysed by an enzyme which we now call nucleoside-triphosphate-adenylate kinase [EC 2.7.4.10].(Krebs *et al.*, 1955) In 1987 it had been reported that in the presence of F^-, addition of Mg^{2+} can cause a graded increase of fluorescence emission that was associated with a conformational change to represent the activated state of a G protein, and that meant a transition state analogue was formed. Addition of a low concentration of Al^{3+} could decrease the F^- level required to make the same change in fluorescence signal.(Higashijima *et al.*, 1987)

In the 1990's, there was a lot of research focused on the study of aluminium fluoride as TSAs of phosphoryl transfer enzymes. A pH-dependent study of crystallographic structures of 12 phosphoryl transfer enzymes was made which suggested that pH influences the coordination of AlF_x phosphoryl transfer TSAs.(Schlichting *et al.*, 1999) The TSA species shift from octahedral AlF_4^- into tbp 'AlF_3^0' as the pH rises. In the PDB, there are 34 structures reported to contain 'AlF_3^0' as TSA analogues (up to 2011) (**Table 1.3**).

Another example of a tbp transition state analogue interpreted as AlF_3^0 was for phosphoserine phosphatase (PSP) in PSP-AlF_x-H_2O complex.(Wang *et al.*, 2002) The structure has dual occupancy by 60% octahedral AlF_4^- and 40% tbp AlF_3. Now it appears more realistic to interpret their data as 60% octahedral AlF_4^- and 40% tbp MgF_3^-.(Baxter *et al.*, 2008)

MgF_3^- was first formally identified in 2002 as a TSA in the case of a G protein by the use of proton-induced X-ray emission spectroscopy (PIXE) on a crystal whose X-ray structure had been determined by Graham and coworkers in 2002.(Graham *et al.*, 2002) The confidence in the identity of the central metal is based on three facts. Firstly,

deferoxamine (DFO) was used to remove all aluminium (Al^{3+}) in the crystal trial. Secondly, because PIXE can give a minimum detectable limits in the range of 1 ~ 100

Table 1.3 Deposited structures containing bound 'AlF_3'.

PDB	Protein complex	Res Å	pH
1e2e	Human thymidylate kinase with thymidine monophosphate, ADP, Mg^{2+} and AlF_3	2.0	
1he1	GAP domain of the pseudomonas aeruginosa Exos toxin and human Rac	2.0	
1grn	cdc42/cdc42GAP with Mg^{2+}, GDP and AlF_3 (0.2 M Mg^{2+}, 0.1 mM Al^{3+})	2.1	
2b92	N-terminal large GTPase domain of human guanylate binding protein 1 (hGBP1) with Mg^{2+}, GDP, and AlF_3	3.2	
1n6k	Human Rab5a A30P mutant with Mg^{2+}, GDP and AlF_3	1.55	
2g77	Gyp1 TBC domain with Rab33 GTPase bound to Mg^{2+}, GDP and AlF_3	2.26	
1k3d	Phosphoenolpyruvate carboxykinase with ADP and AlF_3	2.0	
1pcq	Crystal structure of groEL-groES with ADP and AlF_3	2.81	
3cf1	Structure of P97/vcp with ADP/ADP.AlFx	4.4	
5ukd	UMP/CMP kinase complex with Mg^{2+}, ADP, UMP and AlF_3	1.9	8.5
1a6e	Thermosome-Mg-ADP-AlF_3	3.2	
1tuy	Acetate Kinase with ADP, AlF_3 and acetate	3.0	
2f48	Pyrophosphate-dependent phosphofructo-1-kinase with fructose 1,6-bisphosphate and AlF_3	2.11	
3ex7	The exon junction complex (EJC) with ADP and AlF_3	2.30	
1kdn	Nucleoside diphosphate kinase with ADP and AlF_3	2.0	
1vfx	Kif1A motor domain with Mg^{2+}, ADP and AlF_3	2.55	8.5
2g08	Pyrimidine 5'-nucleotidase type 1, product-transition complex analog with AlF_3	2.35	
1kh9	*E. coli.* alkaline phosphatase mutant (D153G,D330N) with phosphate	2.5	8.0
3ukd	UMP/CMP kinase from slime mold with ADP, CMP, and AlF_3	1.9	8.5
1l3r	The catalytic subunit of cAMP-dependent protein kinase with ADP, AlF_3 and SP20 substrate peptide	2.0	8.0
1k3c	Phosphoenolpyruvate carboxykinase with ADP, AlF_3 and pyruvate	2.0	
1rlt	YbiV (HAD phosphatase) from *E. coli.* K12 with AlF_3 and H_2O (5 mM Mg^{2+}, 2 mM Al^{3+})	2.2	5.3
1wq1	Ras-RasGAP with Mg^{2+}, GDP and AlF_3	2.5	8.0
1k5g	Ran-GDP-AlFx-RanBP1-RanGAP	3.1	
2ngr	Cdc42/cdc42GAP$_{R305A}$ mutant with Mg^{2+}, GDP and AlF_3 (0.2 M Mg^{2+}, 0.1 mM Al^{3+})	1.9	6.0
1ihu	*E.coli.* arsenite-translocating (ArsA) ATPase with Mg^{2+}, ADP and AlF_3 (AlF_4^- at pH 6.2)	2.15	8.0
1kh5	*E.coli.* alkaline phosphatase D330N mutant with two Zn^{2+}, H_2O and AlF_3	2.0	8.0
1l7n	Phosphoserine phosphatase with Mg^{2+}, H_2O and AlF_3	1.8	
1svt	GroEL-ADP-AlFx-GroES	2.81	
1khj	*E.coli.* alkaline phosphatase mutant (D153H,D330N) with 2 Zn^{2+}, H_2O and AlF_3	2.3	8.0
2hf7	AphA class B Acid Phosphatase/Phosphotransferase with Mg^{2+}, AlF_3 and H_2O	1.6	
1e1r	Bovine mitochondrial F1-ATPase with Mg^{2+}, ADP and AlF_3	2.5	
2d33	γ-Glutamylcysteine synthetase with glutamate, cysteine, Mg^{2+}, and AlF_3	2.6	
1g4u	SPTP (has a Salmonella tyrosine phosphatase domain and GAP domain) bound to Rac1 with Mg^{2+}, GDP and AlF_3 (0.1 mM Al^{3+}, 2 mM Mg^{2+})	2.3	5.7

parts per million (ppm) by weight for all elements with Z >11 in the Periodic Table, the measurement can be carried out to compare ratios between different atoms. Obviously,

in this case, the complex should contain two phosphorus atoms (in GDP) and two magnesium ions (one is the catalytic magnesium associated with GDP and one is the magnesium in MgF_3^-). The experimental result gave a P/Mg ratio of 0.86 (±0.1), effectively unity. Thirdly, at high resolution, the difference between the bond lengths of central metals to the axial O (O–M–O) and to equatorial Fs (M–F) can be easily measured in MgF_3^- and AlF_4^-.(Graham *et al.*, 2002) This result was soon rigorously proved by the use of ^{19}F NMR work to show that the tbp moieties in PSP, βPGM, human PGK are indeed trifluoromagnesate.(Baxter *et al.*, 2008; Cliff *et al.*, 2010)

The switch between octahedral AlF_4^- into tbp 'AlF_3^0' as the pH rises is now clear. As Al^{3+} and Mg^{2+} are both present in these solutions, those crystallised at or below pH 7.0 were square planar AlF_4^- which cannot be replaced by MgF_3^- because of the large difference in the intrinsic stability of these two metal fluoride species and the binding constant difference. But Al^{3+} precipitates out as $Al(OH)_3$ as the pH rises in some crystallisation experiments so that MgF_3^- takes over and occupies the same position in the TSA complex. Also, MgF_3^- is a good explanation of why the charge distribution of AlF_3^0 complexes is unable to match the charge in the actual transition state PO_3^- as it is not a real transition state analogue. So we have reason to suspect that many, perhaps most, of the tbps previously assigned as AlF_3^0 in the PDB are in fact MgF_3^-.

There is also some additional confusion on the assignment of trigonal planar structures. The direct observation of a pentacoordinate phosphorane has been claimed as an intermediate building up in a reaction that can be monitored by time-resolved ^{31}P cryo-MAS NMR.(Cherepanov *et al.*, 2008) This result is not convincing as there is no direct experimental data to show that the change of ^{31}P chemical shifts are not caused by other factors, for instance, pH. Another claim on having captured a pentacovalent phosphorane intermediate in the reaction of βPGM phosphoryl transfer in an X-ray crystal structures sought to establish that the transfer reaction proceeded by an associative pathway.(Lahiri *et al.*, 2003; Allen *et al.*, 2004) However, this has also been challenged and a conflict of opinion has arisen focusing on whether the XY_3 complex seen in a βPGM complex was PO_3^- or really MgF_3^-.(Blackburn *et al.*, 2003; Baxter *et al.*, 2009; Baxter *et al.*, 2010) Blackburn pointed out that if the trigonal planar structure in βPGM was truly PO_3^- then the apical P–O bond lengths for this pentacovalent phosphorus intermediate (2.0 and 2.1 Å) were longer than those in small-molecule

phosphoranes (~1.67 Å), while they are similar to the apical bond lengths of the MgF_3^- complex in the G protein (2.0 and 2.3 Å).(Allen, 2003; Blackburn *et al.*, 2003; Lahiri *et al.*, 2003; Allen *et al.*, 2004)

The P/Mg ratio appeared to be an independent way to identify the true situation. If the tbp is PO_3^-, the P/Mg ratio of the PGM tbp complex would be 2:1; if it is MgF_3^-, as seen in the small G protein, RhoA, the ratio would be 1:2.(Graham *et al.*, 2002) The Bradford protein assay and Malachite Green phosphate assays were used by Allen to give an enzyme:phosphate ratio of 1:2. However, this result is not reliable due to the fact that the Bradford protein assay always gives a lower protein concentration than other protein assays and so leads to an apparent lowering in the protein/phosphate ratio.(Berges, 1993; Tremblay *et al.*, 2005) Also, the proposed P–O bond length in Allen's structure was erroneously preset in the data analysis to be 1.7 Å, as was later discovered by Webster.(Webster, 2004) Based on these obervations, it should be appeciated that the whole controversy on the publication claiming the observation of the pentaoxyphosphorane intermediate of phosphotransfer in βPGM (Lahiri *et al.*, 2003) was a result of the fact that their crystal data was analysed in isolation and was not backed up by any other means. A full study of very short-lived TSs should involve all these synergistic biological structural and kinetic tools to ensure the interpretation of the exciting data is truly open to challenge. By this means, the confusion on the assignment of trigonal planar structures at the active sites for a significant number of phosphoryl transfer enzymes has been resolved.

The use of transition state analogues (TSAs) gives insight into the profile of active sites that are essential to catalysis. Those enzymes that produce the biggest rate acceleration and TS affinities should become the most sensitive targets for inhibitor design. There are already some successful examples of using the TSA agents to control hypertension, the spread of HIV, the maturation of insects, and the growth of weeds.(Schramm, 1998) There are quite a number of analogues which are capable of imitating appropriately either the geometry and/or the charge of the TS for phosphoryl transfer.

There are two types of metal fluoride complexes that can mimic the phosphoryl group in transfer: tetrafluoroaluminate (AlF_4^-) and trifluoromagnesate (MgF_3^-) (**Figure 1.22**). Tbp MgF_3^- is isoelectronic and isosteric to PO_3^-, but very low populated in the solution

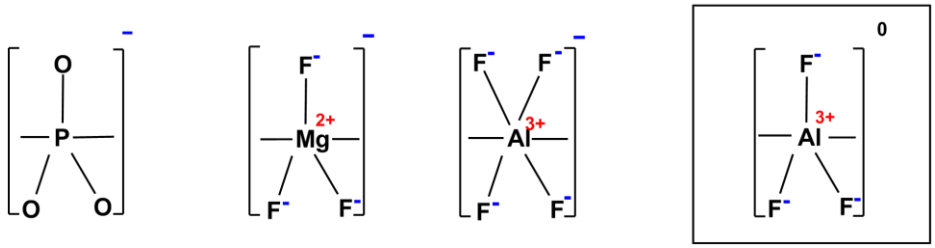


Figure 1.22 The mimics of phosphoryl group in transfer by metal fluoride moiety: MgF_3^-, AlF_4^-, and so called 'AlF_3^{0}' in some structures in PDB.

because it is an inherently unstable species. By comparison, octahedral AlF_4^- can be observed by ^{19}F NMR in a solution of Al^{3+} and F^- is enormously more stable, while the closer proximity of Al and P in the Periodic Table makes the character of Al–F bond more like P–O bond. Indeed, the Al–F bond (1.6 Å ~ 1.7 Å) is more similar in length to a P–O bond (1.5 Å ~ 1.6 Å).(Chabre, 1990)

1.9 The Concept of Charge Balance

Differences in the binding affinity for related substrate analogue inhibitors of some nucleotide kinases has been known for a long time ago. A key example is that in the active site of adenylate kinase, diadenosine polyphosphate compounds competitively inhibit the binding of ATP plus AMP. Ap_5A binds 1000 fold tighter than Ap_4A (Feldhaus *et al.*, 1975) and this seems to be because it has only one negative charge less than the six negative charges of the transition state.(Williams *et al.*, 1998) Ap_6A has the same number charges as the true TS but the hexa-phosphate chain is now too large for the space between the two nucleotide binding sites, therefore the affinity is compromised (**Figure 1.23**). However no one interpreted such results in the light of the need for the charges on the inhibitors to be closely similar to the net charge in the TS to achieve optimum binding.

The Charge Balance Concept for the transition state was first established by Williams and Blackburn when they tested the binding of stable bisubstrate ligands for phosphoglycerate kinase.(Williams *et al.*, 1998) In this work, they highlighted the requirement that, in order to gain better charge approximation to the TS for stronger affinity, an additional phosphate residue could be introduced in the bisubstrate ligand but with consequent loss of steric compatibility.

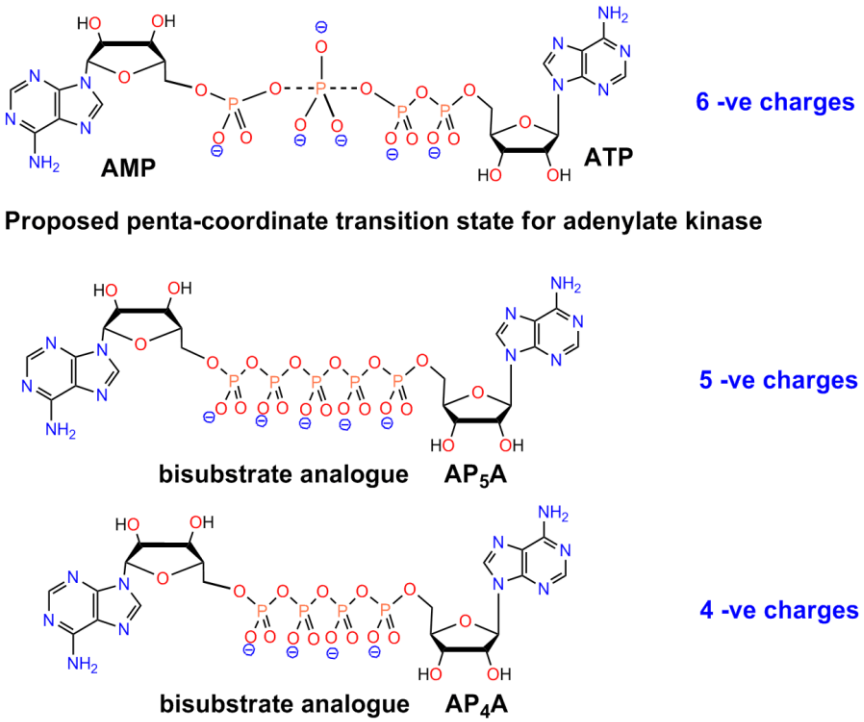


Figure 1.23 Bisubstrate analogues Ap_5A and Ap_4A mimic the transition state of the adenylate kinase reaction: AMP + ATP → 2ADP.

This concept has been developed further by Blackburn in addressing Charge Balance for phosphoryl transfer enzymes as being essential for TS stabilisation during catalysis.(Bowler *et al.*, 2010) The highly localised negative charge on phosphate residues demands that phosphoryl transfer enzymes have to overcome a charge repulsion barrier, the "anionic shield", to bring the two substrates together and stabilise the TS by neutralising its net negative charge (**Figure 1.24**).(Bowler *et al.*, 2010)

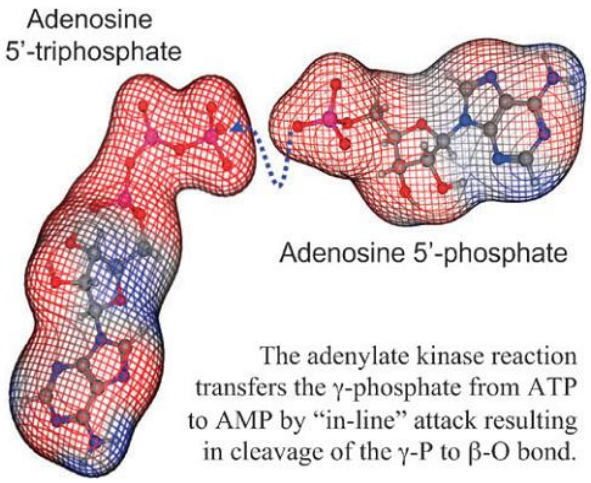


Figure 1.24 Electrostatic surfaces calculated for the ATP tetra-anion (left) and the ADP dianion (right). (Negative polarity – red; positive polarity – blue).(Bowler *et al.*, 2010)

Encouraging experimental support for this proposal was later provided for the K219A mutant of human phosphoglycerate kinase to establish the importance of Charge Balance in catalysis. This mutant cocrystallises with an octahedral trifluoroaluminate TSA in the active site, as confirmed by ^{19}F NMR.(Cliff *et al.*, 2010) The deletion of one positive charge from Lys 219, previously proved to be a key residue interacting with the 1-phosphate of 1,3-bisphosphoglycerate in the TS, directly leads to the corresponding loss of a negative charge from the bound TSA aluminium fluoride species, from AlF_4^- to $AlF_3(H_2O)^0$ (**Figure 1.25c**). The average chemical shifts move downfield because of the loss of electron density from the $AlF_3(H_2O)^0$ moiety, which is consistent with the replacement of F by a less electronegative species H_2O (**Figure 1.25a** and **b**).(Cliff *et al.*, 2010)

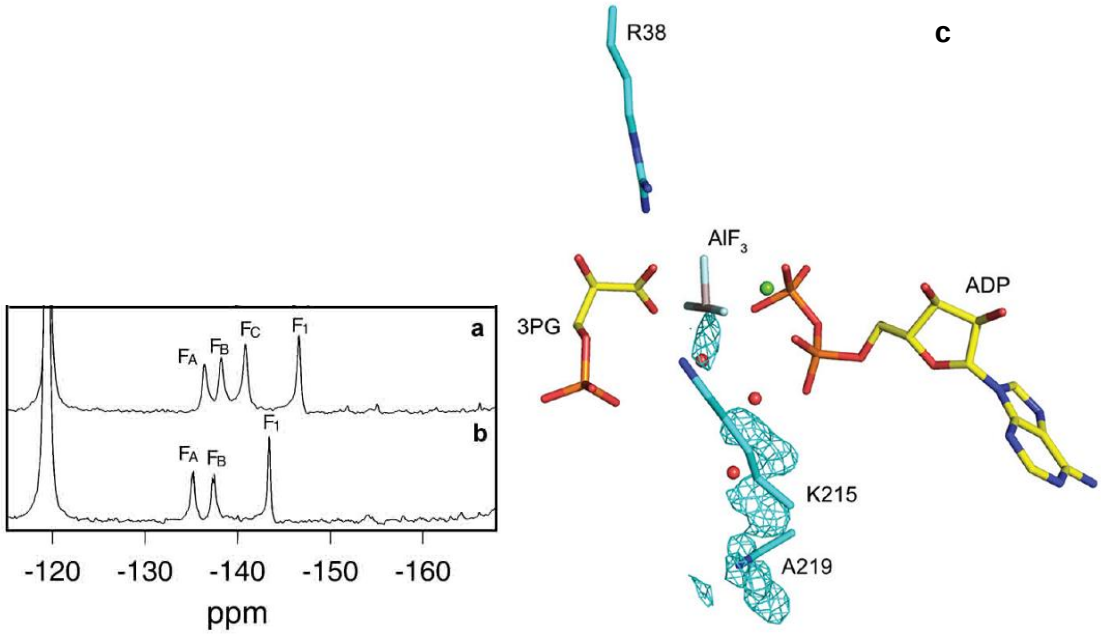


Figure 1.25 (**a**) ^{19}F NMR spectrum of HsPGK-3PG-AlF_4^--ADP TSA complex (F_A = –136, F_B = –138, F_C = –141, F_1 = –146 ppm). (**b**) ^{19}F NMR spectrum of HsPGK(K219A)-3PG-$AlF_3(H_2O)^0$ -ADP TSA complex (F_A = –135, F_B = –137, F_1 = – 143 ppm). (**c**) The $AlF_3(H_2O)^0$ moiety in the HsPGK(K219A)-3PG-$AlF_3(H_2O)^0$-ADP TSA complex is octahedral. The $2F_o – F_c$ electron density difference Fourier map between the HsPGK-3PG-AlF_4^--ADP TSA and the HsPGK(K219A)-3PG-$AlF_3(H_2O)^0$-ADP complexes is shown as a cyan mesh contoured at 2σ, with the 3PG and ADP ligands and the side chains coordinating the $AlF_3(H_2O)^0$ moiety in the HsPGK(K219A)-3PG-$AlF_3(H_2O)^0$-ADP active site shown as sticks. The water and Mg^{2+} are shown as red and green spheres, respectively.(Cliff *et al.*, 2010)

An equivalent reciprocal chemical mutation has also been observed as a consequence of the charge-reducing protein mutation in βPGM_{K145A}.(Griffin, PhD Thesis, 2011) This likewise leads to the observation that in ^{19}F NMR in the absence of aluminium, two peaks can be seen corresponding to βPGM_{K145A}-G6P-MgF_2^0. An X-ray structure (to be

published) shows the magnesium adopts octahedral geometry of which the sixth ligand is a coordination to the backbone carbonyl of Ser 114 along with a water molecule and two fluorine ligands (**Figure 1.26**). For aluminium fluoride complexes, there is some inconsistency between the ^{19}F NMR results and crystal data (**Figure 1.27**). In ^{19}F NMR spectra, when NH_4F and $AlCl_3$ are added to the sample, the dominant complex is βPGM_{K145A}-AlF_4^--G6P-NH_4^+ complex which gives four fluoride signals because an ammonium cation sits in the cavity of the deleted lysine and coordinates to one of the fluorides. The minor species (< 10%) was identified as a βPGM_{K145A}-AlF_4^--G6P-Na^+/H_3O^+ complex by means of a control using NaF as fluoride source. Because D-glucose 6-phosphate used for all the experiments as purchased from Sigma is a disodium salt, even though NH_4F was used, when 10 mM G6P was added unavoidably there was 20 mM Na^+ in the sample and so the minor form was generated. The integration difference of the two forms shows the binding of NH_4^+ is at least 20 times stronger than Na^+/H_3O^+ for the same site. This may be because of the relative sizes and the hydrogen bonding capability between the two different cations. Na^+ has an ionic radius of 116 pm which is too small for the space gap but hydrated sodium cation is too big. By contrast, that for NH_4^+ is 175 pm.

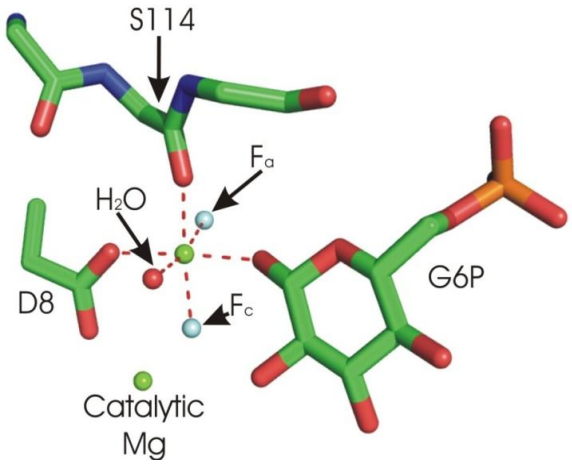


Figure 1.26 Crystal structure of the active site of βPGM_{K145A}-MgF_2^0-G6P TSA complex showing an octahedral geometry around the central magnesium ion. The four equatorial ligands are two fluorides, one water molecule and the backbone carbonyl group of Ser 114.

To identify the commonality of charge balance in the active site of enzymes that catalyse phosphoryl group transfer, structures have been downloaded from the PDB that contain metal fluoride TSAs (AlF_x and MgF_3^-). Some analyses have also been made using MgF_4^{2-} complexes). The metal fluorides have been assigned as having the Pauling charge of +2.51 on the central metal as having one net negative charge on the PO_3^- and assigning the Pauling charge for an oxygen anion to each of the three fluorines.

Standard Pauling charges can then be assigned to all the additional atoms in the active site. The sum of all the partial charges on atoms within a sphere of a given radius from the transferring phosphorus atom (or its surrogate metal ion) is plotted as a function of the radial distance from the phosphorus atom as origin.

The first results were obtained for a selection of three small molecule kinases (nucleoside-diphosphate kinases, NDP kinases, human phosphoglycerate kinase, HsPGK, and UMP-CMP kinase) and four G proteins (Cdc42, RhoA, Ras and $G_{i\alpha1}$) along with nitrogenase as a representative non-kinase, ATP-using enzyme (**Figure 1.28**). The results are strongly supportive of the CBH. Within a 15 Å shell, the net charge for these phosphoryl transfer enzymes remains focused at less than ±3. (Nitrogenase is a notable exception). The reason that the CB curves become more scattered outside 10 Å may reflect that the demand on CB is not as great as within the active site sphere. In addition, in most cases the distance from the phosphate binding site to the nearest surface of the globular enzymes is often not bigger than 15 Å because of the limiting size (20 ~ 30 kDa) of these small molecule kinases and small G proteins.

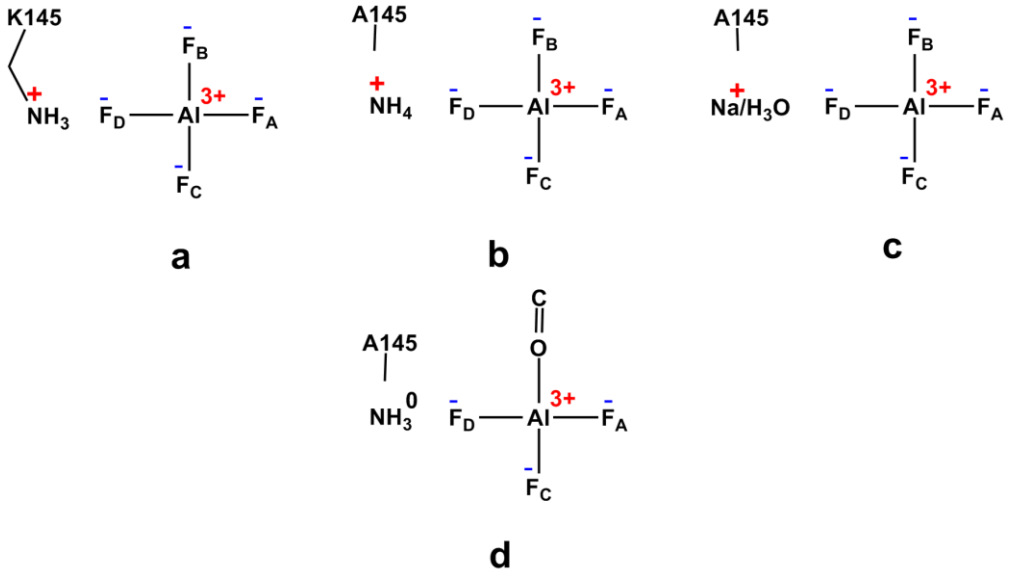


Figure 1.27 Schematic diagram of the fluoroaluminate complexes of βPGM. (**a**) wt βPGM-AlF_4^--G6P TSA complex showing the side chain of Lys 145 coordinating the F_D position to balance the charge of the metal fluoride moiety. (**b**) The dominant species in the ^{19}F NMR spectra of the βPGM_{K145A}-AlF_4^--G6P TSA complex, showing how ammonium ions are positioned in the F_D coordination site and therefore balance the charge of the AlF_4^- metal fluoride moiety. (**c**) The dominant species in the NMR spectra of the βPGM_{K145A}-AlF_4^--G6P TSA complex formed in the absence of ammonium ions: four fluorine nuclei are present, and therefore the charge of the AlF_4^- moiety must be balanced by a negatively charged ion coordinating F_D (the nature of this coordinating ion is not known). (**d**) The dominant species in the crystal structure of the βPGM_{K145A}-AlF_3^0-G6P TSA complex: the backbone carbonyl group

of Ser 114 coordinates the aluminium in the F_B position and neutral ammonia coordinates F_D in order to balance the charge of the neutral AlF_3^0 moiety.(Griffin, PhD thesis, 2011).

Based on this, the Charge Balance Hypothesis (CBH) was advanced.(Bowler *et al.*, 2010) The CBH proposes that local charge balance, rather than native geometry, dominates the stabilisation of the TS for phosphoryl transfer enzymes. While the results of CB analysis will be described in detail in due course for the five enzymes that have been studied in this work (**Chapters 3 ~ 7**), it has also been extended to a range of other key enzymes.

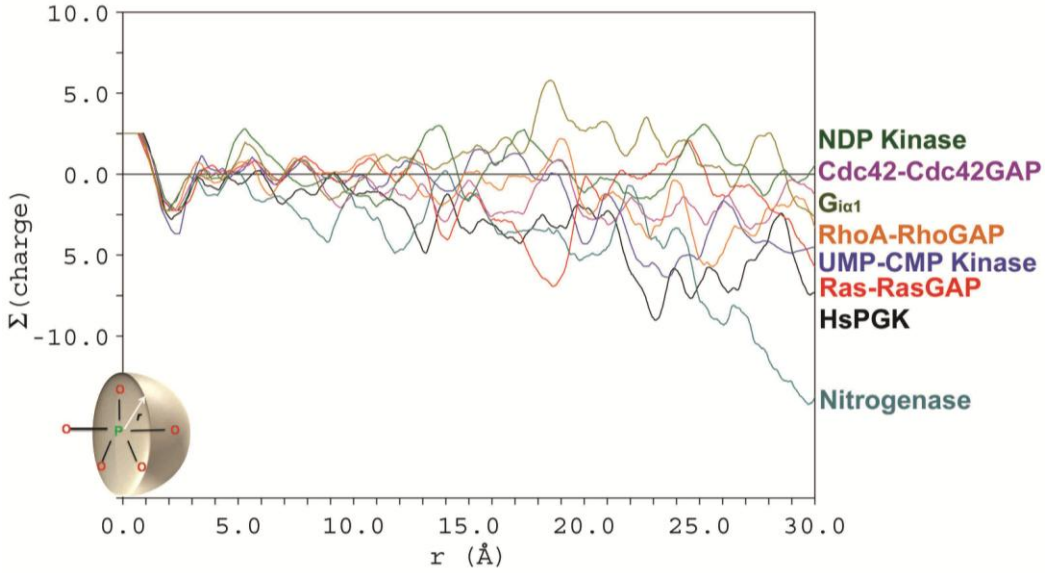


Figure 1.28 Charge balance for seven phosphoryl transfer enzymes as listed and one non-phosphoryl transfer enzyme to compare the charge distribution within certain radius. Nitrogenase (cyan curve) apparently drops out at around 9, 11, and 25 Å.(Cliff *et al.*, 2010)

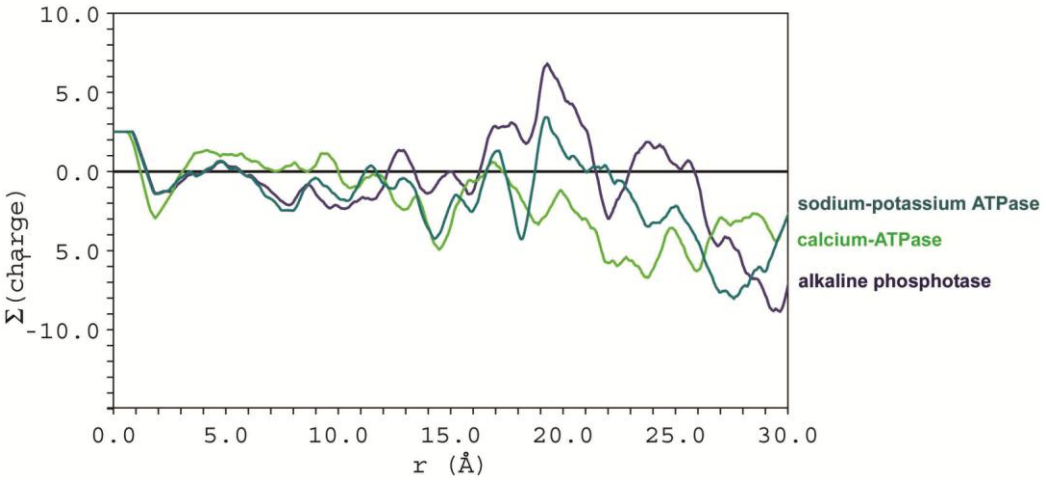


Figure 1.29 Charge balance for two ATPases: sodium-potassium ATPase (PDB: 2ZXE) and calcium ATPase (PDB: 3FGO) and alkaline phosphatase (PDB: 1KH9).(Jin, unpublished)

Figure 1.29 shows the CB analysis for alkaline phosphatase and for two ion-transporting membrane ATPases. For all of them, the net charge stays within ±3 out to 15 Å. All of these calculated results share the common feature that the prevalence of CB among many phosphoryl transfer enzymes is so effective that there is always local, near zero net charge within some 15 Å radius of the transferring phosphate.

1.10 ^{19}F NMR is Selected to Resolve our Problems

Nuclear magnetic resonance (NMR) is a technique that relies on nuclear spin, which is specific to each isotope of an element. So only nuclei with an odd mass number that have half-integral spin numbers and nuclei with even mass numbers and odd atomic numbers that have an integral spin number are NMR active. Nuclei with an even mass number and even atomic numbers have zero spin, so they are NMR inactive.(Cavanagh *et al.*, 2007) The most common nuclei are the ones with a spin number of ½, such as ^{1}H, ^{13}C, ^{15}N, ^{19}F and ^{31}P. ^{9}Be has a spin number of $^{3}/_{2}$ and therefore possesses a quadrupole moment and that has an effect on the fluorine resonance signal in the beryllium fluoride complex. The advantage of NMR is it is able to probe time-dependent phenomena such as reaction kinetics and intramolecular dynamics in a solution. Compared to crystallography, it is particularly useful to study proteins in their native state in solution.

The ^{19}F nucleus has a spin number of ½ and this 100% abundant nucleus with its high sensitivity (83% of ^{1}H), spectral simplicity, and large chemical shift range is attractive for protein NMR *in vitro*.(Gerig, 1989) The chemical shift of the fluorine signals can reflect the local Van der Waals environment as well as local electrostatic fields. Solution ^{19}F NMR was first applied to protein systems in the late 1960s.(Spotswood *et al.*, 1967) ^{19}F NMR has been most used to probe the protein conformational change in their native solution environment by labelling the aromatic rings in tyrosine, phenylalanine, and tryptophan.(Salopek-Sondi *et al.*, 2002) The extreme sensitivity of the ^{19}F resonance to its atomic surroundings can reveal important structural and kinetic features of protein conformational changes.(Danielson *et al.*, 1996) It was then found to be useful simply for observing the formation of aluminium fluoride complex in G proteins.(Higashijima *et al.*, 1991; Hoffman *et al.*, 1998; Graham *et al.*, 1999; Praefcke *et al.*, 1999)

Multidimensional NMR is useful to probe structures in solution for small proteins (M_r < 40 kDa), but many proteins of known crystallographic structure are too large for

multidimensional NMR though they still fall within the range of molecular weights accessible to solution ^{19}F NMR (M_r < 100 kDa).(Danielson *et al.*, 1996)

The first direct observation of an MgF_3^- TSA complex with ^{19}F NMR was done in Sheffield, with measurements of ^{19}F SIIS, {^{19}F}^{1}H-NOE, and {^{19}F}^{1}H,^{15}N-HSQC. (Baxter *et al.*, 2006) Later on, the application of ^{19}F has been extended to study the switch between central metals in TSAs in pH titrations, the use of presaturation to observe the obstructed signals in TSA complexes, and so on. It has now become a major, powerful tool to study metal fluoride complexes in phosphoryl transfer (and also sulfuryl transfer) enzymes from all angles.

Throughout this thesis, NMR and X-ray crystallography have been used as complementary techniques to report on the features of a protein complex both in solution and in crystal states. But more importantly, in the study of transition state analogues of metal fluoride complexes, these two techniques are able to validate each other. This is especially important where the number of fluorines coordinating to the metal directly determines the net charge of the TSA, as this is proving to be critical for the interpretation of the mechanism of phosphoryl transfer in enzymes.

1.11 Aims of this Thesis

The primary objective of this thesis is to understand the largely-structural details of enzyme catalysis on phosphoryl group transfer for five archetypical phosphoryl transfer enzymes: one phosphomutase – βPGM, one phosphatase – PSP, one protein kinase – cAPK, one GTPase – small G protein RhoA-RhoGAP, and one small molecule kinase – UMP/CMP kinase.

Throughout this thesis, ^{19}F NMR as a tool for studying TSs is further developed. Changes of conformation and charges in the active site can be identified by changes of the chemical shift and solvent isotope shifts.

This thesis also seeks to correct errors in the published assignment of the trigonal bipyramidal TSA complexes designated as AlF_3 in solid state structures at high pH, to establish that at least some of these tbp metallofluoride moieties observed in the crystal structures are indeed MgF_3^-, irrespective of pH.

Based on these analyses of TSAs, this thesis generates good evidence to support the Charge Balance Hypothesis which endorses the prioritisation of charge over geometry in these phosphotransfer enzymes.

The specific goals for each protein are as follows:

For βPGM (Chapter 3):

1. To examine the selectivity of the wildtype protein for six stable synthetic analogues of βG1P on 1-α-OH and stereo configuration of the CHFP.

2. To investigate the effects of single point mutations on general acid/base Asp 10 and nucleophilic Asp 8 in activity assays and TSA complexes, using NMR and crystallography as tools.

3. To identify some of the ambiguous TSA and GSA complexes observed in the wildtype protein by mutation.

For PSP (Chapter 4):

1. To investigate the effects of single point mutations on general acid/base Asp 13 in activity assays and TSA complexes, using NMR.

2. To compare between PSP_{D13N} and βPGM_{D10N} as well as their wild types are made to identify functional and constitutional differences between PSP and βPGM.

For cAPK (Chapter 5):

1. To resolve inconsistencies in the literature to show if cAPK conforms to the CBH to form an MgF_3^- TSA complex and an AlF_4^- TSA complex in solution.

2. To measure the binding constant of MgF_3^- and AlF_4^- TSA complexes by aluminium titration monitored by ^{19}F NMR.

3. To produce a trifluoroberyllate GSA complex for composing the whole reaction coordination of cAPK.

For RhoA-RhoGAP (Chapter 6):

1. To confirm the PIXE analysis of RhoA-GDP-MgF_3^--RhoGAP TSA complex and make RhoA-GDP-AlF_4^--RhoGAP complex to test the role of charge balance in activation of G protein by GAP.

2. To test if RhoA obeys CBH on its own without the support of RhoGAP making a TSA complex using ^{19}F NMR.

3. To test if RhoA-$RhoGAP_{R85A}$ obeys CBH using ^{19}F NMR and crystallography.

4. To test the hydrolytic activity of a substrate analogue GTPγF for RhoA, with or without the contribution from an arginine finger in RhoGAP using ^{31}P NMR.

For UMP/CMP kinase (Chapter 7):

1. To prove the trigonal bipyramidal AlF_3 observed at high pH in the active site of a solid state structure is MgF_3^- by pH titration monitored by ^{19}F NMR.

2. To test if Charge Balance Hypothesis is validated for UMP/CMP kinase by ^{19}F NMR using metal fluoride TSA complex.

3. To explore the residual adenylate kinase activity in UMP/CMP kinase.

Chapter 2. General Experimental Methods

2.1 Measurement of pH

For NMR samples a Russell Thermo CMAW711/3.7/180 combination, single junction pH electrode microprobe was used at 25˚C unless otherwise indicated.

For general buffers a 'large probe' attached to a Russell RL150 pH meter was used.

Both of the pH meters were calibrated before use with Scientific Laboratory Supplies pH buffer at pH 4.0, pH 7.0, and pH 10.0 until the R factor was above 0.93.

In the case of pH measurement for small volume in pH titration experiments using high concentration of protein in Tris buffer, another strategy was used as described in **Chapter** 7 UMP/CMP Kinase.

2.2 Polymerase Chain Reaction (PCR)

The PCRs were carried out in a Techne Progene™ thermal cycler (fitted with hot top assembly). The QuikChange™ Site-Directed Mutagenesis Kit was used for the reactions. All the components were kept on ice during the setup and added to a sterile 0.5 mL Eppendorf tube in order before the addition of DNA polymerase. All the reactions had a total volume of 50 μL. The detailed reaction conditions for each protein were listed in the corresponding chapters.

After primers design for site-directed mutagenesis, the melting temperature was calculated using website (http://www.basic.northwestern.edu/biotools/oligocalc.html), which also takes into account potential hairpin formation and self-annealing sites in the primers.

2.3 Agarose Gel Electrophoresis

Electrophoresis grade agarose (1.0 g) was made up to 100 mL with 1× TAE or 1× TBE buffer and heated in a microwave oven. When the solution had cooled, 0.001% DNA stain Gel Red (Biotium, Cat. No. 41003) was added. The agarose solution was poured into the taped BioRad minisub DNA cell and left to cool flatwise for about 30 min until the gel had set.

Table 2.1 Compositon of buffers and solutions for the agarose gel electrophoresis.

50 × TAE Buffer	2 M	Tris-HCl
	1 M	AcOH
	0.05 M	EDTA (pH 8.0)
	filtered	
5 × TBE Buffer	0.45 M	Tris-HCl
	0.45 M	boric acid
	0.01 M	EDTA (pH 8.0)
	filtered	
Loading Dye	40% (w/v)	sucrose
	0.25% (w/v)	bromophenol blue
	0.25% (w/v)	xylene cyanol
Agarose Gel	1–1.5% (w/v)	agarose

Agarose electrophoresis gel was run for 70 min at 110 V. 1 μL Fermentas or New England Biolab 1 kb DNA ladder was used as a marker.

2.4 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Table 2.2 Composition of buffers and solutions for SDS-PAGE.

4× Lower Buffer	1.5 M	Tris-HCl, pH 8.8
	14 mM	SDS
4× Upper Buffer	0.5 M	Tris-HCl, pH 6.8
	14 mM	SDS
4× Loading Buffer	200 mM	Tris-HCl, pH 6.8
	400 mM	DTT(not needed if no cysteine)
	8%	SDS (w/v)
	0.4%	bromophenol blue (w/v)
	40%	glycerol (v/v)
5× Tris-glycine Electrophoresis Buffer	125 mM	Tris
	960 mM	glycine
	0.5%	SDS (w/v)
16% Resolving Gel	2.5 mL	4× Lower Buffer
	4 mL	40% acrylamide/Bis (37.5:1)
	3.5 mL	ddH_2O
	100 μL	10% APS
	10 μL	TEMED
4% Stacking Gel	2.5 mL	4× Upper Buffer
	1.125 mL	40% acrylamide/Bis (37.5:1)
	6.375 mL	ddH_2O
	110 μL	10% APS
	11 μL	TEMED
0.05% Coomassie Blue Stain	0.05% (w/v)	Coomassie Blue
	7% (v/v)	glacial acetic acid
	40% (v/v)	methanol
	filtered	
Destain	45% (v/v)	methanol
	10% (v/v)	glacial acetic acid

The SDS-PAGE gels were made several hours before use. All the samples were heated with SDS loading buffer in a 1.5 mL Eppendorf at 95°C for 5 min before being loaded

into the 16% SDS-PAGE gel. All gels were run on a BioRad Power Pac 300 at 80 V for 10 min then 160~180 V until the blue loading dye ran out of the gels in Tris-glycine electrophoresis buffer. Gels were stained in Coomassie Blue stain solution for 1 h or overnight then destained until the background of the gel was colourless.

2.5 Urea Gel

Urea gels were made several hours before use. The protein sample was mixed with loading buffer without heat denaturation. All gels were run on a BioRad Power Pac 300 at 22 mA under 'constant current' mode for 90 ~ 100 min until the blue loading dye ran out of the gels in Tris-glycine running buffer. The gels were stained in Coomassie Blue stain solution for 1 h or overnight then destained until gel background was colourless.

Table 2.3 Composition of buffers and solutions for urea gels.

4× Lower Buffer	500 mL	ddH_2O
	182g	Tris-HCl
	4.5 M	urea
	pH to 9.1	
	make up to 1 L	
4× Upper Buffer	500 mL	ddH_2O
	60.5g	Tris-HCl
	4.5 M	urea
	pH to 9.1	
	make up to 1 L	
4× Loading Buffer	100 mM	Tris-HCl, pH 6.8
	0.2%	bromophenol blue (w/v)
	20%	glycerol (v/v)
	400 mM	DTT
5× Tris-glycine Electrophoresis Buffer	125 mM	Tris
	1250 mM	glycine, pH 9.1-9.3
	0.5%	SDS (w/v)
7% Resolving Gel	2.5 mL	4× Lower Buffer
	1.75 mL	40% acrylamide/Bis (37.5:1)
	5.75 mL	ddH_2O
	100 µL	10% APS
	10 µL	TEMED
4% Stacking Gel	2.5 mL	4× Upper Buffer
	1.125 mL	40% acrylamide/bis (37.5:1)
	6.375 mL	ddH_2O
	110 µL	10% APS
	11 µL	TEMED

2.6 Production of Competent Cells ($CaCl_2$ Method)

Aliquots of cells were thawed on ice and 1 µL aliquots were transferred into four 20 mL universals each of which contained 5 mL LB media. The culture was incubated at 37°C with shaking at 250 rpm overnight. LB media (100 mL in a 250 mL flask) was

inoculated by 2 mL of overnight culture and incubated at 37°C at 250 rpm until OD_{600} reached 0.6~0.8 which indicated that the cells had grown to log phase. The cell culture was aseptically transferred to 2 sterile ice-cold 50 mL Falcon centrifuge tubes which were then stored on ice for 10 min. The cells were recovered by centrifugation at 4000 rpm, 4°C for 10 min and the pellets were resuspended in 2 mL ice-cold 0.1 M fresh $CaCl_2$. 70 µL DMSO was added to the resuspended cells and the tubes were swirled gently before being returned to the ice bath. After 15 min, another 70 µL DMSO was added. The ice-cold suspension were divided into 200 µL aliquots in chilled Eppendorf tubes and stored at –80°C.

2.7 Transformation of Competent Cells

Aliquots (200 µL) of competent cells were thawed on ice and 1 µL plasmid was added. The Eppendorf tubes with cells were incubated on ice for 30 min. Then the tubes were placed in a preheated circulating water bath at 42°C for exactly 45 s and immediately put onto an ice bath for 1~2 min to chill the cells. 800 µL LB media was added to the tubes and the culture was incubated at 37°C, 250 rpm for 1 h to allow the bacteria to recover and to express the antibiotic resistance marker encoded by the plasmid. After 1 h, 50 µL, 100 µL, and 200 µL transformed cells were transferred by pipette respectively onto the LB agar plate containing the appropriate antibiotic and incubated at 37°C overnight.

2.8 Minipreps

A single colony was picked up by a tip from the agar plate and incubated at 37°C for 12~16 h and dipped into the universal containing 5 mL LB media with 100 µg/mL ampicillin. The inoculated media was incubated at 37°C overnight on a 250 rpm shaker. Then all the harvest and extraction procedures were carried out by following the Qiagen QIAprep Mini-prep Kit strategy to obtain up to 50 µg/mL of plasmid.

2.9 Large Scale Protein Overexpression

The large scale overexpression was carried out at different conditions, such as LB or minimal media, *E. coli.* strain, incubation time length and temperature, OD_{600} value for expression induction, antibiotics used, *etc*. In this thesis, the details are listed in each chapter in the experimental section.

2.9.1 LB media

For each litre, the following ingredients were added to 1 L ddH_2O:

10% (w/v)	tryptone
10% (w/v)	NaCl
5% (w/v)	yeast extract

The solution was sterilised by autoclaving at 126°C.

2.9.2 M9 minimal media

For each litre, the following compounds were added to 950 ml ddH_2O[5]:

6 g	Na_2HPO_4
3 g	KH_2PO_4
0.5 g	NaCl

The pH was adjusted to 7.4 and the volume was made up to 1 L with ddH_2O.[5]The media was then sterilised by autoclaving and the following chemicals, along with the appropriate antibiotics, were added when the media had cooled and just before use:

650 μL	Trace elements	
3 g	Glucose[2,3]	
1 mL	Thiamine	(1 mg mL^{-1})[3]
2 mL	$(^{15}NH_4)_2SO_4$	(0.5 mg mL^{-1})[3]
1 mL	1 M $MgSO_4$	(autoclaved)
0.1 mL	1 M $CaCl_2$	(autoclaved)[4]

Notes:

1 – Trace elements. For 100 mL final volume the following compounds were added to 70 mL ddH_2O:

550 mg	$CaCl_2 \cdot 2H_2O$
140 mg	$MnSO_4 \cdot H_2O$
40 mg	$CuSO_4 \cdot 5H_2O$
220 mg	$ZnSO_4 \cdot 7H_2O$
45 mg	$CoCl_2 \cdot 6H_2O$
26 mg	$Na_2MoO_4 \cdot 2H_2O$
40 mg	H_3BO_3
26 mg	KI

The pH was adjusted to 8.0 before adding 500 mg EDTA, then the pH was readjusted to 8.0 before adding 375 mg $FeSO_4 \cdot 7H_2O$. The solution was then made up to 100 mL volume with ddH_2O before being sterilised by autoclaving.

2 – Uniformly labelled glucose ($^{13}C_6$-D-glucose or $^{13}C_6$-D-glucose-1,2,3,4,5,6,6-D7) was used here if required.

3 – 0.2 μm syringe-filter sterilised.

4 – $CaCl_2$ was added last and the flask immediately agitated to disperse the white precipitate that formed. The preparation was abandoned if the precipitate was not successfully dispersed.

5 – For D, ^{15}N, ^{13}C triple labelled proteins, D_2O was used in all the buffers except trace elements.

2.10 Measurement of Optical Density of Cell Cultures

The optical densities of the cell cultures were measured using a WPA lightwave S2000 UV/vis spectrophotometer at a wavelength of 600 nm. The OD_{600} of the cells cultures was determined by comparison with a blank which contained only media.

2.11 Cell Harvest and Protein Extraction

2.11.1 Harvesting cells

After the large scale expression, the culture was transferred into six 0.5 L centrifuge bottles and spun at 5,000 rpm for 30 min at 4°C. The supernatant was discarded and the pellets were resuspended in 10 mL autoclaved LB. The suspension was poured into a 50 mL Falcon centrifuge tube and centrifuged at 7,000 rpm for 30 min at 4°C. The supernatant was discarded and the pellets were stored at 80°C before purification.

2.11.2 Cell resuspension and lysis

The frozen cell pellets were suspended in 50 mL lysis buffer (the ingredients of the lysis buffer refer to the experimental section of each chapter for each protein) dissolving a small amount of RNase and a protease inhibitor cocktail tablet (EDTA or EDTA-free depends on the next purification step). The suspension in Falcon tube was sonicated for 30 s on an ice bath 5 ~ 10 times with 60 s pauses in between to let the suspension cool. The lysate was centrifuged at 12,000 g, 4°C for 30 min. The pellets were discarded and supernatant was filtered using 0.22 µm syringe filter. For phosphoserine phosphatase (PSP) or its mutant D13N encoded from the thermophilic archaea *Methanocaldococcus jannaschii*, the supernatant was heated at 70°C for 30 min to denature most of the proteins from *E.coli.* before filtration. The heated suspension was then spun at 70,000 g, 10°C for 30 min followed by 0.22 µm filtration.

2.11.3 Concentration of protein and buffer exchange

The protein solution was concentrated at 4°C in a 10,000 Da MWCO Sartorius Vivaspin^{TM} which was spun at 5,500 g until the required final volume was reached. For buffer exchange, the concentrated protein solution was diluted with the buffer to be changed into and then concentrated again. The same procedure should be repeated several times according to the purpose of the experiment.

2.12 The Measurement of Protein Concentration

2.12.1 UV Spectroscopy

The principle of this method is based on the Beer-Lambert Law: $A = \varepsilon c \lambda$ where A is the absorbance of the protein sample and ε is the extinction coefficient of the protein with the unit such as M^{-1} cm^{-1} or mg cm^{-2}; λ is the path length which is usually 1 cm; c is the concentration of the protein sample whose unit depends on which extinction coefficient unit is chosen. Absorbance of purified protein solution was measured in 500 µL or 1 mL volume quartz cuvettes at 200-350 nm on a WPA lightwave S2000 UV/vis spectrophotometer The concentrated samples were diluted to $A_{280\ nm}$ within the range 0.1 ~ 1.

2.12.2 The BioRad Bradford assay

The BioRad Bradford assays were carried out on a WPA lightwave S2000 UV/vis spectrophotometer at 595 nm. Test sample (20 µL) or standard BSA was mixed with 1 mL 1× dye reagent within 1 h at room temperature before the test. A standard concentration curve was plotted by reading the absorbance of 0.2 ~ 0.8 at 595 nm against the known concentrations of bovine serum albumin (BSA) protein. The concentration of the sample was then read from the curve. When the unknown protein sample is too concentrated, dilution is needed to make the absorbance lie within the range 0.2~0.8.

2.13 Measurement of DNA/RNA Concentration

A WPA lightwave S2000 UV/vis spectrophotometer was used to measure the absorbance of DNA/RNA solution in 500 µL volume quartz cuvettes at 200~300 nm. The plasmid/circular DNA was diluted until the OD_{260nm} was between 0.1 ~ 1.0 for the accuracy of the measurement.

For ssDNA, if $OD_{260\ nm} = 1$, $C = 33\ \mu g\ /mL = 0.10\ mM$;

For dsDNA, if $OD_{260\ nm} = 1$, $C = 50\ \mu g\ /mL = 0.15\ mM$;

For ssRNA, if $OD_{260\ nm} = 1$, $C = 40\ \mu g\ /mL = 0.11\ mM$.

For example, the dsDNA/plasmid is diluted by 200 fold, and the OD_{260nm} is 0.351, then the concentration of the original solution $C' = OD_{260nm} \times C \times$ dilution factor = $0.351 \times 0.15 \times 200 = 2.317$ mg/mL.

2.14 NMR

All ^{1}H, ^{15}N-HSQCs are carried out using ^{15}N-labelled protein. The 1D ^{1}H NMR spectra were acquired either on a Bruker Avance 500 MHz spectrometer (operating at 470.59 MHz for fluorine) equipped with a 5-mm QXI probe equipped with z-axis gradients. 1D ^{31}P NMR spectra were recorded on a Bruker Avance 500 MHz spectrometer with a 5-mm multinuclear probe tuned to ^{31}P (202.5 MHz). 2D high-sensitivity HSQC NMR spectra with fluorine decoupling were recorded on a Bruker Avance 500 MHz spectrometer equipped with $^{1}H/^{15}N/^{13}C/^{19}F$ QNF probe and pulse field *z* axis gradients. All spectra were recorded at 25°C unless specifically stated. Proton chemical shifts were referenced relative to internal DSS (sodium 2,2-dimethylsilapentane-5-sulfonate) at 0.0 ppm. ^{15}N, ^{31}P and ^{19}F chemical shifts were calculated indirectly using the following gyromagnetic ratios: $^{15}N/^{1}H = 0.101329118$, $^{31}P/^{1}H = 0.404807356$ and $^{19}F/^{1}H = 0.940940080$. All the samples contained D_2O capillary or 10% internal D_2O as deuterium lock. 2D ^{1}H,^{15}N-HSQC, and 1D ^{19}F NMR spectra were processed and analysed using FELIX 2000 software (Felix NMR Inc, San Diego, CA). Integration of peaks was performed with phase adjustment and baseline correction optimized for the chosen peaks. 1D ^{1}H and ^{31}P NMR spectra were processed using Bruker Topspin 1.3.

Solvent-induced isotope shifts (SIIS) on the ^{19}F resonances were measured by comparing spectra for samples prepared in 10% D_2O and 100% D_2O buffer and are defined as $^{19}F\delta(H_2O\ buffer) - {}^{19}F\delta(D_2O\ buffer)$. All samples were incubated within the spectrometer for 30 min prior to NMR acquisition.

2.15 Isotope Shift Analysis

In 1952 Ramsey predicted the isotope effect phenomenon between isotope atoms in NMR.(Ramsey, 1952) In the following year, this phenomenon was proved

experimentally by Wimett in proton NMR resonances.(Batiz-Hernandez, 1967) Isotope shifts were first reported in fluorine NMR in 1957 when the proton in the $–CF_2H$ group was replaced by deuterium and the signal moved 0.60 ± 0.05 ppm upfield.(Van Dyke Tiers, 1957) Early research proposed that deuterium isotope shifts arising from different electric field effects were solely caused by the different bond lengths of X–H and X–D.(Gutowsky, 1959) Nowadays, this theory has been shown to be incorrect, however it may still apply for strongly polarised hydrogen bonds in proteins.(Tuchsen *et al.*, 1991; Sosnicki *et al.*, 2007)

For protein ^{19}F NMR, solvent–induced isotope shifts (SIIS) are defined as SIIS = $\delta X(H_2O) - \delta X(D_2O)$ where $\delta X(H_2O)$ is the chemical shift of nucleus X in 100 % H_2O and $\delta X(D_2O)$ is that of in a 100 % D_2O.(Hansen *et al.*, 1985) As a result of the interaction of the exposed fluorine with the solvent or with the exchangeable protons in protein, the ^{19}F resonance can be shifted upfield. The extent of the isotope shift depends on the number of the substitutable hydrogens around the observed nuclei and also the distance and angle of hydrogen bonds.(Bernheim *et al.*, 1964; Kanazawa *et al.*, 1965) The sum value of the SIIS gives a very informative insight on the chemical environment of the fluorine being examined. It has been shown that for the F···H–N and F···O–H hydrogen bonds between the metal fluoride moiety and the protein, the magnitude of the isotope shift reflects the local proton densities around each fluorine nucleus. Moreover, even though the structures of the proteins in solution and in solid state maybe different, our observation of the isotope effect in ^{19}F NMR are always in line with the scenario in the corresponding crystal structures.(Baxter *et al.*, 2010) This is due to the fact that the metal fluoride complexes formed in the active site are more protected and conformations of the active site in both solution and crystals are very similar or the same. This gives a great opportunity to combine complementary information from NMR and crystallography and even to validate the crystal data objectively.

2.16 Crystallography

All the unpublished crystallography work in this PhD thesis were performed and permitted to use by Dr. Matthew Bowler and Miss Erika Pellegrini (a PhD student registered in University of Sheffield and supervised by Prof. Waltho) from the Macromolecular Crystallography Group, European Synchrotron Radiation Facility (ESRF), Grenoble, France, including: RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ and RhoA-

GDP-AlF_4^--$RhoGAP_{R85A}$ TSA complexes, βPGM_{D10N}-AlF_4^--(H_2O)-βG1P TSA complex, βPGM_{D10N}-βG16BP complex, βPGM_{wt}-MgF_3^--(*S*)-βCHFG1P and βPGM_{wt}-AlF_4^--(*S*)-βCHFG1P TSA complexes, PSP-MgF_3^--Ser TSA complex.

All data were collected at 293 K, at the ESRF beamline ID14-4 (λ = 0.933 Å). The dataset was integrated with XDS.(Kabsch, 2010) All subsequent calculations were carried out with the CCP4 program suite.(MolRep and REFMAC5; Vagin *et al.*, 1997; Murshudov *et al.*, 1997) The structures were solved by molecular replacement with MolRep.(Vagin *et al.*, 1997) Ligands were not included until the final rounds of refinement so that they could be built into unbiased difference Fourier maps. For the final round of refinement restraints for the metal fluoride moiety were relaxed allowing the exact atomic positions to be defined.

2.17 Structure Analysis

All the images of 3D protein structures are made using the open-source molecular visualisation software PyMOL. All the bond lengths and angles labelled in the pictures are measured directly from PDB structures in PyMOL.

Chapter 3. β-Phosphoglucomutase

3.1 Introduction

3.1.1 Introduction of the protein

β-Phosphoglucomutase from *Lactococcus lactis*, βPGM (EC 5.4.2.6), belongs to the Haloalkanoic Acid Dehalogenase (HAD) superfamily, subfamily I. It catalyses the isomerisation between β-D-glucopyranose 1-phosphate (βG1P) and β-D-glucopyranose 6-phosphate (G6P) between which the equilibrium lies strongly toward the latter (K_{eq} = 28 or 25). The reaction involves an aspartyl-phosphoenzyme (βPGM^P) intermediate formed by the phosphorylation of Asp 8 by the 6-phosphate or β-1-phosphate of the sugar or by another intermediate species, β-D-glucose 1,6-bisphosphate (βG16BP), that is formed by phosphorylation of βG1P or G6P by pAsp 8 (**Figure 3.1**).(Golicnik *et al.*, 2009) The mechanism makes use of in-line phosphoryl transfer from enzyme to either βG1P or G6P or water to accomplish the catalytic process.

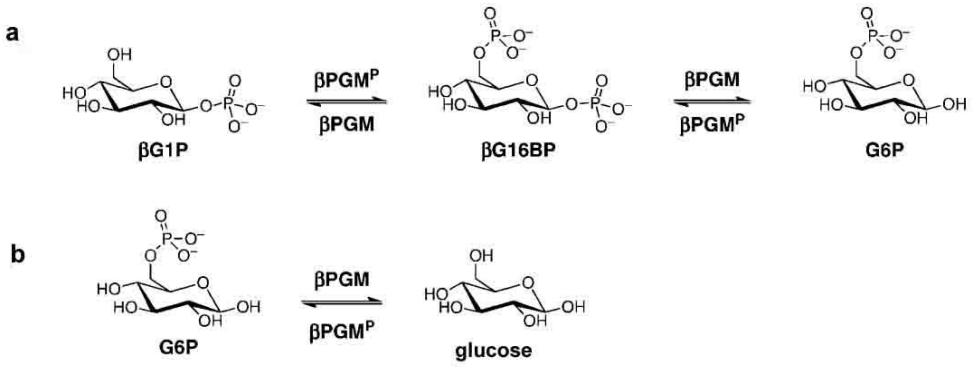


Figure 3.1 Reaction catalysed by βPGM. (**a**) The mutase activity on βG1P and G6P. (**b**) The hydrolase activity on G6P.

The catalytic mechanisms of βPGM have been under investigation for over ten years.(Collet *et al.*, 1998) The whole protein is composed of two domains: a cap domain contributing conserved Lys 45 and Thr 16, and a core domain contributing conserved Asp 8, Asp 10, Val 12, Ser 114, Lys 145, Glu 169, and Asp 170. The two domains switch between closed and open conformations. The closed conformation directly leads to desolvation of the active site and positioning of the general base residue in place which results in enzyme activation (**Figure 3.2**).

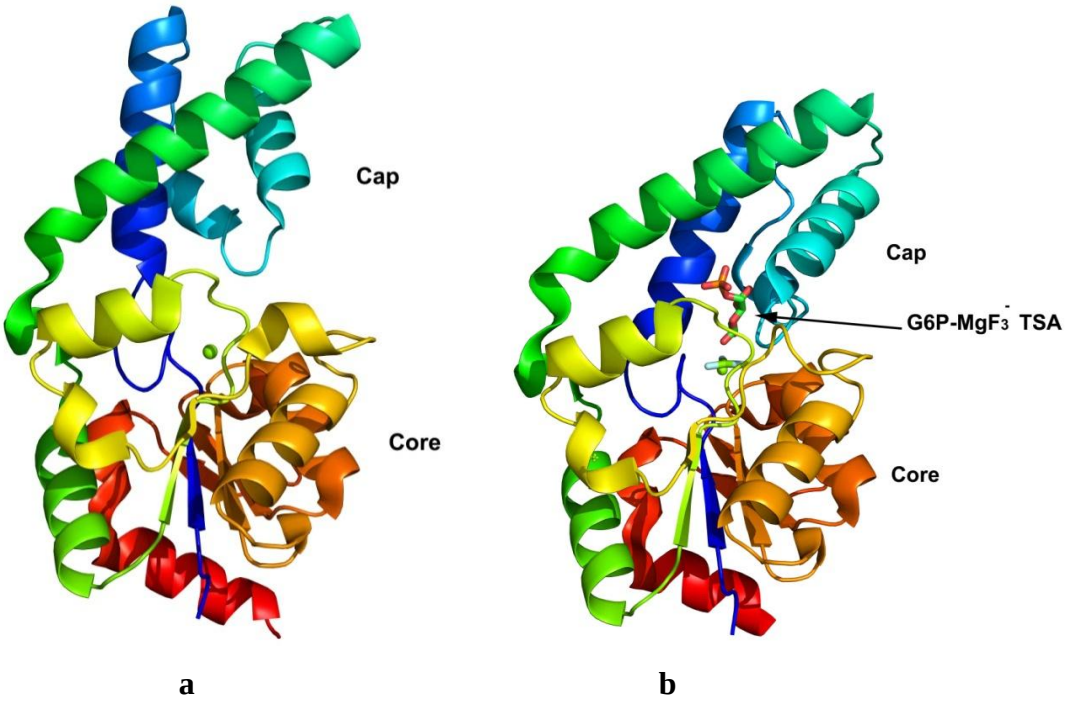
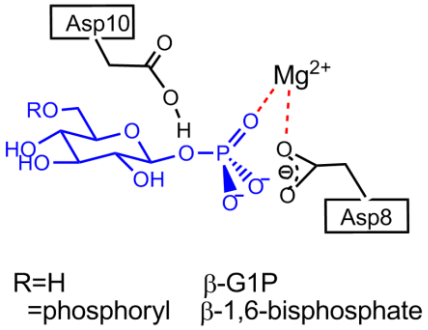


Figure 3.2 Ribbon diagram of the structure of βPGM open form (**a**, PDB:2WHE) and close form (**b**, PDB:2WF5) with the bound G6P-MgF_3^- shown as a stick model and the catalytic Mg^{2+}, represented as a green sphere.

The active site is located at the N-terminal and defined by the side chains of a motif of five highly conserved residues motif: D*X*D*X*(T/V) in which there is a nucleophilic aspartate Asp 8, a general acid/base aspartate Asp 10, and an electrophilic magnesium (**Figure 3.3**). This motif has been used repeatedly to effect a wide range of biological functions.(Aravind *et al.*, 1998; Ridder *et al.*, 1999)

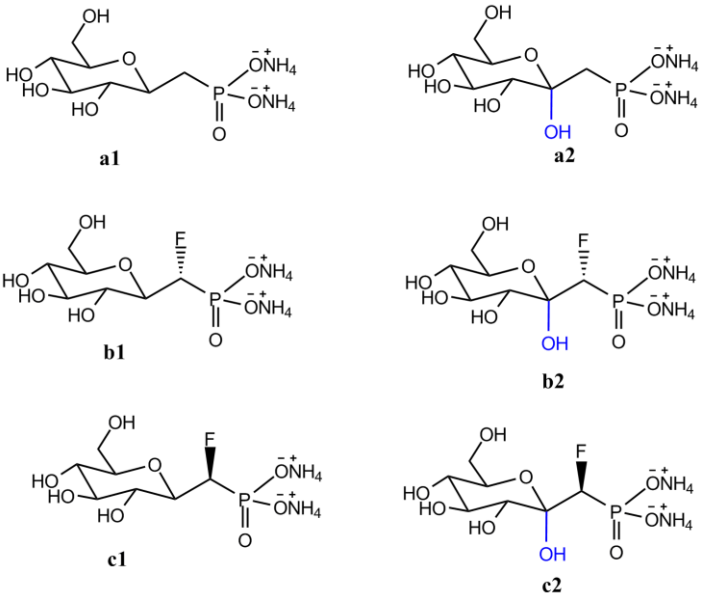
Figure 3.3 βPGM has a nucleophilic aspartate, a general acid/base aspartate and an electrophilic magnesium in the active site.



3.1.2 Aims of the work on wild type βPGM

It is a remarkable fact that 20 ~ 30% of all drugs on the market have fluorine in them.(Muller *et al.*, 2007) The high electronegativity of fluorine means it has a large electronic effect at neighbouring carbon centres, as well as having a substantial effect on the dipole moment of the molecules, the acidity or basicity of other groups nearby, not

to mention the overall reactivity and stability of the molecule. The fluorine substitution on the phosphonate methylene group has been proved to lower the pK_a of phosphonate (Blackburn *et al.*, 1981) which can potentially enhance the binding of the new dianionic phosphonate. In addition, fluorine is a hydrophobic moiety. It has been found that fluorine rarely acts as a hydrogen bond acceptor.(Dunitz *et al.*, 1997) The C–F bond length is 139 pm which means fluorine has a rather larger steric demand than hydrogen (C–H bond length 109 pm).



The wild type βPGM has been the subject of structural studies for many years using NMR in Sheffield and more recently by crystallography in ESRF, Grenoble. Based on all the knowledge, six βG1P analogues were designed by Prof. Mike Blackburn and synthesised by Dr. Debabrata Bhattasali in Professor David Jakeman's lab (Dalhousie University, Nova Scotia, Canada). They are hydrolytically stable because the C(1)-O on the sugar is replaced by $–CH_2$ or –CHF and the resulting phosphonate is resistant to hydrolysis of the C–P bond. That means they can be used as analogues for βG1P in forming TSA complexes for βPGM without undergoing chemical transformation. Compounds **a2**, **b2**, and **c2** have an α-hydroxyl group on C(1) because they were intermediates on the synthetic route before being reduced to their 1-deoxy counterparts **a1**, **b1**, and **c1**. This was designed to show that the influence of the 1-α-hydroxyl on D-glucopyranose and α-fluorine on the 1-methylene carbon of these six phosphonates determines their ability to form a TSA complex with wild type protein.

The work described in this thesis has focused on evaluating the binding of six stable phosphonate analogues of βG1P for the formation of metal fluoride TSAs and GSAs as analysed by ^{19}F NMR. The ^{19}F NMR conditions have been optimised to improve the intensity of the signals. Following successful TSA complex identification by ^{19}F NMR, the two analogues giving the best results have been used in crystal trials, performed by Erika Pellegrini and Dr. Matthew Bowler in ESRF (Grenoble, France).

3.1.3 Aims of the work on βPGM mutants

Arising from previous ^{19}F NMR studies of wild type βPGM metal fluoride complexes are several observations that have not been well explained. Using single point mutants should be helpful in terms of resolving the multiplicity and chemical shifts of signals. First, the original purpose of making the D10N mutant of βPGM two years ago was to test whether $βPGM_{D10N}$, which lacks the GAB of wild type enzyme, can trigger/mimic the minor form ('out' form) in the ^{19}F NMR of $βPGM_{wt}$-MgF_3^--G6P. In fact, this minor form has now been proved to be an octahedral magnesium coordinated by three fluorines and one backbone carbonyl from Ser 114.(Griffin, PhD thesis, 2011) Second, the two rotationally averaged broad peaks at –150.2 ppm and –153.2 ppm in the spectrum of $βPGM_{wt}$-BeF_x^--G6P complex still cannot be explained. The cause could be (a) the in- and out-conformations of Asp 10, (b) the α- and β-anomers of G6P, (c) three fluorines in an Asp 8-COO^--BeF_3^-…O1-G6P complex or in an Asp8-COO^-…$^-F_3$Be-O1-G6P complex.

The D10N mutant should be useful to test the possibility (a) because it lacks the GAB of the wild type enzyme thus can mimic the 'out' form of Asp 10. D8N could test whether the possibility (c) is valid because the change from –COOH into $–CONH_2$ can still provide a ligand to the catalytic Mg^{2+} but at the expense of coordination to the tetrahedral BeF_3^- to 'drive' it onto O1 of G6P.

In addition, we also found $βPGM_{D10N}$ has other properties which have not been seen before in the wild type enzyme that can pinpoint the significance of Asp 10 as a GAB not just on the chemical reaction kinetics but also on the dynamics of the protein as a whole.

The role of Asp 8 in the active site as a nucleophile to the phosphoryl group in the reaction is also studied. A D8E mutant has one CH_2 extension on the side chain of Asp

but maintain nearly the same pK_a for the purpose of nucleophilic attack. The carboxylic group should be able to move closer towards the phosphoryl group which, to some extent, mimics the associative mechanism of phosphotransfer. The effects of this single point mutation are investigated in activity assays and TSA complexes using NMR and crystallography.

3.2 Materials and Methods

3.2.1 Site-directed mutagenesis of βPGM to generate D8E mutant

Table 3.1 PCR setup conditions for βPGM_{D8E}.

Content	Volume / µL			
	Reaction 1 (Control)	**Reaction 2**	**Reaction 3**	**Reaction 4**
dd H_2O	39.5	16	16.5	15
10× Buffer	5		5	
Forward primer 11.0 µg/mL	1.25		12	
Reverse primer 8.9 µg/mL	1.25		14	
Template DNA	2	1	0.5	2
dNTs	1		1	
pfuTurbo polymerase	1		1	
Total Volume/ µL	50	50	50	50
Segment	Temperature / °C	Time Length / min		
1	95	1		
2	55	1		
3	68	6 min 30 s		

3.2.2 Site-directed mutagenesis of βPGM to generate D8N mutant

Table 3.2 PCR setup conditions for βPGM_{D8N}.

Content	Volume / µL			
	Reaction 1 (Control)	**Reaction 2**	**Reaction 3**	**Reaction 4**
dd H_2O	39.5	12	12.5	11
10× Buffer	5		5	
Forward 7.4 µg/mL	1.25		17	
Reverse primer 9.7 µL/mL	1.25		13	
Template DNA		1	0.5	2
dNTs	1		1	
pfuTurbo polymerase	1		1	
Total Volume/ µL	50	50	50	50
Segment	Temperature / °C	Time Length / min		
1	95	1		
2	55	1		
3	68	6 min 30 s		

3.2.3 Cell growth and expression in LB media with glucose and minimal media with glycerol

The growth and expression conditions are exactly the same apart from the difference in the supplementation by glucose or glycerol according to **Table 3.3** to test if it has anything to do with *E. coli.* CRP (C-reactive protein) in response to stress towards glucose.

Table 3.3 The protocol of media for cells grown with different carbon sources.

LB media	**LB media with glucose**	**Minimal Media with glucose**	**Minimal Media with glycerol**
trypton 10.0 g/L	trypton 10.0 g/L	Na_2HPO_4 6.0 g/L	Na_2HPO_4 6.0 g/L
yeast extract 5.0 g/L	yeast extract 5.0 g/L	KH_2PO_4 3.0 g/L pH 7.4	KH_2PO_4 3.0 g/L pH 7.4
NaCl 10.0 g/L	NaCl 10.0 g/L	NaCl 0.5 g/L	NaCl 0.5 g/L
	glucose 3.0 g/L	trace element 650 µL/L	trace element 650 µL/L
		glucose 3.0 g/L	glycerol 6mL/L
		thiamine 1.0 mL/L	thiamine 1.0 mL/L
		NH_4Cl 1.0 g/L	NH_4Cl 1.0 g/L
		$MgSO_4$ 1 mM	$MgSO_4$ 1 mM
		$CaCl_2$ 0.1 mM	$CaCl_2$ 0.1 mM
Results (analysed by ^{31}P NMR)			
βG16BP, βG1P, G6P, phosphoaspartate	None	βG16BP	βG16BP

The expression results show that β-D-glucose 1,6-bisphosphate (βG16BP) binding can be inhibited and produce a sugar-free enzyme only when βPGM_{D10N} is expressed in LB media with supplement of glucose.

3.2.4 Large scale expression

Each of the 8 × 2 L flasks containing 1000 mL LB media (or minimal media) which had been autoclaved in advance was treated with 100 µg/mL ampicillin as the antibiotic and then inoculated with 10 mL of starter culture. The large scale culture was incubated at 37°C on a 250 rpm shaker until OD_{600} 0.6~0.8 was reached. The expression was induced after adding 0.5 mM IPTG. The flasks were then moved to a 25°C room on a 250 rpm shaker and incubated for 4~6 h. The cells were harvested at 5,000 rpm, 4°C for 20 min. The pellets were stored at –80°C until use.

3.2.5 Protein purification

All βPGM_{wt}, βPGM_{D10N}, βPGM_{D8E}, and βPGM_{D8N} proteins were purified following the same procedure. The cells were resuspended in lysis buffer (HEPES 50 mM, pH 7.5, EDTA 2 mM) with a protease inhibiter cocktail tablet and lysed on ice by sonication 20

s for 3 times, with 1 min interval. The protein extract was made clear by centrifugation at 24,500 rpm, 4°C for 45 min. The supernatant was filtered through a 0.2 μm filter syringe before being loaded on to pre-equilibrated DEAE-Sepharose column.

3.2.5.1 Ion exchange chromatography

An ion exchange column of dimensions 5×2 cm, packed with DEAE-Sepharose resin (Sigma, Cat. No. DFF100) was equilibrated with ion-exchange buffer 1. The column was run at 3 mL/min and the flow through was collected. A NaCl salt gradient (0 – 0.5 M) was applied to the column over 300 mL length using an ÄKTAprime™ plus FPLC until the UV trace returned to baseline. All fractions were collected. The components of the fractions were determined by SDS-PAGE and those containing ~24 kDa size protein were collected and concentrated to 2~3 mL using a Sartorius Vivaspin™ at 5,500 g, 4°C. The DEAE-Sepharose column was cleaned with 2M NaCl, 0.1 M NaOH and preserved in 20 % ethanol.

Table 3.4 Protocol for Ion Exchange Buffer.

Ion Exchange Buffer **1**	50 mM	HEPES, pH 7.2
	5 mM	$MgCl_2$
	1 mM	EDTA
	1 mM	NaN_3
Ion Exchange Buffer **2**	50 mM	HEPES, pH 7.2
	5 mM	$MgCl_2$
	1 mM	EDTA
	1 M	NaCl
	1 mM	NaN_3

3.2.5.2 Gel filtration chromatography

A HiLoad 26/60 Superdex™ G75 Gel Filtration column (GE Healthcare) was equilibrated with gel filtration buffer. The concentrated protein was loaded onto the column through a 3 mL sample loading loop. The column was eluted at a flow rate of

Table 3.5 Protocol for Gel Filtration Buffer.

Gel Filtration Buffer	50 mM	HEPES, pH 7.2
	5 mM	$MgCl_2$
	1 mM	EDTA
	1 mM	NaN_3
	filtered and degassed before use	

2.0 mL/min and fractions (5 mL) were collected from the beginning. Samples (20 μL) from the fractions identified by UV trace as containing proteins were put on SDS-PAGE to test the purity and then the fractions containing pure βPGM were collected and

concentrated to ~2 mM using a Sartorius Vivaspin™ at 5500 g, 4°C. The final protein was stored at –80°C in 1 mL aliquots.

3.2.6 The unfolding and refolding process of βPGM_{D10N}

Concentrated βPGM_{D10N} stock (300 µL, 3.72 mM) was made up to 10 mL with unfolding buffer. If convenient, ^{1}H NMR is a quick way to show if the protein is denatured. Another 10 mL unfolding buffer was first spun through a Sartorius Vivaspin™ concentrator (20 mL, 10 kDa MWCO) to get rid of glycerol. The 10 mL protein solution was concentrated to 2 mL using Sartorius Vivaspin™ then made up to maximum 20 mL with unfolding buffer at 4°C. This procedure was repeated 3~5 times to make sure that all bound sugars were washed off when the protein was in open-form. Finally, the concentrated protein solution (2 mL) was added little by little to a 200 mL refolding buffer with strong stirring. Some precipitate formed at this step, but it was better to get rid of the precipitate by filtration or centrifugation before the next step. The diluted solution was concentrated again to 10 mL using a Sartorius Vivaspin™ concentrator and dialyzed with 2 L refolding buffer overnight for 6 times at rt using dialysis tubing (MWCO 6-8,000). The solution was concentrated to the required final concentration. ^{1}H, ^{31}P NMR were used to check if the protein was folded and no sugar bound. A ^{1}H,^{15}N-HSQC was acquired to compare with that of the open form βPGM.

Table 3.6 Buffers for unfolding and refolding βPGM_{D10N}.

Unfolding Buffer	50 mM	HEPES, pH 7.5
	1 mM	EDTA
	2 M	GdnHCl
	1 mM	NaN_3
Refolding Buffer	50 mM	HEPES, pH 7.5
	1 mM	EDTA
	5 mM	$MgCl_2$
	1 mM	NaN_3

3.2.7 Crystallography Methods

Erika Pellegrini, (ESRF, Grenoble, France) carried out all crystallisation and data collection, solved and refined the crystal structures. For the crystallisation experiments, frozen βPGM was thawed and the buffer exchanged to one containing 50 mM HEPES (pH 7.5) and 5 mM $MgCl_2$. The protein was concentrated to 15 mg/ml and supplemented with the 5 mM sugar analogue, 10 mM NaF, and 2 mM $AlCl_3$.

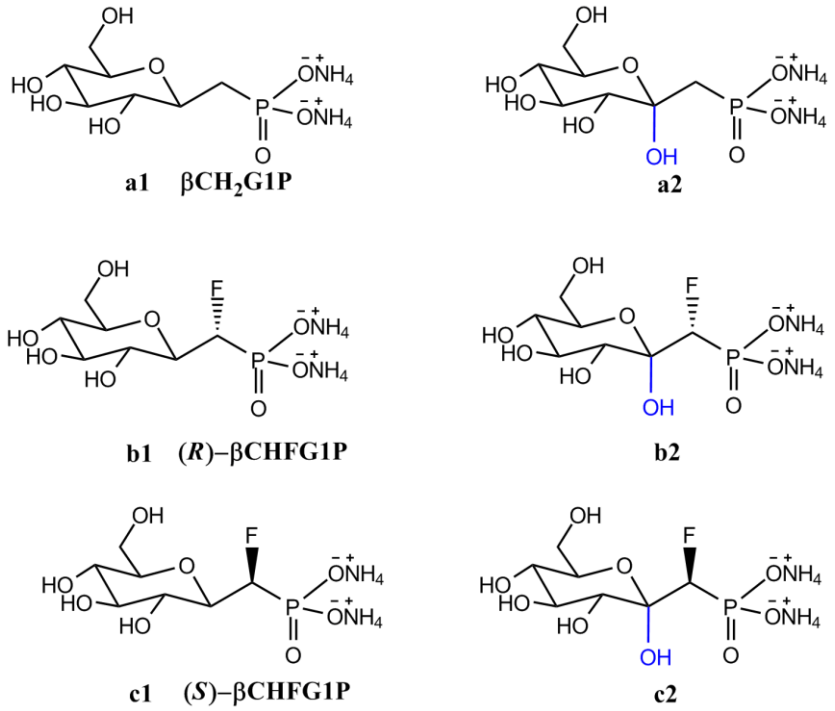
For crystallisation, 2 µl of the complex solution was mixed 1:1 with precipitant solution (27 ~ 32% PEG 3350, 50 mM magnesium acetate) and placed in sitting-drop crystallisation plate. Large plate crystals appear usually in 24 ~ 48 h. The crystals were harvested and cryprotected as described previously.(Pellegrini *et al.*, 2011)

Diffraction data were collected to 1.5 Å resolution at beamline ID14-2 at the ESRF (Grenoble, France). Data were processed with XDS. All subsequent calculations were carried out using the ccp4 program suite. The βPGM-G6P-MgF_3^- complex structure (PDB: 2WF5) was used as a search model for molecular replacement using MolRep (Vagin *et al.*, 1997) without ligands. Refinement was carried out as described previously.(Baxter *et al.*, 2010)

3.3 Results and Discussion

3.3.1 βPGM_{wt} with βG1P analogues

3.3.1.1 NMR



All six compounds were tested by ^{19}F NMR for their potential to form a trifluoromagnesate and/or tetrafluoroaluminate TSA complex. The β-1-phosphonomethylene-D-glucopyranose (**a2**) and (*R*)-1-β-phosphonofluoromethylene-D-glucopyranose (**b2**) compounds did not show any sign of TSA complex formation. (*S*)-

1-β-Phosphonofluoromethylene-D-glucopyranose (**c2**) forms a very weak complex with aluminium fluoride but not with magnesium fluoride. The deoxyglucose compound (*R*)-1-β-phosphonofluoromethylene-1-deoxyglucopyranose (**b1**) doesn't form a complex with either magnesium or aluminium fluoride. By contrast, β-1-phosphonomethylene-1-deoxy-D-glucopyranose (**a1**, shortened as **βCH_2G1P**) and (*S*)-1-β-phosphonofluoromethylene-1-deoxy-D-glucopyranose (**c1**, abbreviated to **(*S*)-βCHFG1P**) do form very stable TSA complexes with both magnesium and aluminium fluoride. Comparing the intensity of the resonances of the two spectra recorded under the same conditions, it shows that (*S*)-βCHFG1P binds stronger than CH_2G1P (**Figure 3.4**). However, NMR can only indicate the rank order of the G1P analogue binding abilities, but can't quantify their magnitude. The accurate binding constants can be measured by isothermal titration calorimetry (ITC), which is proposed to be carried out in the Chemistry Department, University of Sheffield, in the near future.

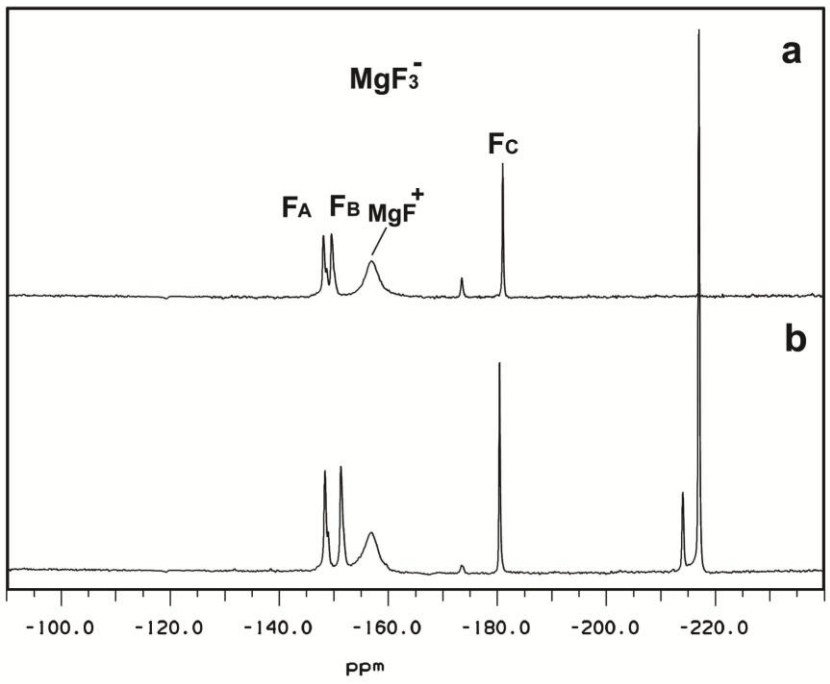


Figure 3.4 ^{19}F NMR spectra of (**a**) βPGM-MgF_3^--βCH_2G1P TSA complex (F_A = –148.2 ppm, F_B = –149.7 ppm, and F_C = –180.0 ppm.) and (**b**) βPGM-MgF_3^--(*S*)-βCHFG1P TSA complexes (F_A = –148.4 ppm, F_B = –153.4 ppm, and F_C = –180.4 ppm) recorded for the same concentration of fluoride, magnesium, and G1P analogue, showing (*S*)-βCHFG1P binds stronger than βCH_2G1P. Both samples contain 1 mM βPGM, 5 mM $MgCl_2$, 10 mM NH_4F, 5 mM (*S*)-βCHFG1P or βCH_2G1P at pH 7.3. The free and bound (*S*)-βCHFG1P in (**b**) show peaks at –214.0 and –217.0 ppm. The weak signal at –174.0 ppm is unbound sugar-MgF_x species.

Most of the ^{19}F NMR spectra were recorded with the transmitter frequency offset at –65852 Hz (–140 ppm), and a 19.5 μs pulse, which is an effective 90 degree pulse up to

about 50 ppm on either side of the transmitter frequency at no less than 90% of the power. This ^{19}F NMR setup is suitable for those spectra in which the dispersion of the peaks is between –100 and –190 ppm. In the βPGM-MgF_3^--(*S*)-βCHFG1P TSA complex, the fluorine nucleus of the methylene group on the bound and free (*S*)-βCHFG1P have chemical shifts at –214 and –217 ppm respectively. Apparently the intensity of these off-resonance fluoride signals at –214 and –217 ppm are to some extent attenuated due to the lack of excitation power. To improve accuracy of the integration of all the resonances of the metal fluoride complexes and substrates, the transmitter frequency was changed to –190 ppm enabling a spectral range of –150 to –240 ppm. That sacrifices the strength of free fluoride at –119 ppm (in this case its intensity is not important for experimental purposes) but without effect on other signals between –140 to –240 ppm. The difference between the two setups is illustrated in **Figure 3.5**. In spectrum **Figure 3.5b**, the intensity of the resonance at –214 ppm is almost as big as those of the other three MgF_3^- signals so this can also act as an internal standard to judge the occupancy of the complex.

Table 3.7 Comparison of the chemical shifts in the metal fluoride complexes using βG1P, CH_2G1P, and (*S*)-βCHFG1P.

Complex	**F_A**	**F_B**	**F_C**	**F_D**
βPGM$_{wt}$-MgF_3^--βCH_2G1P	–148.2	–149.7	–181.0	–
βPGM$_{wt}$-MgF_3^--(*S*)-βCHFG1P	–148.4	–153.4	–180.4	–
βPGM$_{wt}$-AlF_4^--βG1P	–128.6	–138.3	–141.9	–159.9
βPGM$_{wt}$-AlF_4^--βCH_2G1P	–131.2	–138.8	–140.6	–160.0
βPGM$_{wt}$-AlF_4^--(*S*)-βCHFG1P	–131.8	–138.5	–139.8	–159.7

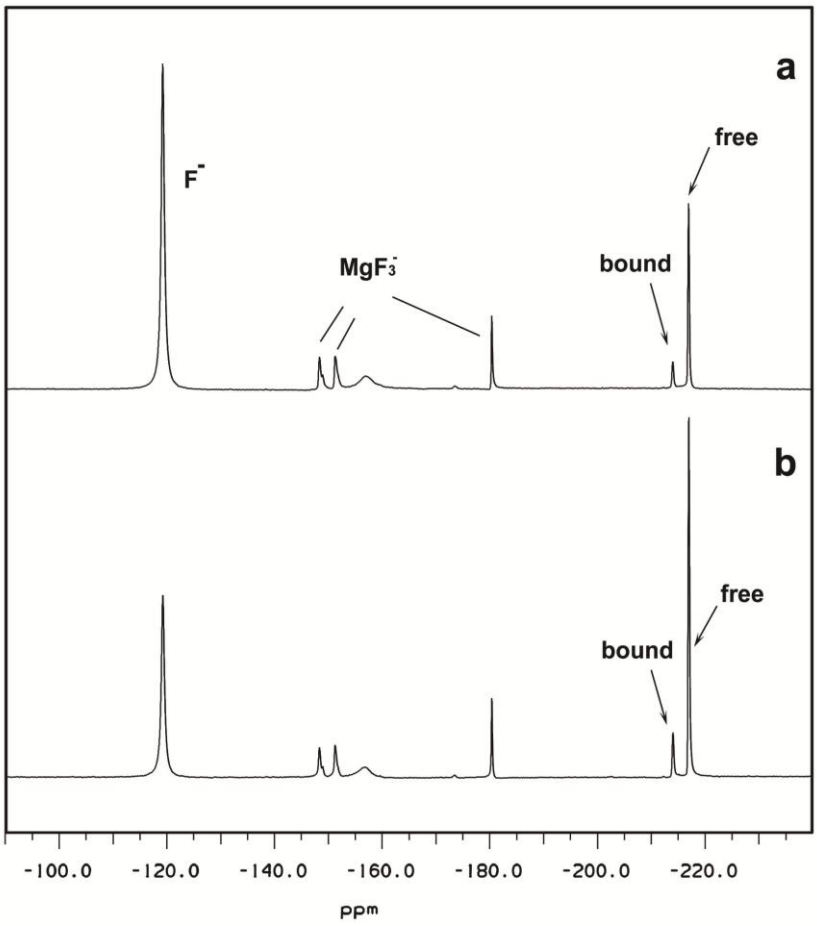


Figure 3.5 Comparison of the effect on peak intensity in ^{19}F NMR spectra for different carrier frequencies set at: (**a**) –140 ppm, the intensity of the signals for bound (*S*)-βCHFG1P (–214 ppm) and free (*S*)-βCHFG1P (–217 ppm) is attenuated; F_A = –148.4 ppm, F_B = –153.4 ppm, and F_C = –180.4 ppm. (**b**) –190 ppm, the intensity of the free F^- signal (–119 ppm) is attenuated.

In the ^{19}F NMR of βPGM-MgF_3^--(*S*)-βCHFG1P TSA complex, the most upfield shift resonance is at –180.4 ppm, suggesting a lack of hydrogen bonding partners for the 2-OH in (*S*)-βCHFG1P.(Baxter *et al.*, 2010) The same is the case for F_D in the βPGM-AlF_4^--(*S*)-βCHFG1P TSA complex (**Table 3.7**, **Figure 3.6**). As $βPGM_{wt}$ does not give a crystalline MgF_3^- complex with βG1P because of turnover degradation of βG1P, only the MgF_3^- complexes with $βCH_2G1P$ and (*S*)-βCHFG1P can be compared. The most notable changes in chemical shift between these two complexes is on F_B, at 3.7 ppm. This is likely to be induced by the substitution of fluorine in (*S*)-βCHFG1P and is later proved by the structure showing this fluorine perturbs the conformation of the loop containing Ser 116 and Ala 115 the backbone of which has direct interactions with F_B (**Figure 3.7**).

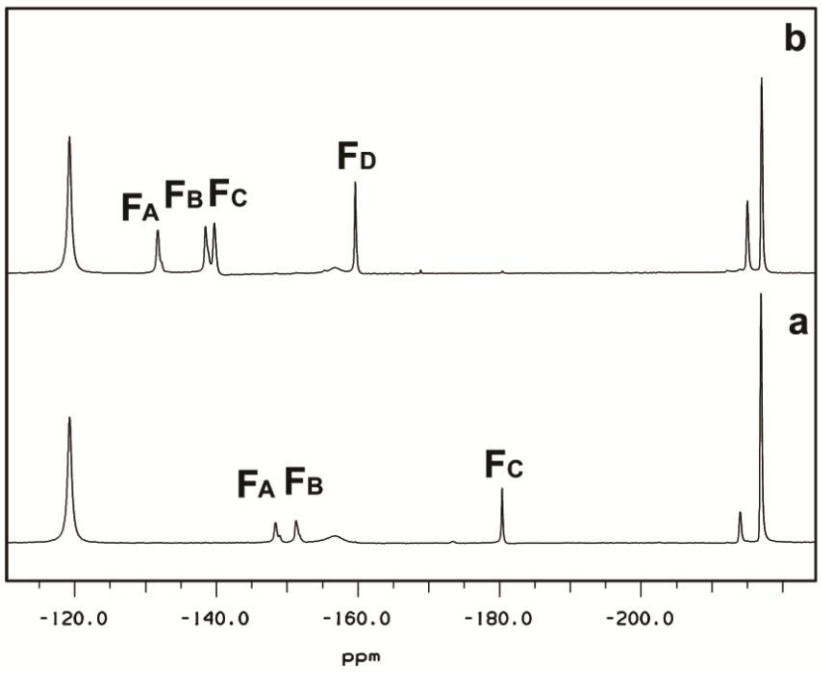


Figure 3.6 ^{19}F NMR spectra of (**a**) βPGM_{wt}-MgF_3^--(*S*)-βCHFG1P TSA complex, containing 1 mM βPGM, 5 mM $MgCl_2$, 10 mM NH_4F, 5 mM (*S*)-βCHFG1P at pH 7.3. F_A = –148.4 ppm, F_B = –153.4 ppm, and F_C = –180.4 ppm. The free and bound (*S*)-βCHFG1P show peaks at –214.0 and –217.0 ppm. (**b**) βPGM_{wt}-AlF_4^--(*S*)-βCHFG1P TSA complex, recorded in the same condition of **a** with supplement of 1 mM $AlCl_3$. F_A = –131.8 ppm, F_B = –138.5 ppm, F_C = –139.8 ppm and F_D = –159.7 ppm. The free and bound (*S*)-βCHFG1P show peaks at –215.1 and –217.1 ppm. The change of the integration of the free and bound (*S*)-βCHFG1P signals in spectra **a** and **b** clearly show that tetrafluoroaluminate has a higher occupancy than trifluoromagnesate in the active site, due to the stronger binding of aluminium fluoride TSAs.

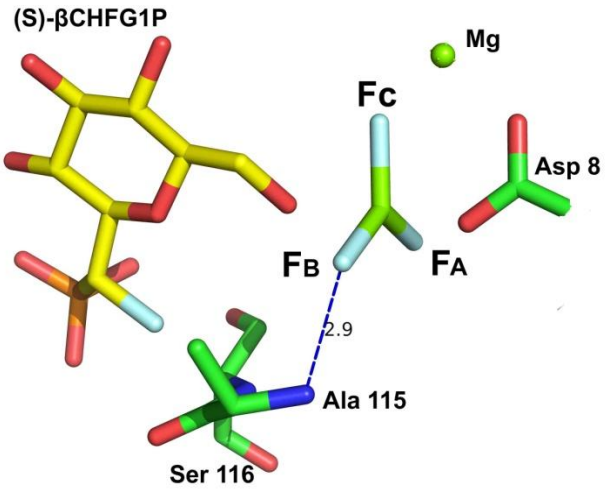


Figure 3.7 Assignment of the fluorine NMR resonances of MgF_3^- TSA in **Figure 3.6a** in the structure of βPGM-MgF_3^--(*S*)-βCHFG1P TSA complex (M.Bowler, unpublished). (*S*)-βCHFG1P is yellow sticks, protein residues are green sticks, magnesium trifluoride moiety is kelly and light blue sticks. The catalytic magnesium is a kelly sphere.

3.3.1.2 Crystallography

The whole set of ^{19}F NMR data reflects that both βCH_2G1P and (*S*)-βCHFG1P are good analogues of βG1P, therefore we can trust the structures of these analogues to a large

extent to reflect the binding of wild type βG1P. Thanks to Erika Pellegrini and Dr. Matthew Bowler in ESRF, Grenoble, we have obtained the only structure of wild type βPGM so far co-crystallised with (*S*)-βCHFG1P in their AlF_4^- and MgF_3^- TSA analogues. The crystal structures of βPGM-AlF_4^--βCH_2G1P and βPGM-AlF_4^--(*S*)-βCHFG1P will give us insights of how βG1P behaves as a substrate in βPGM in solution and it will also facilitate the analysis on other βPGM mutants complexes with βG1P.

Both PGM complexes crystallised in the conditions already described (Baxter *et al.*, 2010), giving large and well-ordered plates, diffracting to high resolution (**Table 3.8**). The difference Fourier map showed clearly the geometry of the ligands. In particular, the electron density for the sugar analogues is well defined, in comparison to that observed for G6P, implying tighter binding in the active site. The structures were solved to 1.5 Å resolution, and reassembled to give the overall geometry already described for the closed conformation of βPGM, with an rmsd Cα of 0.084.(Baxter *et al.*, 2010)

Table 3.8 Data collection and refinement statistics.

Structure	**βPGM-MgF_3^--(*S*)-βCHFG1P**	**βPGM-AlF_4^--(*S*)-βCHFG1P**
Space group	$P2_12_12_1$	$P2_12_12_1$
Wavelength	0.933 Å	0.933 Å
Unit cell dimensions (Å) *a,b,c*	37.54, 54.34, 104.35	37.19, 54.28, 104.50
Resolution range (Å)	37.63 - 1.5	48.17 - 1.5
Number of unique reflections	32136	33972
Multiplicity	3.4 (2.9)	4.2 (3.4)
Completeness (%)	97.1 (88.7)	98.1 (92.1)
R_{merge}	0.042 (0.338)	0.085 (0.595)
<I/σI>	19.8 (3.2)	12.5 (2.1)
Wilson B factor ($Å^2$)	13.6	11.9
Water molecules	293	455
R-factor (%)	16.9	14.8
Free R-factor (%)	19.5	19.8
Bonds (Å)	0.01	0.01
Angles (°)	1.53	1.26

In both structures the sugar is coordinated in the same way. The contact of sugar phosphate with protein is mediated through hydrogen bonds to the guanidinium protons of Arg 49 and side chain hydroxyl of Ser 114 (not shown). The interaction between the fluorine on 1-methyl to its surroundings is amazingly modest. It is doubtful whether this fluorine is hydrogen bonded to the backbone amide nitrogen of Arg 49 at 2.8 Å

distance because the latter forms a 2.7 Å hydrogen bond to a water molecule. While this water molecule is within 3.1 Å distance of the fluorine, it is also hydrogen bonded to another three water molecules (**Figure 3.9**). Therefore, for the (*S*)-fluorine, apart from three hydrogens that are just within Van der Waals radii (**Figure 3.10**, right), the interaction of this fluorine is not perceptibly strong with any atoms around it. It is possible that the enhanced binding is derived from displacement of a water molecule from this binding site and this achieving an entropic gain.

3.3.1.3 What's the conclusion?

The results show very clear discrimination against the recognition of the 1-α-hydroxyl on D-glucopyranose and for the presence of fluorine on the 1-methylene carbon, dependent on its stereochemistry. It is very obvious that the 1-α-hydroxyl group on the sugar is a negative factor for complex formation, because the distance between C(1) on (*S*)-βCHFG1P and the imidazole ring of His 20 is too short to accept a hydroxyl substituent on C(1) (**Figure 3.9**). Whereas the stereochemistry of the 1'-β-fluoromethylene carbon is decisive. Therefore **a1 (**βCH_2G1P) lacking the 1-hydroxyl and interference of fluorine can bind well and **c1** ((*S*)-βCHFG1P) lacking the 1-hydroxyl and with fluorine in the (*S*)-orientation binds even better (see **Figure 3.4**).

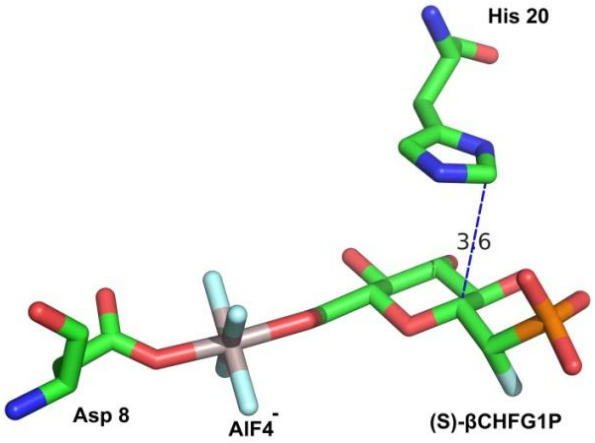


Figure 3.8 Structure to show that in βPGM-AlF_4^--(*S*)-βCHFG1P TSA complex, the distance between C(1) on (*S*)-βCHFG1P and the imidazole ring of His 20 is too short to accept a hydroxyl substituent on C(1). The protein residues are green, aluminium is grey and fluorines is light blue.

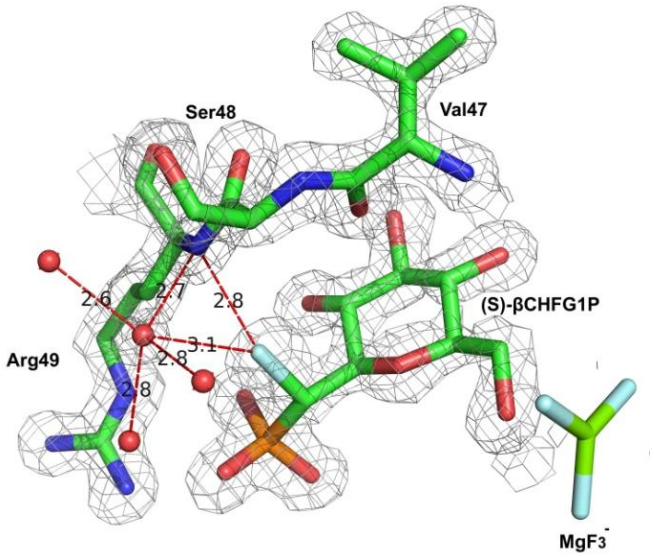


Figure 3.9 Electron density map to show the distance of 1-(*S*)-fluorine to the nearby potential hydrogen bonding partners in βPGM-MgF_3^--(*S*)-βCHFG1P TSA complex. The protein residues are green, magnesium is kelly and fluorines are light blue.

But why does the other stereoisomer (*R*)-βCHFG1P not bind? Without a structure of a βPGM complex with CH_2G1P yet available, it's difficult to tell the impact of fluorine on the molecular conformation and draw a firm conclusion. As this work is in an unknown territory, we can only seek to find the possibilities from the available structures and try to postulate the reason for the configurational binding preference of the βG1P analogues. In the (*S*)-conformer of βCHFG1P, the nearest neighbouring atoms are all hydrogens and the shortest distances from fluorine to them are 3.0, 2.6, and 2.4 Å. This is still a long enough distance without clash given the Van der Waals radius of fluorine is 1.47 Å, and that of hydrogen is 1.2 Å (**Figure 3.10**, left). By contrast, if the fluorine is inverted in PyMOL to the (*R*)-configuration in the same structure, it shifts into the neighbourhood of a carbonyl oxygen and another two C-H hydrogens with shortest distances of 2.6 Å, 2.6 Å and 1.5 Å. Considering the Van der Waals radius of oxygen is 1.57 Å, two out of three interactions are clashes because 2.6 Å to a carbonyl oxygen and 1.5 Å to a backbone CH_2 hydrogen are both too short to be acceptable (**Figure 3.10**, right).

Given this is the first time a TSA complex structure of βPGM with a stable βG1P analogue has been obtained, the structures of βPGM-MgF_3^--(*S*)-βCHFG1P TSA and βPGM-MgF_3^--G6P TSA complexes are compared here, showing there is little conformational change in the protein when it binds different sugar phosphates (**Figure 3.11**). The protein can organise the binding of G6P and (*S*)-βCHFG1P in a way that

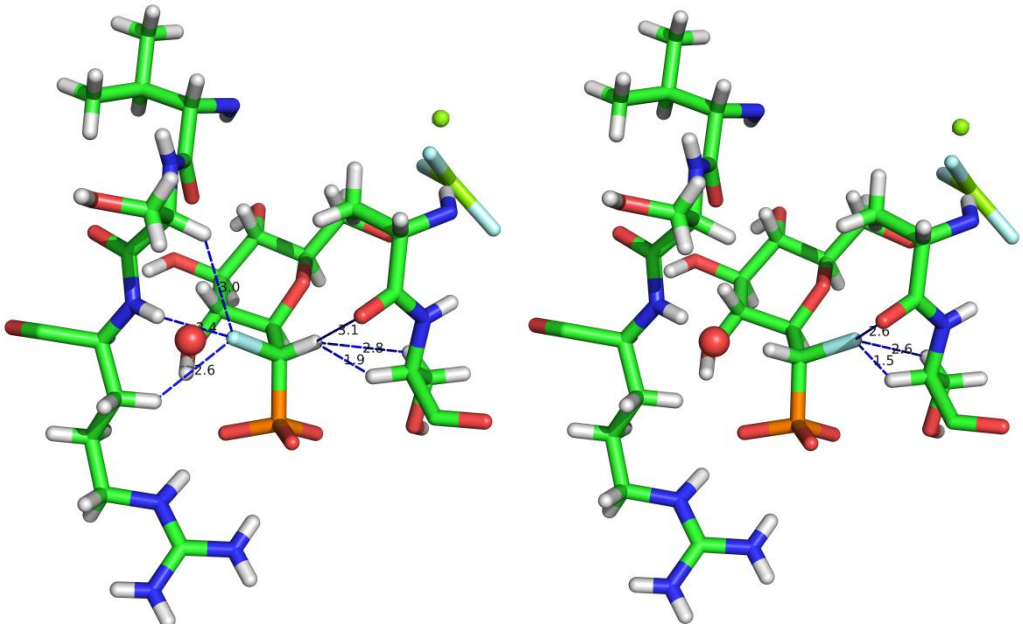


Figure 3.10 Structure of βPGM-MgF_3^--(*S*)-βCHFG1P to show the distances between the fluorine in (*S*)-βCHFG1P (left) and the simulated structure of βPGM-MgF_3^--(*R*)-βCHFG1P (right) to show that the potential interactions of the (*R*)-fluorine with its neighbours are too short to avoid clashing. The protein residues are green sticks, magnesiums are kelly and fluorines are light blue.

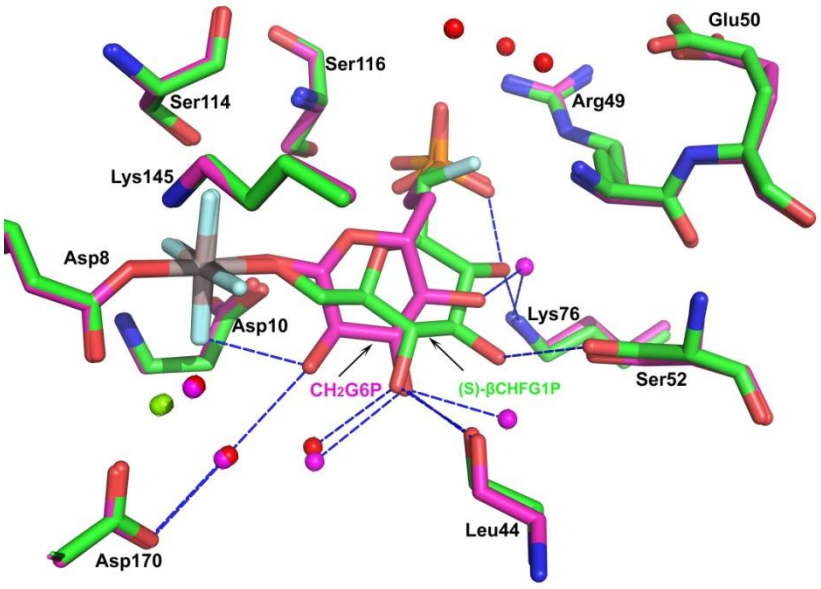


Figure 3.11 Comparison of the conformational change of the βPGM in the active site by overlay of the structures of βPGM-AlF_4^--(*S*)-βCHFG1P TSA (green, M.Bowler, unpublished) and βPGM-AlF_4^--βCH_2G6P (magenta, PDB: 2WF7) TSA complexes. The magnesium ions are kelly, aluminium ions are grey, fluorines are light blue and waters are red spheres.

C(1)-OH in G6P and the C(6)-OH in (*S*)-βCHFG1P are in the same position, and the 6-phosphate in G6P and 1-phosphonate in (*S*)-βCHFG1P are also well aligned.

3.3.2 The unfolding and refolding process of βPGM$_{D10N}$ monitored by 1D ^{1}H and ^{1}H,^{15}N-HSQC NMR

It is found that βG16BP and βG1P occupy the active site of the βPGM$_{D10N}$ protein because they co-purify through the preparative process and this complicates the formation of TSAs with the experimentally planned occupancy. To get rid of these phosphates from the purified protein, denaturation and refolding of the βPGM was performed. The folding status was determined by comparing the individual ^{1}H NMR signals located between 0.5 to –1 ppm where internal methyl groups in folded proteins are normally observed. This is because a methyl group located near an aromatic ring can give a large conformation-dependent upfield chemical shift change when the protein is folded. Therefore a folded protein usually gives well resolved resonances in the chemical shift range 0.5 to –1 ppm. By contrast, an unfolded protein gives poor chemical shift dispersion with broad peaks or even no upfield peaks above 0 ppm.

From a 1D proton NMR, it is clear that the protein has gone through an unfolding-refolding process which can be determined from the profile of the peaks at about 0 to –1 ppm and 6 to 10 ppm (**Figure 3.12**).

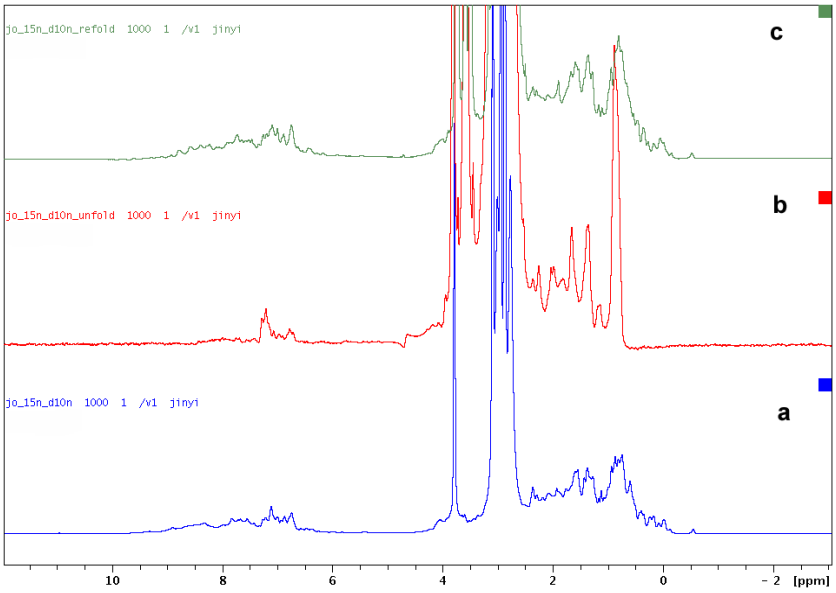


Figure 3.12 ^{1}H NMR spectra of the unfolding and refolding of βPGM$_{D10N}$. (**a**) βPGM$_{D10N}$ before denaturation. (**b**) Unfolded βPGM$_{D10N}$. (**c**) Refolded βPGM$_{D10N}$.

However, 1D ^{1}H NMR can only approximately identify if the protein is folded or not. To see whether the conformation has recovered or not after the refolding process, a 2D ^{1}H,^{15}N-HSQC of the open form βPGM$_{D10N}$ was acquired and overlaid with the open

form apo-βPGM which shows that the majority of the peaks match each other (**Figure 3.13**). The minor peaks which do not match can only be figured out by doing the whole assignment of a TROSY of βPGM. However, as the proteins are changing conformation between open and closed form in the time scale of NMR acquirement, most of the residues at critical parts of the protein, especially around active site, are relatively exposed in the solution which leads to a faster signal relaxation and line broadening, so the assignment cannot be completed.

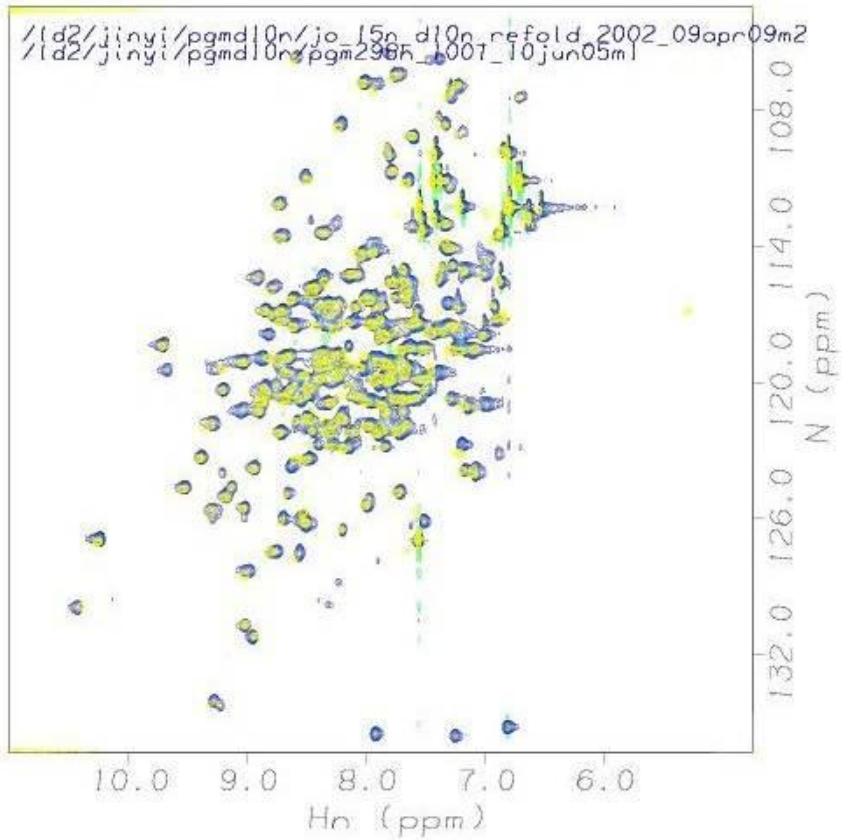


Figure 3.13 Overlay of 2D $^{1}H,^{15}N$-HSQC spectra of βPGM in its open form (blue) and the refolded form $βPGM_{D10N}$ (yellow).

3.3.3 Phosphorylation assay of $βPGM_{D10N}$ using acetyl phosphate ($CH_3COOPO_3^{2-}$)

An aspartate residue may have been selected by evolution to function in phosphoryl transfer in part because of its nucleophilicity but also in part because of the moderate kinetic stability of the acyl phosphate group. The lifetime of the aspartyl phosphate (pAsp) group in a protein might be increased through water exclusion from surroundings, while its lifetime can be decreased in an incremental manner through the positioning of a residue nearby which, depending on its structure, can serve to bind and

activate water with a predetermined efficiency. At neutral pH, phosphoaspartate residues in signaling proteins have half-lives ranging from seconds to several minutes, depending on the individual protein and especially on the presence or absence of magnesium.(Goudreau *et al.*, 1998)

As described in the experimental section, the observation of a –9 ~ –10 ppm phosphorus signal could be from pAsp 8 in βPGM_{D10N} though it is not certain why its lifetime should be longer than other aspartyl phosphates. It is also noticed in ^{31}P NMR (**Figure 3.14**) that the ratio between the bound sugar phosphate species in βPGM_{D10N} from different preparations may vary due to several reasons, such as the expression conditions and the length of purification. Nonetheless, the sum of the intensity of βG1P and G6P always seems equal to that of pAsp 8 which leads to the suggestion that the active site is either occupied by βG16BP or the shared occupancy by βG1P-pAsp 8 and G6P-pAsp 8 (**Figure 3.15**).

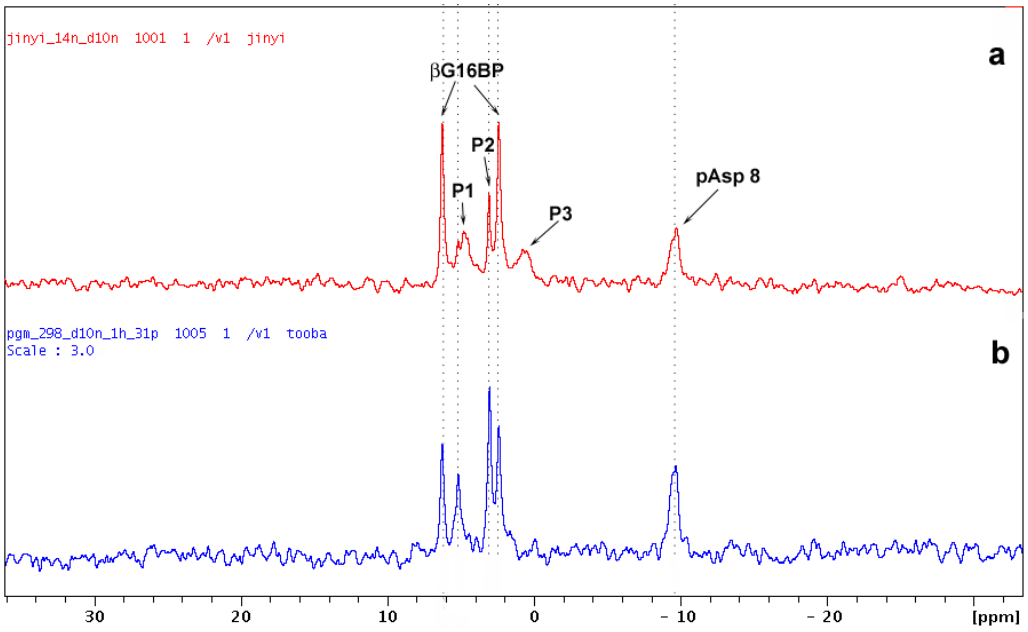


Figure 3.14 ^{31}P NMR spectra of βPGM_{D10N} from two different preparations to show the change of ratio between the bound sugar phosphate species is variable. The two samples have the same protein concentration (by ^{1}H NMR) but (**b**) (ns = 16k) was scaled up 3 fold for comparison with (**a**) (ns = 48k). The two peaks at 6.3 and 2.4 ppm belong to βG16BP, the peak at –8.6 ppm is assigned as pAsp 8. Peak **P1** was assigned as βG6P (Alizadeh, T., PhD thesis, 2009) but it is broader in the preparation (**a**) than (**b**) indicating there may be α- and β-anomer of G6P co-exist or even other phosphate species present. Peak **P2** was assigned as bound βG1P. **P3** is another unknown species only present in preparation (**a**) but not (**b**).

The aim of this experiment is to validate the assignment of phosphate species in the βPGM_{D10N} expressed in LB media and to measure roughly the phosphorylation rate on

Asp 8 in an inactive mutant βPGM_{D10N} using acetyl phosphate (acetyl-P) as the phosphorylating agent. If possible, it could also permit measuring the dephosphorylation rate on the pAsp 8. Hopefully, in the end, the function of Asp 10 as a GAB can be established through this reaction.

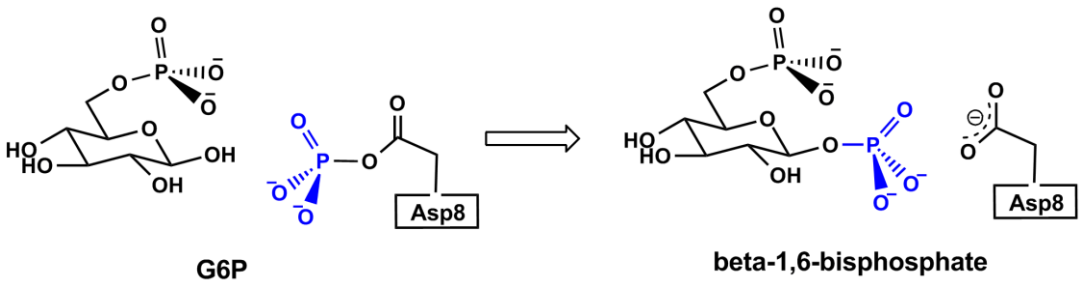


Figure 3.15 Depiction of the phosphates bound in the active site of βPGM_{D10N}.

3.3.3.1 Controls

A control was performed to test the life time of acetyl-P in the experiment buffer. It has been reported (Di Sabato *et al.*, 1961) that the rate of hydrolysis of acetyl-P is pH independent between pH 2.5 and 7.5, but is influenced by the temperature. For instance, at pH 6.9, 39°C, the $t_{1/2}$ is 2.7 h, and at 25°C the $t_{1/2}$ is 21 h. The pH of the phosphorylation assay began at 7.5 at 25°C and after about 18 h the pH dropped to 6.7. This was because of the limited buffering capacity from 50 mM HEPES and the

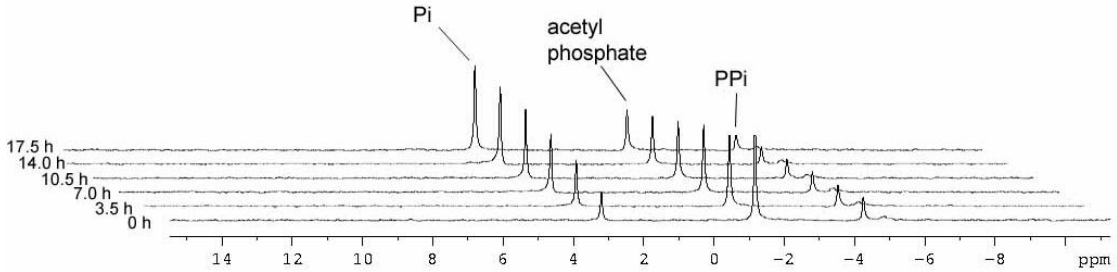


Figure 3.16 ^{31}P NMR spectra of the time course of acetyl phosphate hydrolysis without G6P and βPGM_{D10N} in 50 mM HEPES, 5 mM $MgCl_2$ at pH 7.3. Acetyl-P 30 mM was added finally and the acquisition was started within 5 min allowing time for locking, tuning, and shimming. Each spectrum contains 5120 scans and takes 1h 45min. The peak at 3.2 ppm is Pi, –1.2 ppm is acetyl-P, and – 4.2 ppm is PPi.

generation of acetate and phosphate by the hydrolysis of acetyl-P. At the end of the time course, 70% of the acetyl phosphate was intrinsically hydrolysed into inorganic phosphate and 10% was turned into pyrophosphate by self-phosphorylation (**Figure 3.16**). In a word, our control experiment matches the literature data that acetyl-P itself

in the protein buffer has long enough lifetime in water to be used to carry out the proposed phosphorylation on Asp 8.

Another control experiment was carried out to show in the presence of potential nucleophile G6P, the hydrolysis rate of acetyl-P was not increased which means excess G6P in the buffer is not consuming acetyl-P and it is not a nucleophile for the acetyl-P. This also proves that the C(1)-hydroxyl on G6P has to be deprotonated to be a good nucleophilic for pAsp 8.

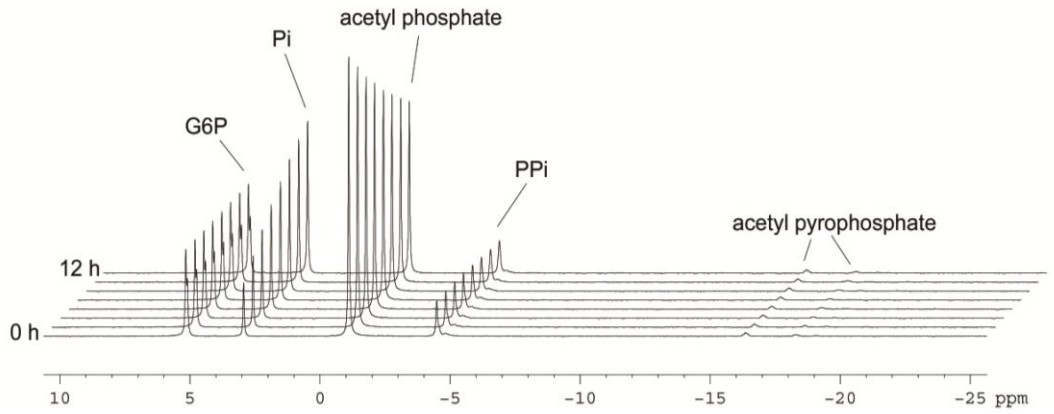


Figure 3.17 ^{31}P NMR spectra of the time course of acetyl-P hydrolysis without βPGM_{D10N} in 50 mM HEPES, 5 mM $MgCl_2$, and 5 mM G6P at pH 7.3. Acetyl-P 30 mM was added finally and the acquisition was started within 5 min allowing time for locking, tuning, and shimming. Each spectrum contains 5120 scans and takes 1h 45m.The peak at 3.2 ppm is Pi, –1.2 ppm is acetyl-P, and – 4.2 ppm is PPi.

3.3.3.2 Phosphorylation assay

The phosphorylation of βPGM_{D10N} on Asp 8 was first done without G6P, to test if this phosphorylated mutated protein has a substantial life time without any substrate bound. However, 30 mM acetyl-P was depleted within 2 min from the time the acquisition began (excluding 1 min dead time for locking and shimming, 3 min total from the moment acetyl-P was added). No valid spectrum can be recorded during that period of time. This much faster hydrolysis rate comparing to the intrinsic rate indicates there are activated nucleophiles from βPGM_{D10N} which could access to the excess acetyl-P in the solution and one of them could sensibly be the Asp 8.

Therefore, 5 mM G6P was added to facilitate the formation of the closed form, leading to water exclusion from the active site. By this means the life time of pAsp 8 should be prolonged. To the solution of 50 mM HEPES, 5 mM $MgCl_2$, 1 mM EDTA, 5 mM G6P, and 1.5 mM refolded βPGM_{D10N} at pH 7.3, 30 mM acetyl-P was added last and

acquisition began after 1 min allowing for locking and shimming. The result shows it was successful (**Figure 3.18**). The signal of acetyl-P (–1.2 ppm) decreased at a much slower rate that allowed the ^{31}P NMR to accumulate scans for about 4 h until all the acetyl-P had disappeared. The two singlets at 6.3 and 2.4 ppm which are assigned as βG16BP appeared in the first 1024 scans (21 min 18 s), as did inorganic phosphate (Pi, 3.2 ppm), pyrophosphate (PPi, –4.7 and –5.0 ppm, corresponding to Mg^{2+} bound and unbound species), and acetyl pyrophosphate (acetyl-PPi, –16.5 and –18.5 ppm). There was no sign of pAsp 8 forming within the same time. To be able to observe a low population of phosphate on pAsp 8, 5120 scans (1 h 46 min) were summed and a chemical shift appeared at –8.6 ppm as a singlet (a little different from the one observed in original βPGM_{D10N} at –9.3 and –9.6 ppm that were proposed to be two conformers

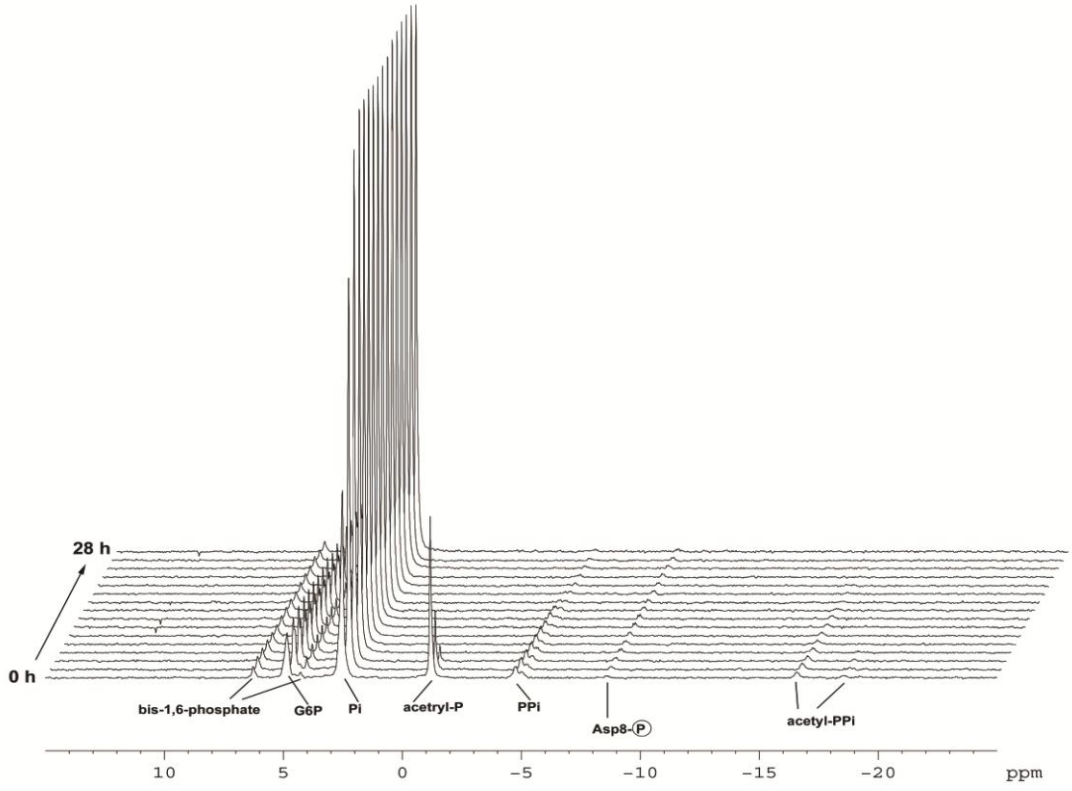


Figure 3.18 ^{31}P NMR of the hydrolysis of acetyl phosphate in the presence of βPGM_{D10N}, recorded in 50 mM HEPES, 5 mM $MgCl_2$, 1 mM EDTA, 5 mM G6P, with 1.5 mM refolded βPGM_{D10N} at pH 7.3. Acetyl phosphate 30 mM was added finally and the acquisition started within 5 min allowing time for locking, tuning, and shimming. Each spectrum contains 5120 scans and takes 1 h 46 min. Peaks at 6.3 and 2.4 ppm belong to βG16BP, 3.2 ppm is Pi, –1.2 ppm is acetyl-P, – 4.2 ppm is PPi and –8.6 ppm is pAsp 8.

paired with bound βG1P and G6P). The characteristic chemical shift at –8.6 ppm is very likely to be pAsp 8 suggested to be paired with αG6P, the lifetime of which is

longer than 28 h during the time course. At the end of the time course (28 h), all the pyrophosphate species and acetyl phosphate were hydrolysed to Pi, but protein bound βG16BP and pAsp 8 were still present. The concentration of G6P in solution was almost constant and Pi was increasing throughout the time course.

The result shows, the synthesis of βG16BP is catalysed by the βPGM_{D10N} by comparison to the control experiment without protein. Its production is faster than pAsp 8. This could be rationalised in the model if Asp 8 is first phosphorylated by acetyl-P, but protected from hydrolysis and stabilised by αG6P which is 3 fold of βG6P in the solution. The pAsp 8 then in return only phosphorylates C(1)-OH of G6P when the β-anomer binds, but not the α-anomer. This process goes on until the completion of the acetyl-P. After that, the reaction of βG16BP phosphorylating Asp 8 and pAsp 8 phosphorylating βG6P reaches the equilibrium. This explains why they can survive during the 72 h purification time. The whole process was slowed down because of the D10N mutation, rather than the turnover rate in the wild type βPGM. When the GB Asp 10 is not available after the D10N mutation, the phosphorylation of G6P in the active site should theoretically be slowed down by the factor of $10^{(pKaAsp- pKaAsn)} = 10^{[4-(-1)]} = 10^5$. Alternatively, the protein may simply provide an opportunity for G6P in proximity to acetyl-P. Both of these two possibilities can be achieved by catalysis by another general base (either in protein or in the solution) rather than by Asp 10. The candidates can be the phosphate in acetyl phosphate itself, acetate from the hydrolysis of acetyl phosphate, or other potential residues in the active site.

3.3.4 βPGM_{D10N}-AlF_4^--G6P TSA complex

A series of TSAs experiments on βPGM_{D10N} was first done by Tooba Alizadeh when the multi-sugar occupancy in the active site of βPGM_{D10N} as expressed in LB media was not known (see T. Alizadeh PhD Thesis, 2008). In my experiments, several ways of removing the bound sugar species were tried and the two successful methods are described here. One is to denature the protein, buffer exchange to remove the sugar, and refold protein back to the original conformation. The other method is to express the βPGM_{D10N} in LB media with supplement of glucose 3.0 g/L and this expression condition can successfully suppress the production of sugar phosphates to make sugar-free βPGM_{D10N} protein. The βPGM_{D10N} used for all the experiments in the present work was made using these two methods.

To make the βPGM_{D10N}-MgF_3^--G6P complex, 10 mM NH_4F was added to the sample containing 50 mM HEPES, 1 mM EDTA, 1 mM βPGM_{D10N}, 5 mM $MgCl_2$, and 5 mM G6P at pH 7.3. However, the ^{19}F NMR result shows βPGM_{D10N} cannot form a βPGM_{D10N}-MgF_3^--G6P complex, because MgF_3^- is not stable enough to overcome the energy barrier for the deprotonation of the sugar. But the βPGM_{D10N}-AlF_4^--G6P TSA complex was successfully formed with protein generated with both methods (**Figure 3.19**). The formation of an AlF_4^- TSA complex also indirectly proves that βPGM_{D10N} has been refolded correctly back into the native conformation and so is able to synthesis the metal fluoride complex at the active site.

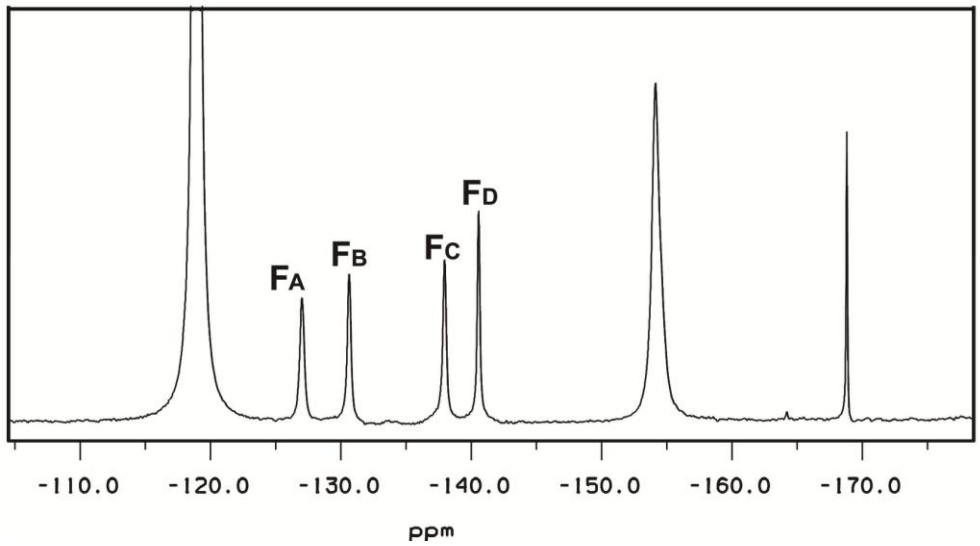


Figure 3.19 ^{19}F NMR of the βPGM_{D10N}-AlF_4^--G6P complex in 50 mM HEPES, 1 mM EDTA, 1 mM βPGM_{D10N}, 5 mM $MgCl_2$, 10 mM NH_4F, 5 mM G6P, 1 mM $AlCl_3$ at pH 7.3. Protein-bound TSA F_A = –127.1 ppm, F_B = –130.7 ppm, F_C = –138.0 ppm, and F_D = –140.5 ppm. F^- = –119.2 ppm, free AlF_x = –155.0 ppm, unbound sugar-AlF_x species= –168.9 ppm

Table 3.9 ^{19}F chemical shifts (ppm) and SIIS (ppm) of the βPGM_{wt}-AlF_4^--G6P and βPGM_{D10N}-AlF_4^--G6P complexes for comparison.

Complex	**F_B**	**F_A**	**F_D**	**F_C**
βPGM_{wt}-AlF_4^--G6P	–130.6	–137.0	–140.6	–144.0
SIIS	0.8	1.3	0.9	0.8
Complex	**F_A**	**F_B**	**F_C**	**F_D**
βPGM_{D10N}-AlF_4^--G6P	–127.1	–130.7	–138.0	–140.5
SIIS	–	–	–	–

Comparing the chemical shifts of βPGM_{D10N}-AlF_4^--G6P with those of the βPGM complex, the impact of the removal of the Asp 10 is significant. As we have no structure of this complex, the full assignment of the peaks is based on the chemical shift order. Because

the backbone amide nitrogen of Asp 10 has a direct hydrogen bonding interaction F_A, its mutation can cause some apparent conformational change of the backbone which induces the biggest chemical shifts downfield by 7 (137.0 – 130.0) or even 10 (137.0 – 127.1) ppm depending on the assignment (**Figure 3.20**). This scale of change indicates the formation of additional hydrogen bonds. All the downfield shifts of the resonances show the overall electron deshielding of the AlF_4^- moiety in the D10N mutant complex.

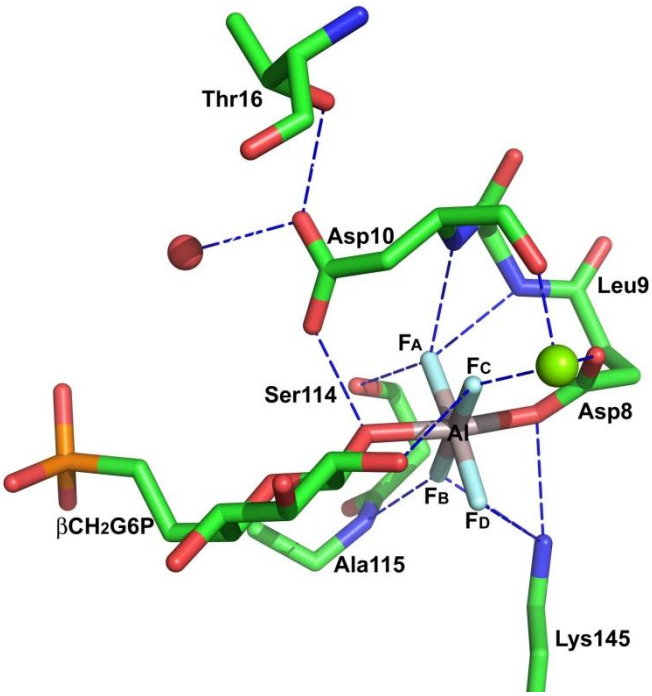


Figure 3.20 Structure of βPGM_{wt}-AlF_4^--βCH_2G6P TSA complex (PDB: 2WF7). βPGM_{wt} and βCH_2G6P are green sticks, aluminium is grey, catalytic magnesium is kelly sphere, water molecule is red sphere and fluorines are light blue.

The orientation of the Asn 10 amide side chain determines if it acts as a hydrogen bond donor or acceptor, which can also influence the chemical shifts of the fluorines. As there is only one set of resonances in the spectrum, there should only be one orientation. No structure is yet available for this βPGM_{D10N}-AlF_4^--G6P, but from the structure of the wild type complex (PDB: 2WF7), it is uncertain whether Asn 10 should donate a hydrogen bond to the C(1)-oxygen because the possibility of the carbonyl group in Asp 10 can either accept or donate a hydrogen bond to Thr 16 and a water molecule (**Figure 3.20**). The average chemical shifts of all the resonances are more downfield for the D10N complex than for the wild type complex (**Table 3.9**). This indicates that the aluminium is more electropositive in the D10N mutant complex and is best explained by postulating that in this mutant TSA complex, the glucose 1-oxygen is protonated and

thus cationic. (A similar case of this is in tetrafluoroaluminate TSA complexes for RhoA-RhoGAP, the apical water is cationic and protonated as shown by its forming two hydrogen bonds to neighbouring backbone carbonyls. PDB: 1OW3 and 1TX4). That is consistent with the absence of Asp 10 as a GB to cause its deprotonation. By contrast, in the wild type TSA complex, the sugar 1-oxygen is hydrogen bonded to the protonated carboxyl group of Asp 10 (**Figure 3.20**, **Figure 3.21**) and so carries far less positive charge, making the aluminium less electropositive and its four fluorines less deshielded. Thus the whole electron density of the metal fluoride moiety rises in the wild type complex.

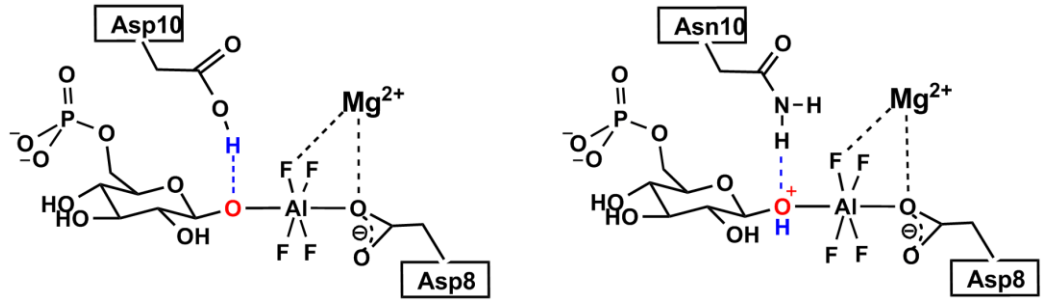


Figure 3.21 The scheme to show the protonation status of C(1)-hydroxyl group of G6P in complex with βPGM_{wt} (left) and βPGM_{D10N} (right). In βPGM_{D10N}, the C(1)-oxygen of G6P is likely to be protonated.

3.3.5 Analysis of the ^{19}F NMR of βPGM_{D10N}-AlF_4^--βG1P and βPGM_{D10N}-AlF_4^--βCH_2G1P TSA

In normal physiological conditions, the concentration of G6P is 28 times higher than that of βG1P. For wild type βPGM, thanks to the fact that C(6)-hydroxyl on βG1P is a better nucleophile than C(1)-hydroxyl on G6P, the activation energy for phosphorylating βG1P is lower. Therefore, βG1P binds more tightly than G6P in a TSA complex. But when wild type βPGM was used to form βG1P TSA complexes, the aluminium fluoride inhibition of the turnover from βG1P to G6P is not complete so that signals for multiple fluoroaluminate TSA complexes in the spectrum of one sample was observed. That raises difficulties in the assignment of multidimensional spectra of all the complexes with βG1P. One way to avoid this is to use an inactive enzyme like βPGM_{D10N}. Another is to use a stable βG1P analogue such as βCH_2G1P or (*S*)-βCHFG1P.

3.3.5.1 ^{19}F NMR analysis of βPGM_{D10N}-AlF_4^--$\beta G1P$ and βPGM_{D10N}-AlF_4^--βCH_2G1P TSA complexes

The samples for ^{19}F NMR analysis were set up as described for other TSA experiments using βG1P and βCH_2G1P. But to avoid inducing domain closure (half-closure) by AlF_4^-, which may prevent sugar binding, and in order to reproduce the results, the order of adding each ingredient was followed strictly: sugar first, then fluoride and aluminium. The ^{19}F NMR data show both βG1P and βCH_2G1P are able to form AlF_4^- TSA complexes with βPGM_{D10N} with very similar fluorine chemical shifts (all chemical shift changes are within 0.6 ppm) and coupling patterns (**Figure 3.22**). The βPGM_{D10N}-AlF_4^--βG1P TSA complex has chemical shifts at F_A = –134.6 ppm, F_B = –136.6 ppm, F_C = –141.4 ppm and F_D = –157.0 ppm. Similarly, values appear in the βPGM_{D10N}-AlF_4^--βCH_2G1P TSA complex at F_A = –134.4 ppm, F_B = –136.9 ppm, F_C = –141.2 ppm, and F_D = –157.6 ppm. Among the four fluorine signals, F_A and F_B have broader line widths than F_C and F_D, which indicates in the proposed structure the angle between F_A and F_B is slightly bigger than 90 degree but the angle between F_C and F_D is smaller.

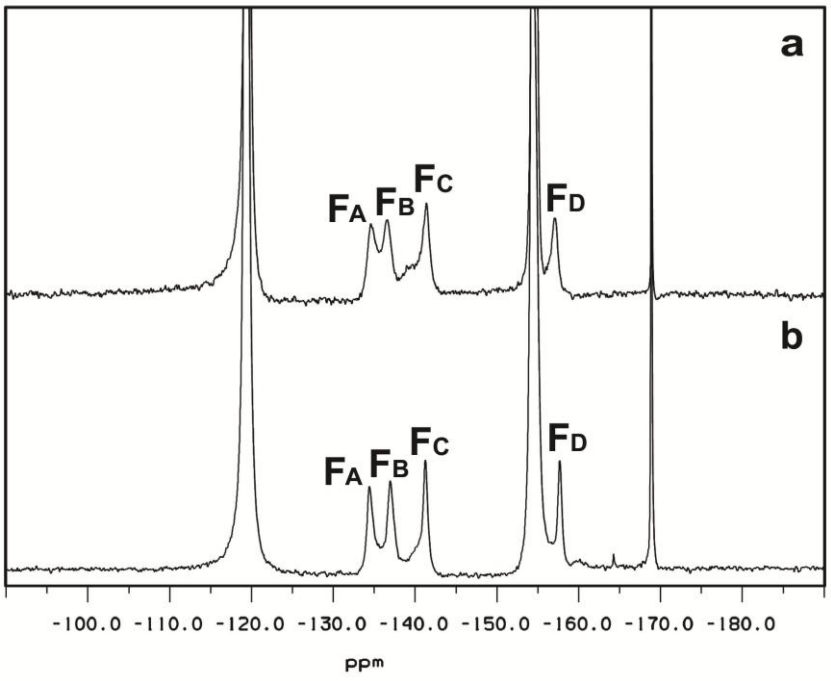


Figure 3.22 ^{19}F NMR spectra showing: (**a**) βPGM_{D10N}-AlF_4^--βG1P TSA complex is formed in 50 mM HEPES, 1 mM EDTA, 1 mM βPGM_{D10N}, 5 mM $MgCl_2$, 2 mM $AlCl_3$, 10 mM NH_4F, 10 mM βG1P, 10% D_2O at pH 7.2; F_A = –134.6 ppm, F_B = –136.6 ppm, F_C = –141.4 ppm and F_D = –157.0 ppm. (**b**) βPGM_{D10N}-AlF_4^--βCH_2G1P TSA complex is recorded in 50 mM HEPES, 1 mM EDTA, 1 mM βPGM_{D10N}, 5 mM $MgCl_2$, 2 mM $AlCl_3$, 10 mM NH_4F, 10 mM βCH_2G1P, 10% D_2O at pH 7.3. F_A = –134.4 ppm, F_B = –136.9 ppm, F_C = –141.2 ppm and F_D = –157.6 ppm. In

both spectra, free F^- = –119.2 ppm, free AlF_x species = –155.0 ppm, unbound sugar-AlF_x species = –168.9 ppm

3.3.5.2 Protein structures of βPGM$_{D10N}$-AlF$_4^-$-H$_2$O-βG1P and βPGM$_{wt}$-AlF$_4^-$-βCH$_2$G6P complexes

One unpublished structure of βPGM$_{D10N}$-AlF$_4^-$-H$_2$O-βG1P complex at 1.4 Å resolution was determined in advance of the later knowledge of the existence of co-purified bound

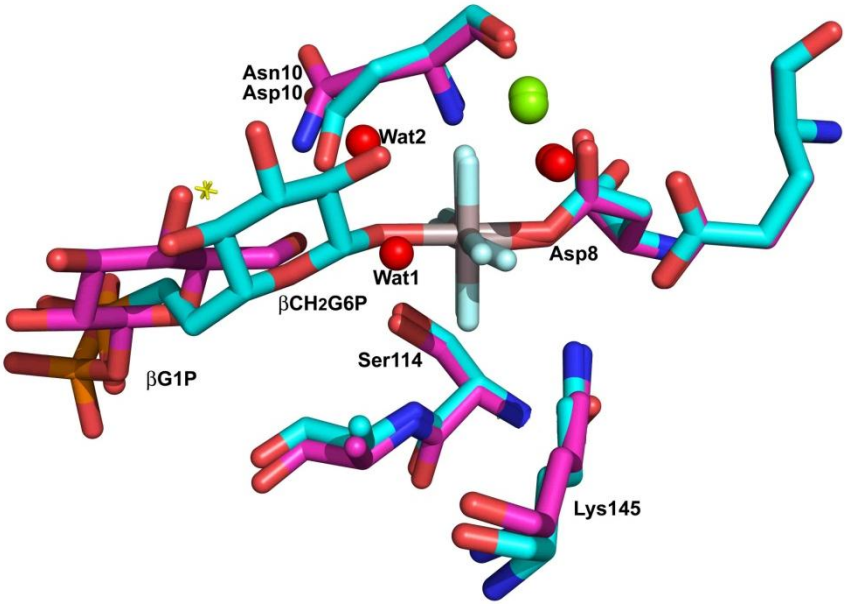


Figure 3.23 Overlaid structures of βPGM$_{D10N}$-AlF$_4^-$-H$_2$O-βG1P (magenta, M.Bowler, unpublished) with βPGM$_{wt}$-AlF$_4^-$-βCH$_2$G6P (cyan, PDB: 2WF7). Aluminium is grey, fluorines are light-blue, catalytic magnesiums are kelly spheres and water molecules are red spheres. A water molecule Wat1 in axial co-ordination to AlF$_4^-$displaces the 6-OH of βG1P. Another water molecule, Wat2, occupies the position occupied by the C(2)-OH of the βCH$_2$G6P analogue in βPGM$_{wt}$-AlF$_4^-$-βCH$_2$G6P.

sugar phosphates. This complex was obtained by addition of excess βG1P to the mutant protein. A comparison of this unpublished structure of the βPGM$_{D10N}$-AlF$_4^-$-H$_2$O-βG1P complex with that of the βPGM$_{wt}$-AlF$_4^-$-βCH$_2$G6P complex (PDB: 2WF7) clearly shows a water molecule (designated Wat1 in **Figure 3.23**) is at 1.8 A from the aluminium and in line (174˚) and therefore is coordinated to the planar AlF$_4^-$ replacing the expected C(6)-OH of βG1P and hydrogen bonded to the C(6)-OH. (It should be noted that this Wat1 may actually be a hydroxide ligand because it clearly receives a hydrogen bond from the backbone amide of Gly 46 and likely receives a hydrogen bond from the C(6)-OH) A second water molecule, Wat2, occupies the position of the C(2)-OH group of the wild type complex with βCH$_2$G6P. This result suggests that βG1P

outcompetes the βG16BP, present from copurification with the protein, in binding to the active site of the D10N mutant with its 1-phosphate occupying the passive binding site. However, because the 6-OH group of βG1P is not activated by GB catalysis from Asp 10, it fails to coordinate to the aluminium and, instead, a water molecule occupies that axial site. The 2-OH group of the displaced βG16BP leaves a site into which Wat 2 binds, forming a hydrogen bond to one of the fluorides. Asn 10 takes up an intermediate position (neither 'in' or 'out') and primarily interacts with the C(6)-OH of βG1P rather than the attacking water (**Figure 3.23**).

3.3.5.3 Interpretation of ^{19}F NMR of βPGM_{D10N}-AlF_4^--βG1P in the light of the protein structure of βPGM_{D10N}-AlF_4^--H_2O-βG1P

Because we haven't got a protein structure of a TSA complex for D10N, we cannot directly make ^{19}F NMR comparisons with a wild type TSA complex. However, it is possible to relate the chemical shifts of βPGM_{D10N}-AlF_4^--βG1P TSA and βPGM_{wt}-AlF_4^--βG1P complexes to establish that the D10N complex studied by NMR is indeed a TSA complex. This calls for indirect comparisons using the available crystal data from a wild type TSA complex, βPGM_{wt}-AlF_4^--(*S*)-βCHFG1P.(Bowler, unpublished) What can be confirmed is that F_D, which has the most upfield chemical shift and the smallest line width, is the one binding to the catalytic magnesium and is missing a hydrogen bond contact to the 2-OH compared to the G6P structure. The most downfield F_A has the biggest chemical shift change because of its hydrogen bonding to the backbone amide of Asn 10 which is likely to persist through the D10N mutation.

Table 3.10 ^{19}F chemical shifts (ppm) of AlF_4^- in βPGM_{wt} complexes with βG1P, βCH_2G1P, and (*S*)-βCHFG1P, and in the βPGM_{D10N} complexes with βG1P and βCH_2G1P for comparison.

Complex	F_A	F_B	F_C	F_D
βPGM_{wt}-AlF_4^--βG1P	–128.5	–138.2	–141.8	–159.9
βPGM_{wt}-AlF_4^--βCH_2G1P	–131.0	–138.6	–140.4	–159.8
βPGM_{wt}-AlF_4^--(S)-βCHFG1P	–131.8	–138.5	–139.8	–159.7
βPGM_{D10N}-AlF_4^--βG1P	–134.6	–136.6	–141.4	–157.0
βPGM_{D10N}-AlF_4^--βCH_2G1P	–134.4	–136.9	–141.2	–157.6

The structure of wt βPGM βPGM_{wt}-AlF_4^--(*S*)-βCHFG1P TSA in solution is almost the same as that of βG1P according to chemical shift analysis, with the exception of for F_A (**Table 3.10**). So when it is aligned with the structure of βPGM_{D10N}-AlF_4^--H_2O-βG1P (aligned structures not shown here), some fundamental differences can be seen. These

changes should result in quite a big chemical shift change, larger than the 3 ppm observed (**Table 3.10**) which is largely the result of the replacement of Asp 10 by Asn 10. Therefore, we must conclude this βPGM_{D10N}-AlF_4^--H_2O-βG1P crystal structure is not a good reflection of the same complex in solution.

3.3.6 βPGM_{D10N}-BeF_3^--H_2O and βPGM_{D10N}-BeF_3^--G6P GSA complexes

It has been noticed that in most structures Asp 10 adapts an 'in' conformation except for the βPGM_{wt}-BeF_3^--G6P GSA (PDB: 2WF8), βPGM_{wt}-BeF_3^--H_2O (PDB: 2WFA), and apo-βPGM_{wt} structures (PDB: 2WHE). When the two βPGM_{wt}-BeF_3^--G6P GSA structures 2WF8 and 2WF9 are aligned, Asp 10 is seen in 'in' and 'out' conformations, respectively. When Asp 10 is in the in-conformation, the change in torsion angle of the backbone allows the side chain of Asp 10 to form two well-defined 2.6 Å hydrogen

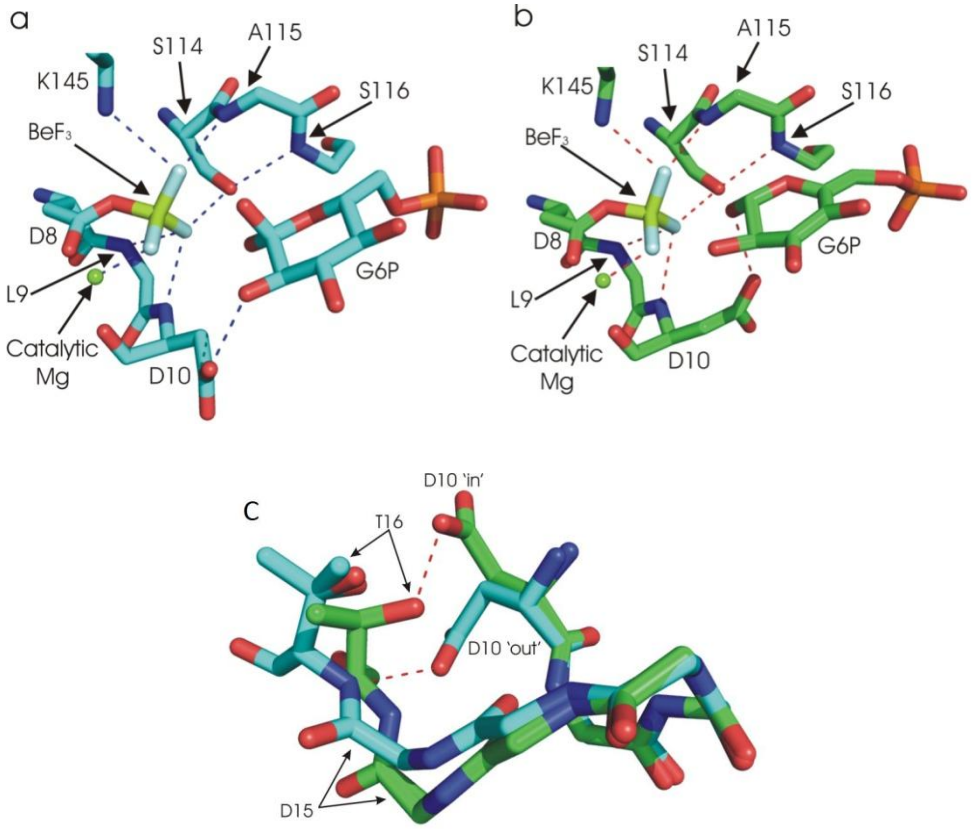


Figure 3.24 Crystal structures of the active site of (**a**) βPGM_{wt}-BeF_3^--G6P TSA complex (PDB: 2WF9) with Asp 10 'out'. The sticks show that both the α- and β-anomers of G6P are present within the 2WF9 crystal structure, suggesting that binding of Asp 10 to the C(1)-OH controls anomer selection. (**b**) βPGM_{wt}-BeF_3^--G6P GSA complex form 1 (PDB: 2WF8) with Asp 10 'in'. (**c**) Crystal structure overlays of the two βPGM_{wt}-BeF_3^--G6P TSA complex structures (2WF8 and 2WF9 as green and cyan sticks, respectively) showing the differences in the backbone torsion angles.

bonds with the Thr 16 side chain hydroxyl and the sugar C(1)-hydroxyl (**Figure 3.24 b** and **c**). In the Asp 10 'out'-conformation, the side chain of Asp 10 forms a 2.9 Å hydrogen bond to the backbone amide of Thr 16 and a 2.6 Å hydrogen bond to C(2)-hydroxyl of the sugar (**Figure 3.24 a** and **c**). The movements of this Asp 10 observed in the crystals are interpreted as a conformational change mechanism to prevent the hydrolysis of the phosphoenzyme by water when the sugar is not present (or is being reorientated), because the 'out' Asp 10 is no longer able to act as a GB to extract the proton from water molecule. When the sugar is reorientated to the 'in' conformation in the active site, Asp 10 is able to extract the proton from C(1)-OH of the sugar for catalysis to proceed. For βPGM to be a mutase and not a phosphatase, it needs to transfer the phosphate group to the alternative end of the hexose phosphate and not to water.

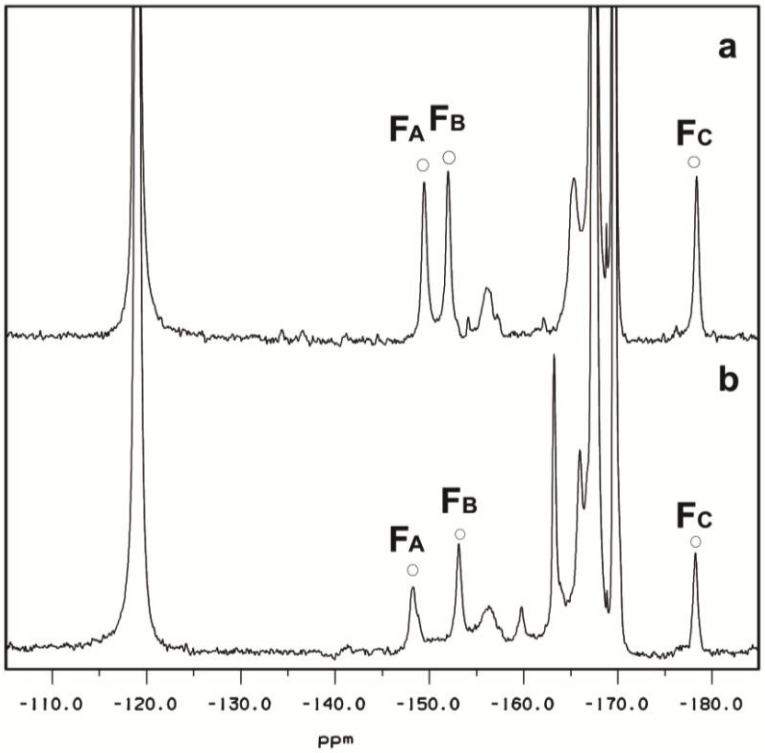


Figure 3.25 ^{19}F NMR spectra showing: (**a**) βPGM_{D10N}-BeF_3^--H_2O GSA complex, recorded in 50 mM HEPES, 0.5 mM βPGM_{D10N}, 5 mM $MgCl_2$, 1 mM EDTA, 5 mM $BeCl_2$, 20 mM NH_4F at pH 7.3; F_A = –149.4 ppm, F_B = –152.0 ppm, F_C = –178.3 ppm; (**b**) βPGM_{D10N}-BeF_3^--G6P GSA complex, recorded in the same conditions with supplement of 5 mM G6P; F_A = –148.3 ppm, F_B = –153.1 ppm, F_C = –178.2 ppm. In both spectra, free F^- = –119.2 ppm, and the resonances between –160.0 to –170.0 ppm are free and sugar-bound BeF_x species.

The mutation of D10N mimics to some extent the Asp 10 in an 'out' position because it is not able to remove a proton from either water or hexose. To test this idea in solution,

two experiments were set up with and without G6P present in 50 mM HEPES, 0.5 mM βPGM_{D10N}, 5 mM $MgCl_2$, 1 mM EDTA, 5 mM $BeCl_2$, and 20 mM NH_4F at pH 7.3. The ^{19}F NMR shows not much change in chemical shift and SIIS between the βPGM_{D10N}-BeF_3^--H_2O GSA complex and the βPGM_{wt}-BeF_3^--H_2O GSA complex (**Figure 3.25, Table 3.11**). F_A and F_C move downfield 0.8 and 0.5 ppm respectively and F_B moves slightly upfield (0.4 ppm). These small changes are due to the mutation of Asp 10 into Asn 10. But because the Asp 10 in the wild type structure is 'out', the mutation doesn't cause any big overall change on the conformation of the GSA complex.

Table 3.11 ^{19}F chemical shifts (ppm) and SIIS (ppm) of the βPGM_{D10N}-BeF_3^--H_2O, and βPGM_{D10N}-BeF_3^--G6P complexes comparing to wild type complexes.

Complex	**F_A**	**F_B**	**F_C**
βPGM_{wt}-BeF_3^--H_2O	–150.2	–151.5	–178.7
SIIS	1.1	1.1	0.2
βPGM_{wt}-BeF_3^--G6P	–150.2(b)	–153.2(b)	
βPGM_{D10N}-BeF_3^--H_2O	–149.4	–152.0	–178.3
SIIS	1.2	1.0	0.3
βPGM_{D10N}-BeF_3^--G6P	–148.3	–153.1	–178.2
SIIS	0.8	1.0	0.1

In wild type protein, the conformational change in the active site, particularly around BeF_3^- moiety between the βPGM_{wt}-BeF_3^--H_2O (PDB: 2WFA) GSA and βPGM_{wt}-BeF_3^--G6P (PDB: 2WF9) GST complexes, is minor (**Figure 3.26**). Asp 10 is in 'out' conformation in both structures. In βPGM_{wt}-BeF_3^--G6P structure, the sugar in G6P is present as both α- and β-conformers and the water molecule in βPGM_{wt}-BeF_3^--H_2O complex overlaps with the C(1)-OH in β conformer of G6P. In the D10N mutant, after the addition of G6P, the chemical shift of F_A moves downfield 1.1 ppm but F_B moves upfield 1.1 ppm. F_C barely moves. This again reproduces the observation in βPGM_{wt} GSA structures.

It can be concluded that the ^{19}F NMR spectrum of βPGM_{D10N}-BeF_3^--G6P complex does not show any similarity to that of βPGM_{wt}-BeF_3^--G6P in terms of profile or chemical shifts. This suggests that the two rotationally averaged signals at –150.2 and –153.2 ppm may not be induced by the in- and out-conformation of Asp 10. After the D10N mutation, Asn 10 in βPGM_{D10N}-BeF_3^--H_2O complex maintains its 'out' conformation as in wild type. But, different from the two possibilities in wild type, in the βPGM_{D10N}-BeF_3^--G6P GSA complex there is only one 'out' conformation.

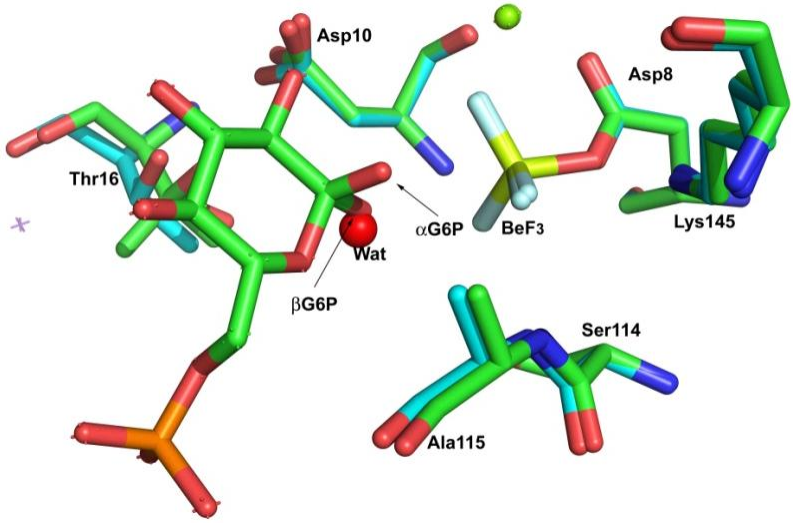


Figure 3.26 Superposed structures of βPGM_{wt}-BeF_3^--H_2O (cyan, PDB: 2WFA) GSA and βPGM_{wt}-BeF_3^--G6P (green, PDB: 2WF9) GSA complexes to show the two structures overlay very well. Beryllium trifluoride GSA moiety is kelly and light blue sticks and catalytic magnesium is green sphere. Asp 10 is in an 'out' conformation in both structures. In βPGM_{wt}-BeF_3^--G6P structure, the sugar in G6P shows occupancy by both α- and β-anomers. The water molecule in βPGM_{wt}-BeF_3^--H_2O complex overlays the C(1)-OH in β-anomer of G6P.

3.3.7 βPGM_{D8E}-AlF_4^--G6P TSA complex

In order to form magnesium and aluminium fluoride TSA complexes of βPGM_{D8E} protein with G6P or βG1P for NMR analysis in solution, experiments were set up as for all other TSA analogue experiments. The ^{19}F NMR shows βPGM_{D8E} does not form a fluoromagnesate complex under these experimental conditions. This may be due to the low stability of MgF_3^- in the solution and probably the longer axial O–Mg–O bond length, which is difficult for the βPGM_{D8E} active site to accommodate. However, ^{19}F NMR of the aluminium fluoride experiment shows three resonances with 1:2:1 integral ratio, thus identifying the βPGM_{D8E}-AlF_4^--G6P TSA complex (**Figure 3.27**). The larger, central resonance is resolved into two peaks of equal intensity in 100% D_2O (used for measuring SIIS). This proves there are four fluorines in the complex coordinating to aluminium in the AlF_4^- moiety. Lastly, βG1P failed to form any TSA complex with βPGM_{D8E}.

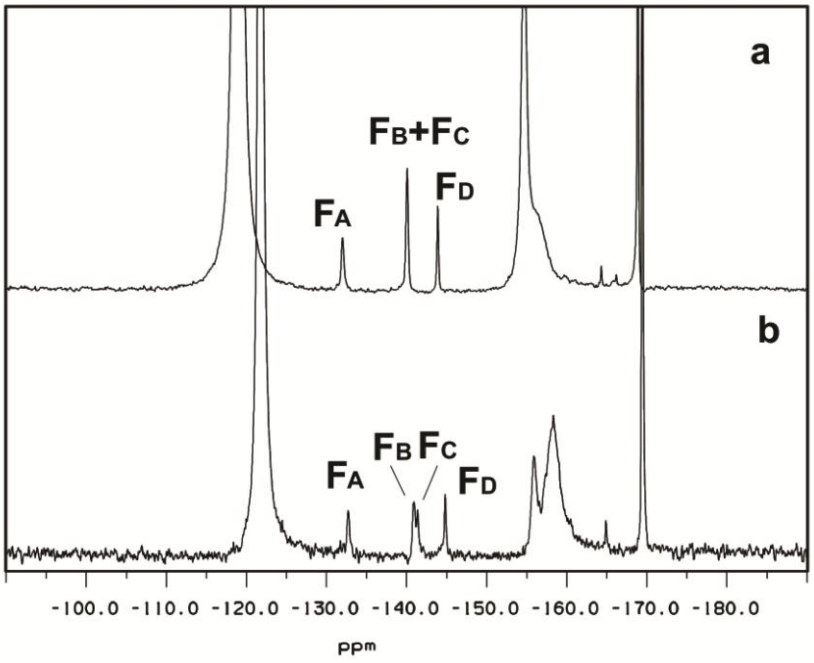


Figure 3.27 ^{19}F NMR spectra showing: (**a**) βPGM_{D8E}-AlF_4^--G6P TSA complex, recorded in 50 mM HEPES, 0.5 mM βPGM_{D8E}, 5 mM $MgCl_2$, 1 mM EDTA, 1 mM $AlCl_3$, 10 mM NH_4F at pH 7.0; free F^- = –119.2 ppm, signals from protein-bound AlF_4^- TSA complex F_A = –131.9 ppm, F_B = F_C = –140.0 ppm, and F_D = –144.0 ppm; (**b**) βPGM_{D8E}-AlF_4^--G6P TSA complex, recorded in the same conditions but in D_2O to separate F_B and F_C; free F^- = –121.9 ppm, signals from protein-bound AlF_4^- TSA complex are F_A = –132.8 ppm, F_B = –140.9 ppm, F_C = –141.4 ppm, and F_D = –144.8 ppm.

In contrast to the big chemical shift and SIIS changes seen in βPGM_{D10N} *vs* wild type protein, in the case of βPGM_{D8E} only F_A and F_B move upfield, 1.3 and 3.0 ppm respectively, relative to βPGM_{wt} in the same complex. The SIIS values are almost the same. If the same assignment is used as in βPGM_{wt}, it is reasonable to say the chemical environment around F_A and F_B does not change much in the TSA complex. The essential role of Asp 8 in nucleophilic catalysis has previously been confirmed by demonstrating that the D8A and D8E mutants are devoid of catalytic activity.(Zhang *et al.*, 2005)

Table 3.12 ^{19}F chemical shifts (ppm) and SIIS (ppm) of the βPGM_{wt}-AlF_4^--G6P and βPGM_{D8E}-AlF_4^--G6P complexes.

Complex	**F_B**	**F_A**	**F_D**	**F_C**
βPGM_{wt}-AlF_4^--G6P	–130.6	–137.0	–140.6	–144.0
SIIS	0.8	1.3	0.9	0.8
βPGM_{D8E}-AlF_4^--G6P	–131.9	–140.0	–140.0	–144.0
SIIS	0.9	1.4	0.9	0.8

βPGM_{D8E} was also put into crystal trials to form an AlF_4^- complex with G6P, so that the details of the conformational change on the nucleophile could be revealed. However, in the resulting 1.6 Å resolution βPGM_{D8E} structure, no catalytic Mg^{2+} or substrate are bound even though the crystal trials were made with Al^{3+}, F^- and G6P. The loop that bears Lys 145 is very disordered (141_V-145_K) (**Figure 3.28**). There is some electron density where the phosphate of G6P should be, but no density from the sugar ring, or from AlF_4^-. It is proposed that the AlF_4^- that is seen in the solution NMR is actually stabilised by the passive phosphate binding site of G6P and the similarity of the chemical shifts of fluorines must be a pure coincidence (**Table 3.12**). If that were the case, the same complex in solution would have been reproduced by using βPGM_{D8E} without the addition of G6P. However, our control experiment shows G6P is essential for the formation of the complex and the signals we see in ^{19}F NMR are truly those of a βPGM_{D8E}-AlF_4^--G6P TSA complex. Therefore this βPGM_{D8E} structure is a structure irrelevant to our observations in solution.

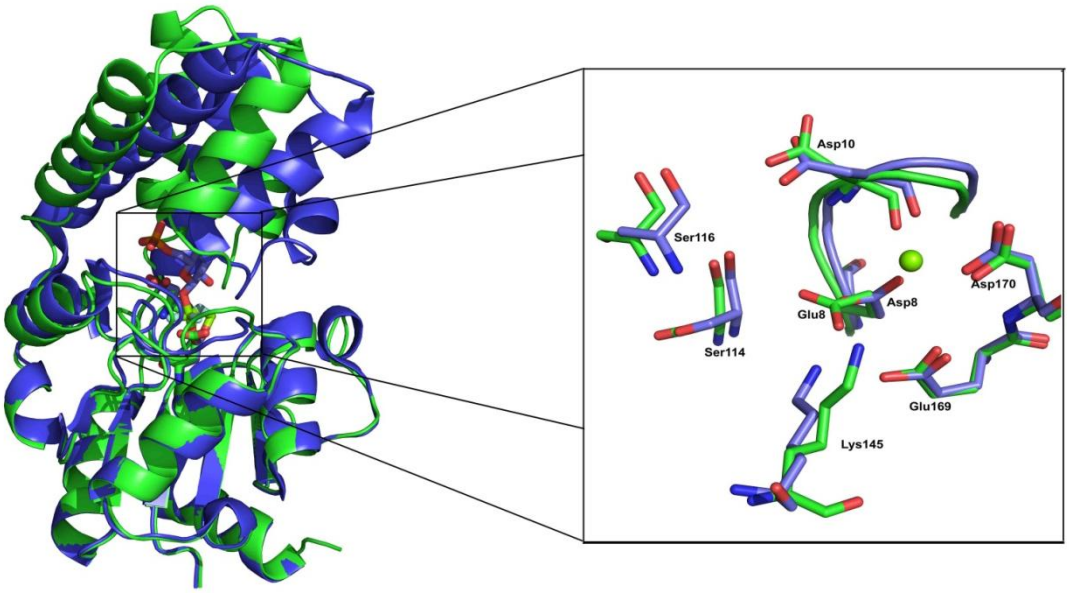


Figure 3.28 Overlay of βPGM_{D8E} with βPGM_{wt}-MgF_3^--G6P TSA complex (in cartoon) to show the substrate-free βPGM_{D8E} is more open (green) than the βPGM_{wt}-MgF_3^--G6P TSA complex (blue). The active site is enlarged to show the detailed difference of residue conformation in which MgF_3^- and G6P are not shown. (The structures are aligned for their core domains).

Based on the likely similar TSA conformation suggested by the similar chemical shifts of the fluorines, we can conclude that the influence of mutating the nucleophilic Asp 8 into a glutamate is not as large as mutating the GB Asp 10 into asparagine. But the binding of G6P in βPGM_{D8E} seems to be weaker than for βPGM_{D10N}.

3.3.8 Attempts to form βPGM$_{D8N}$ metal fluoride complexes

βPGM$_{D8N}$ was designed to prevent formation of a trifluoroberyllate GSA complex on Asp 8, as in ^{19}F NMR there are two broad rotationally averaged peaks which were proposed to be the BeF_3^- moiety mimicking a phosphoryl group attached either on aspartate 8 or C(1)-OH on G6P (Griffin, PhD Thesis, 2011; **Figure 3.29**). The single point mutation of D8N leads to a complete change in the nature of the nucleophile. Asp 8 is a negatively charged carboxylate group which attacks the phosphoryl group in-line on G6P, G1P, or βG16BP whereas the side chain amide of Asn 8 is not charged and the nitrogen in the amide has no nucleophilic activity.

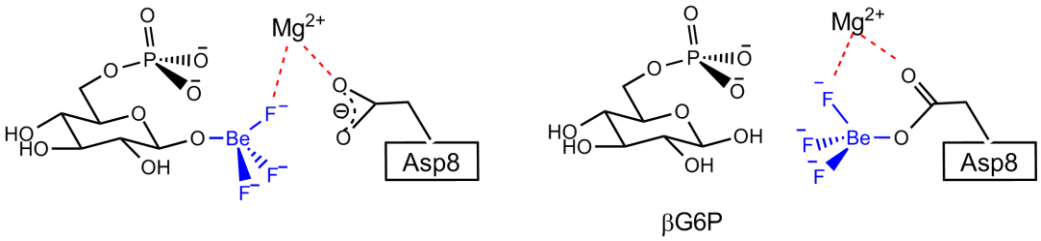


Figure 3.29 Scheme to show the two hypothetic binding modes of BeF_3^- observed in ^{19}F NMR for βPGM$_{wt}$-BeF_3^--G6P GSA complex.

Several conditions were tried to demonstrate the formation of BeF_3^- complexes with βPGM$_{D8N}$. There was no sign of the existence of either broad rotationally averaged or a set of three resolved peaks to represent BeF_3^- complexes except for the single case of wild type βPGM. This result is sensible. In the literature, there is not one case where a BeF_3^- complex can form on an amide nitrogen and there are only rare cases that a BeF_3^- can form on a histidine nitrogen. Because of the nature of beryllium, its small atomic radius leads to largely covalent character in its bonding. Different from fluoromagnesate and fluoroaluminate, fluoroberyllate has near obligate tetrahedral geometry, due to the inaccessibility in beryllium of 3d orbitals required to form higher coordination states.(Chabre, 1990) So to form a bond with an oxygen atom, the oxygen needs to be negatively charged, as in $–COO^-$ and $P–O^-$.

The βPGM$_{D8N}$ mutant gives no MgF_3^- or AlF_4^- complex either under conditions used for wild type or for other mutants, as expected. This result can be explained by the loss of nucleophilicity on the side chain amide of Asn 8 relative to the carboxylate of Asp 8. According to experience previously, the AlF_4^- moiety can only be stabilised when at least one of the two apical oxygen ligands is negatively charged. In this case, the side

chain amide of Asn 8 and the C(1) or C(6)-OH are both non-ionic. Thus no TSA complex can be formed with this mutant.

3.4 Conclusions

Of the six stable synthetic analogues of βG1P, β-1-phosphonomethylene-1-deoxy-D-glucopyranose and (*S*)-1-β-phosphonofluoromethylene-1-deoxy-D-glucopyranose form very stable TSA complexes with both magnesium and aluminium fluoride. The binding of (*S*)-βCHFG1P is stronger than for βCH_2G1P. None of the three 1-α-OH analogues nor the (*R*)-diastereoisomer of the CHFP analogue show signals for complex formation (see **Section 3.3.1**).

The mutant βPGM_{D10N} forms tetrafluoroaluminate complexes with both G6P and βG1P that are TSAs as gauged by ^{19}F NMR analysis (see **Section 3.3.4** and **Section 3.3.5.1)**. It also forms beryllium fluoride complexes with G6P and water (see **Section 3.3.6)**. This means the βPGM_{D10N} mutant can still stabilise the assembly of a GSA and of a TSA. The data do not reproduce either of the two rotationally averaged broad peaks in βPGM_{wt}-BeF_x-G6P. It follows that the essential GABC by Asp 10 during catalysis has an effect on the dynamic properties of BeF_3^-. βPGM_{D10N} can be phosphorylated by acetyl phosphate to give pAsp 8 with a long lifetime (see **Section 3.3.3**). This directly proves Asp 8 is a nucleophile in the catalytic process. The ^{19}F NMR data show good stability for βG16BP in βPGM_{D10N} that shows D10 is essential to complement the nucleophilic activity of Asp 8.

βPGM_{D8E} has a similar potential nucleophile to wild type (Glu 8 rather than Asp 8) although with a different orientation. It forms a tetrafluoroaluminate TSA complex with G6P which reflects similar chemical environments to wild type according to the chemical shift and SIIS (see **Section 3.3.7**). However it fails to form a fluoroberyllate GSA complex as there is no additional stabilisation for the three fluorines. Our ^{19}F NMR data do not show whether βPGM_{D8E} mimics the associative mechanism of phosphotransfer because the only available crystal structure is not relevant to our observations in solution (see **Section 3.3.7**).

βPGM_{D8N} does not form any metal fluoride complexes in solution (see **Section 3.3.8**). Comparing results for βPGM_{D8N} with those for βPGM_{wt}, βPGM_{D8E} and βPGM_{D10N}, we can conclude that an anionic nucleophile is essential for the formation of a beryllium

GSA complex and it must be correctly orientated. Disappointingly, in the ^{19}F NMR spectrum we did not observe any resonance which could possibly resembles the sought for hypothetical Asp8-COO^{-}…BeF_3^{-}-O1-G6P complex as seen in the wild type.

Chapter 4. Phosphoserine Phosphatase

4.1 Introduction

Phosphoserine phosphatase (PSP) (EC 3.1.3.3) belongs to the haloalkanoic acid dehalogenase (HAD) superfamily, subfamily I, as does βPGM. It catalyses the hydrolysis of L-phosphoserine (PLS) to give inorganic phosphate (Pi) and L-serine as shown in **Figure 4.1**.

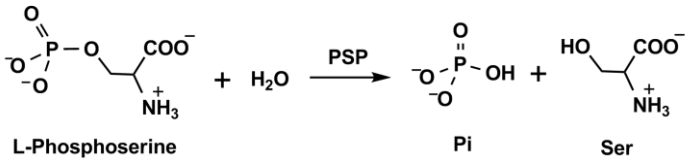


Figure 4.1 The hydrolysis of L-phosphoserine catalysed by phosphoserine phosphatase.

The whole protein is composed of two domains: a cap domain and a core domain (**Figure 4.2**). The active site which consists of a phosphate binding site and a serine-binding site is at the interface between these two domains. When enzyme binds its substrate, it adapts a closed conformation.(Wang *et al.*, 2002)

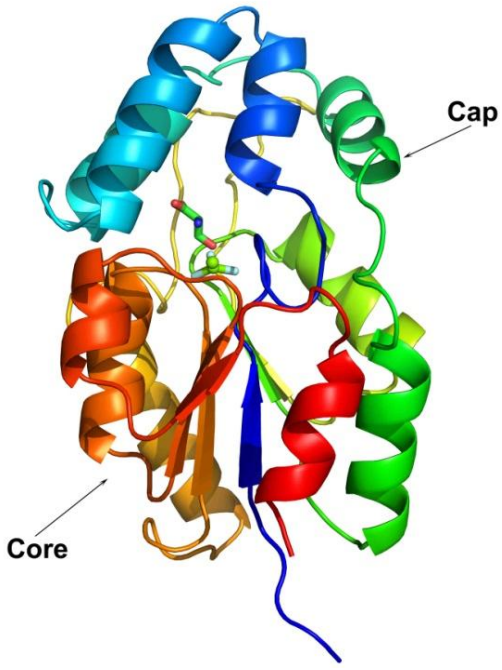


Figure 4.2 Ribbon diagram of the structure of PSP in its closed form with the bound serine and MgF_3^- shown as a stick model and the catalytic Mg^{2+}, represented as a green sphere (Bowler, unpublished).

The study in this thesis is on PSP from *Methanocaldococcus jannaschii*. The structural details of the whole catalytic cycle were originally established by David Wemmer and refined by work in Sheffield from the laboratory of Professor Jon Waltho. The five

stages sequentially utilize (i) a combination of real substrate and single point mutation of the protein to illustrate the substrate binding (PSP_{D11N}-PLS), (ii) the wild type protein, L-serine and MgF_3^- (or AlF_4^-) to mimic the first transition state of the phosphorylation of the enzyme involving Asp 11 (PSP-MgF_3^--Ser), (iii) the wild type protein, BeF_3^- to mimic the phosphoenzyme intermediate pAsp 11 (PSP-BeF_3^-), (iv) the wild type protein, MgF_3^- and water to demonstrate the second transition state of autodephosphorylation of pAsp 11 (PSP-MgF_3^--H_2O), and (v) the wild type protein and inorganic phosphate to demonstrate the product state of the reaction (PSP-P_i).(Wang *et al.*, 2002; Baxter *et al.*, 2008) Some recent computational results indicate that the character of the transition state is geometrically associative yet electronically dissociative.(Re *et al.*, 2011)

So far in work in Sheffield, the formation of the PSP-MgF_3^--H_2O, PSP-MgF_3^--Ser, PSP-AlF_4^--H_2O and PSP-AlF_4^--Ser TSA complexes and subsequent ^{19}F NMR analysis had been achieved by Dr Nicola Baxter. In PSP-MgF_3^- (or AlF_4^-)-Ser complex, the F_C chemical shift moves downfield a hydrogen bonding interaction from the amine of substrate L-serine to F_C. This hydrogen bond does not exist when water is the nucleophile in the second half of the reaction, therefore the binding of L-serine can be indicated directly by the downfield shift of F_C. Later, L-serine titration identified that the binding constant of L-serine is a lot higher in the PSP-AlF_4^--Ser TSA complex than in the PSP-MgF_3^--Ser TSA complex. In the aluminium complex, L-serine even at 70 mM cannot outcompete water (55 M in solution) (Griffin, PhD thesis, 2011). But the same setup in βPGM shows that G6P even below 5 mM can saturate the active site in the βPGM-AlF_4^--G6P TSA complex. This result underlies the difference between βPGM and PSP in their catalytic function and our study has been aimed to pinpoint evidence to illustrate the similarity and difference in the structures which then influence the chemistry and kinetics of the reaction.

Comparing PSP with βPGM, both use a phosphorylated aspartate (Asp 11 in PSP and Asp 8 in βPGM) as an intermediate after dephosphorylation of the substrate, both employ a catalytic magnesium ion to link the transferring phosphoryl group to the nucleophilic aspartate, and both of them use an aspartate (Asp 13 in PSP and Asp 10 in βPGM) as a GAB to facilitate nucleophilic attack (**Figure 4.3**). The conformation of Asp 10 in βPGM is well identified by both crystallography and NMR, as described in **Chapter 3**. During the phosphoryl transfer for βPGM, Asp 10 adopts the 'in'

conformation to form a hydrogen bond with C(1)OH of G6P. After Asp 8 is phosphorylated, Asp 10 flips out of the active site and is held by a strong hydrogen bond to the backbone amide of Thr 16. This rearrangement of hydrogen bonds is proposed to be a mechanism for phosphomutase activity in order to reduce unwanted phosphoenzyme hydrolysis while the phosphosugars are interchanging in the active site. The existence of a relatively stable phospho-aspartate (pAsp 8) in the D10N mutant proves this.

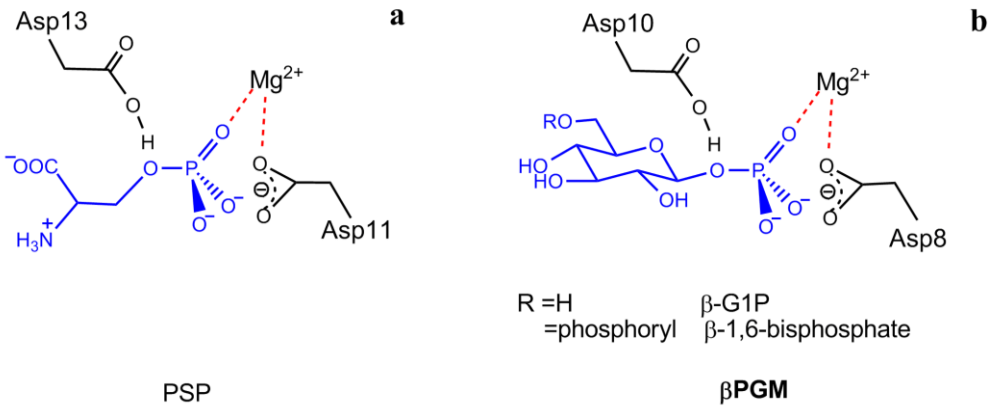


Figure 4.3 (**a**) PSP and (**b**) βPGM both have a nucleophilic aspartate, a general acid/base aspartate and an electrophilic magnesium.

By contrast, because of its role as a phosphatase and in order to complete its catalytic cycle, PSP must favour autodephosphorylation of pAsp 11 once the product L-serine has been released. Based on all the structures that have been published, Asp 13 always stays in an 'in' position and can act as a GABC to facilitate nucleophilic attack by both L-serine and water (**Figure 4.4**).

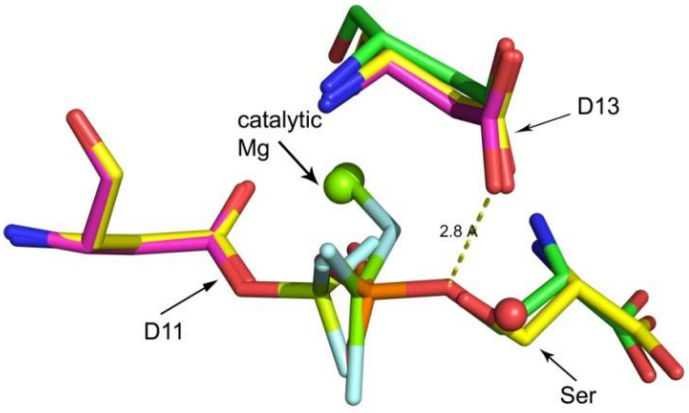


Figure 4.4 Asp 13 stays in the 'in' position in PSP_{D11N}-PLS GS complex (PDB: 1L7P, green sticks), PSP-MgF_3^--Ser TSA complex (yellow sticks), and PSP-BeF_3^--H_2O intermediate state complex (PDB: 1J97, magenta sticks). Magnesium trifluoride TSA moiety is green and light blue sticks, and beryllium trifluoride GSA moiety is kelly and light blue sticks. Catalytic magnesiums are represented as green spheres.

As no existing studies have yet used NMR to look at the effect of modifying Asp 13, our aim in designing a D13N mutant was to investigate the function of Asp 13, considered to be a critical residual in the active site as a catalytic GAB. There has been extensive literature to suggest that Asp 13 (Asp 10 in βPGM) is a GA that protonates the L-serine β-oxygen in bound substrate in order to help the L-serine leave and then comes along as a GA to assist the deprotonation of inorganic phosphate (**Figure 4.5**).(Collet *et al.*, 1998; Wang *et al.*, 2002) We nominated asparagine as the best candidate for replacing Asp 13, to compare it against βPGM_{D10N} mutant as a positive control in two aspects. Firstly, asparagine has the same shape as aspartic acid to maintain the conformation of the active site. Secondly, at neutral pH, unlike the –OH in the carboxyl group of Asp (p*K*a 4.5), the $–NH_2$ in amide group of Asn is not capable of acting as a GAB and that kills the catalytic activity of the enzyme.

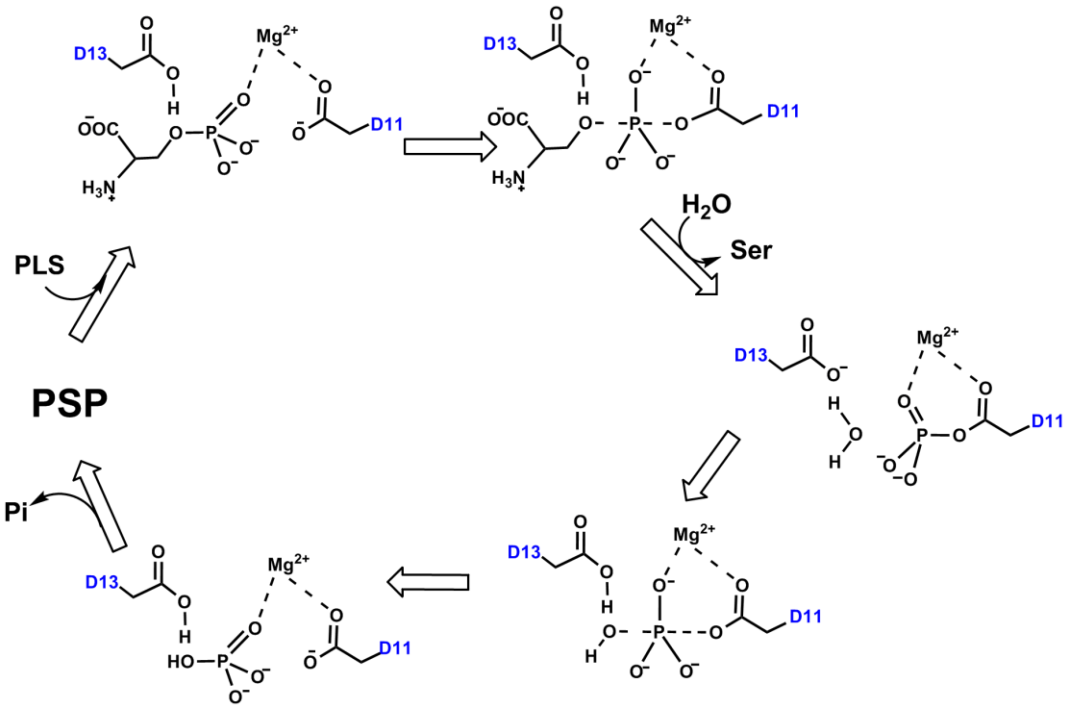


Figure 4.5 The PSP catalytic cycle.

Because of the similarities in structures and general catalytic mechanism to βPGM, PSP_{D13N} was made as a mutant equivalent to βPGM_{D10N}. It has been studied in the same way by using a combination of crystallography and NMR. This chapter describes how the metal fluoride complexes PSP_{D13N}-MgF_3^--H_2O and PSP_{D13N}-AlF_4^--H_2O are formed in solution and the ^{19}F NMR chemical shifts and SIIS for each complex are analysed. ^{31}P NMR is used to detect if there is phosphorylation on Asp 11, and the activity of

mutated enzyme D13N is measured. So far, the setup for all the same GSA, TSA, and PSA complexes have been tried and compare them with wild type enzyme for PSP as has been done for βPGM. In the end of this chapter, comparisons between PSP_{D13N} and βPGM_{D10N} as well as their wild types are made to identify functional and constitutional differences between PSP and βPGM. Encouragingly, the resemblance between them is quite limited.

We would have liked to obtain crystal structures of PSP_{D13N}-MgF_3^--H_2O and PSP_{D13N}-AlF_4^--H_2O complexes, however due to time constraints these results are not available to be included in this thesis.

4.2 Materials and Experiments

4.2.1 Site-directed mutagenesis of PSP D13N mutant

4.2.1.1 Expression vector

The plasmid carrying phosphoserine phosphatase (PSP) gene from *M. Jannaschii* was provided by University of California (Prof. D. Wemmer group). The PSP gene was cloned into the expression system pET-21a(+) vector (Novagen) at restriction cleavage sites *Nde*I 238 and *BamH*I 198 (**Figure 4.6a**).

The sequence of PSP gene in this plasmid was confirmed both by restriction digestion using *Nde*I and *BamH*I and sequencing (Medical School，The University of Sheffield). The agarose gel is shown below (**Figure 4.6b**).

The sequence of PSP is shown below:

```
        10         20         30         40         50         60
MEKKKKLILF DFDSTLVNNE TIDEIAREAG VEEEVKKITK EAMEGKLNFE QSLRKRVSLL

        70         80         90        100        110        120
KDLPIEKVEK AIKRITPTEG AEETIKELKN RGYVVAVVSG GFDIAVNKIK EKLGLDYAFA

       130        140        150        160        170        180
NRLIVKDGKL TGDVEGEVLK ENAKGEILEK IAKIEGINLE DTVAVGDGAN DISMFKKAGL

       190        200        210
KIAFCAKPIL KEKADICIEK RDLREILKYI K
```

Number of amino acids is 211; molecular weight is 23593.6; theoretical pI is 7.60.

4.2.1.2 Site-directed mutagenesis by PCR

The mutagenic oligonucleotide primers were designed to include the desired mutation D13N and synthesized by Sigma Genosys. Oligonucleotide primers were 45 bases long, T_m was 64°C and GC% = 22%.

Primer:

Forward 5'-GCTTATTTTATTTGATTTT**a**ATAG**t**ACATTAATTAATAATGAGAC-3'

Reverse 5'-GTCTCATTATTAATTAATGT**a**CTAT**t**AAAATCAAATAAAATAAGC-3'

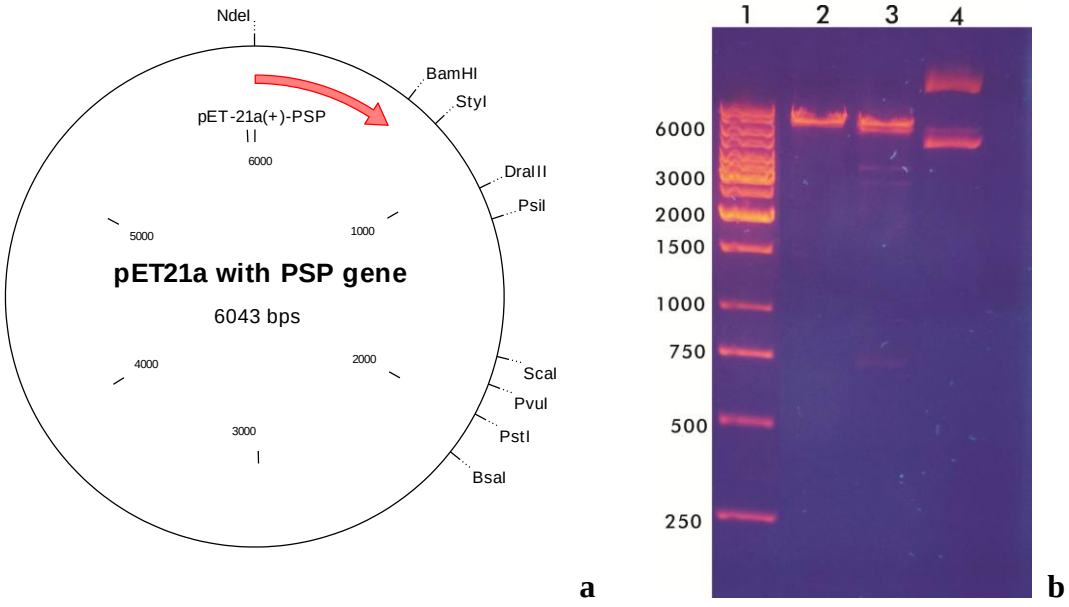


Figure 4.6 (**a**) pET-21a (+) vector with PSP gene. (**b**) Agarose gel showing the restriction digestion to identify the PSP gene in plasmid. Lane 1: DNA 1 kb ladder. Lane 2: PSP plasmid digested by *Nde*I. Lane 3: PSP plasmid digested by *BamH*I. Lane 4: PSP plasmid from miniprep.

The PCR reaction was carried out with 20 cycles, considering the size of the PSP plasmid is 6.04 kb and there was a single point mutation (D13N) and a single point silent mutant (S14), so a 6 min 20 s extension time at 68°C was given presuming the polymerase can synthesize approximate 1 kb per min. Linear DNA (in 2 µL) from the PCR product was transformed into fresh XL1-Blue competent cells following the procedure in section 2.7 and the plasmid was amplified following mini-prep procedure in **Section 2.8**.

The conditions of the PCR are listed in **Table 4.1**.

Table 4.1 PCR setup conditions for PSP_{D13N} single point mutation.

Content	Volume / µL			
	Reaction 1 (Control)	**Reaction 2**	**Reaction 3**	**Reaction 4**
dd H_2O	39.5		24	
10× Buffer	5		5	
Forward primer 12.6 µL/mL	1.25		10	
Reverse primer 16.6 µL/mL	1.25		8	
Template DNA	2	1	0.5	2
dNTs	1		1	
pfuTurbo polymerase	1		1	
Total Volume/ µL	50	50	49.5	51
Segment	Temperature / ℃		Time Length / min	
1	95		1	
2	55		1	
3	68		6	

4.2.1.3 Mutative plasmid DNA sequencing

The single point mutation of PSP (D13N) and the silent mutant (S14S) were first confirmed by restriction digestion using *Rsa*I then by sequencing (Medical School，Sheffield University). The agarose gel is showed below (**Figure 4.7**).

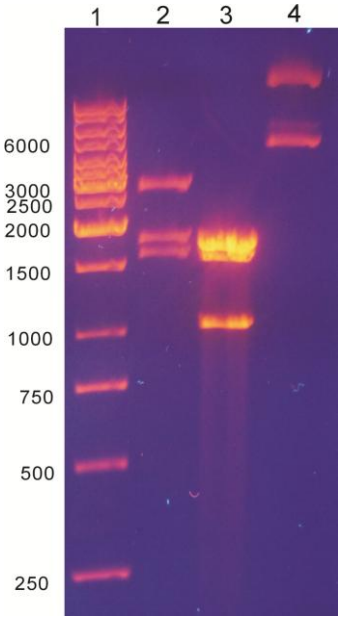


Figure 4.7 Agarose gel showing the restriction digestion of mutant plasmid. Lane 1– DNA 1kb ladder. Lane 2– PSP plasmid digested by *Rsa*I into 3 pieces: 2716, 1760, 1567 bp. Lane 3 – PSP_{D13N} mutative plasmid digested by *Rsa*I into 4 pieces: 1760, 1702, 1567, 1024 bp. Lane 4 – PSP_{D13N} plasmid from miniprep.

4.2.2 Cell Growth and Expression

Phosphoserine phosphatase (PSP) from *M. jannaschii* was subcloned into pET21a vector and expressed in *E. coli.* strain BL21-GOLD(DE3). The expression of PSP and

D13N were both carried out in minimal media with $^{15}NH_4Cl$ to get ^{15}N-labelled protein. To help the bacteria adapt to the lean expression environment, it is better to make the starter culture in M9 media and pre-warm the large flasks of M9 media to 37°C before inoculation. Each 500 mL minimal media with in 8 × 2 L was inoculated by 5 mL overnight culture and incubated at 37°C at 250 rpm until OD reached 0.8. The expression was induced by 0.5 mM IPTG being followed by another 6 h incubation at 30°C. Cells were harvested by centrifugation at 7,000 rpm at 4°C for 30 min. Cell pellets were kept at –80°C for long time storage or lysed straight away.

4.2.3 Purification

PSP_{wt} and PSP_{D13N} were purified following the same protocol. The cells were resuspended in lysis buffer and lysed by sonication in 50 mM HEPES, pH 7.5 and protein extract was made clear by centrifugation at 24,500 rpm, 4°C for 45 min. The supernatant was then heated at 70°C before putting it on to the Q Sepharose Fast Flow column. The purpose of heating is to remove most of the non-PSP proteins from *E. coli.* which can be denatured at high temperature.

4.2.3.1 Ion exchange chromatography

An ion exchange column of dimensions 5×2 cm, packed with Q Sepharose Flat Flow resin (GE Healthcare, Cat. No. 17-0510-01). was equilibrated with ion-exchange buffer. The cell supernatant which had been filtered using a 0.45 μm filter was loaded onto the column. The column was run at 3 mL/min and all the flow through was collected using a 50 mL pot. A NaCl salt gradient (0~1 M) was put on the column using an ÄKTAprime™ plus FPLC until the UV trace returned to baseline. All fractions were collected. The component of the fractions were determined by SDS-PAGE (**Figure 4.8**) and those containing PSP_{wt} and PSP_{D13N} were collected and concentrated to 2 ~ 3 mL using a Sartorius Vivaspin^{TM} at 5,500 g, 4°C.

Table 4.2 Protocol for ion exchange buffers for PSP_{wt} and PSP_{D13N}.

Ion Exchange Buffer **1**	20 mM	Tris-HCl, pH 8.3
	10 mM	DTT
	2 mM	EDTA
	1 mM	NaN_3
Ion Exchange Buffer **2**	20 mM	Tris-HCl, pH 8.3
	10 mM	DTT
	2 mM	EDTA
	1 M	NaCl
	1 mM	NaN_3

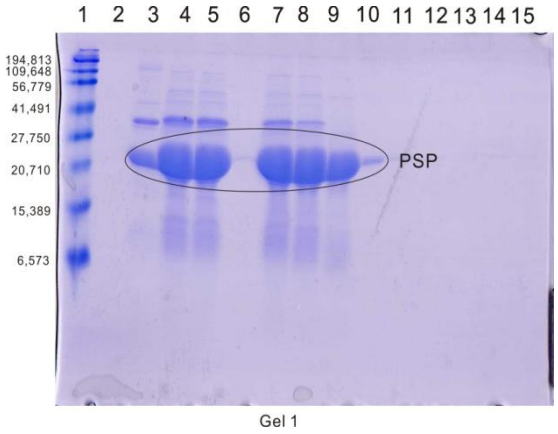


Figure 4.8 SDS-PAGE of fractions collected from ion-exchange chromatography of PSP_{wt}.

4.2.3.2 Gel filtration chromatography

PSP has a low content of residues containing UV chromophoric groups (tryptophan 0%, phenylalanine 3.3%, tyrosine 1.4%), and this makes the PSP peak in the UV trace looks not as high as some of the other peaks, even though from the gel it looks major.

Table 4.3 Protocol for Gel Filtration Buffers for purification of PSP_{wt} and PSP_{D13N}.

Gel Filtration Buffer	20 mM	Tris-HCl , pH 7.5
	1 mM	EDTA
	300 mM	NaCl
	10 mM	DTT
	1 mM	NaN_3
	filtered and degassed before use	

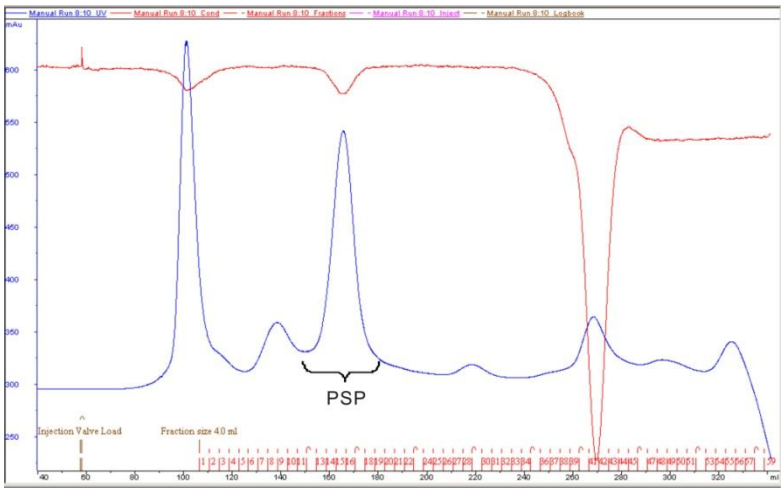


Figure 4.9 UV trace (280 nm) from gel-filtration chromatography of PSP_{wt}.

After gel filtration purification, SDS-PAGE shows that the purity of PSP_{wt} and PSP_{D13N} are pure enough for crystallization and NMR purposes (**Figure 4.10**).

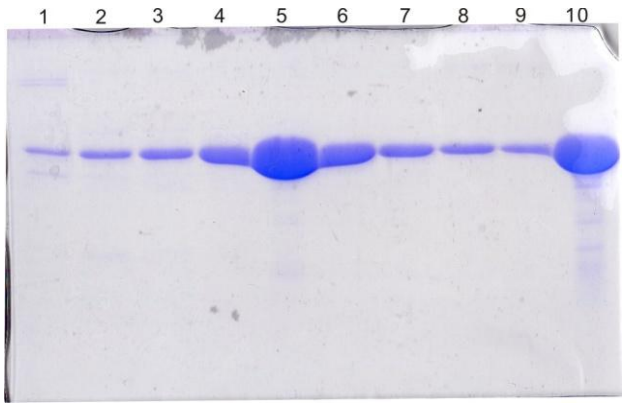


Figure 4.10 SDS-PAGE of fractions collected from gel-filtration chromatography of wild type PSP. Lane 1-9 – Fractions 10-19 from gel-filtration chromatography; Lane 10 – concentrated stock taken forward from ion exchange column.

4.3 Results and Discussion

4.3.1 Activity assay of PSP_{D13N} monitored by ^{31}P NMR

The activity of PSP_{D13N} has not been reported before. So an activity assay monitored by ^{31}P NMR is set up to establish the role of Asp 13 as a GAB. A control was used first to test the stability of L-phosphoserine at neutral pH. No new peaks were observed even after 0.5 M L-phosphoserine was kept at rt for one week, which rules out the possibility of its spontaneous hydrolysis.

For the activity assay, PSP_{D13N} (1 mM) was added to the PLS sample and the acquisition of ^{31}P NMR data was begun immediately after locking and shimming. (**Figure 4.11**). As expected, in the activity assay of PSP_{D13N}, the peak intensity of PLS did not change over 20 h, which shows that this mutant has no activity towards PLS.

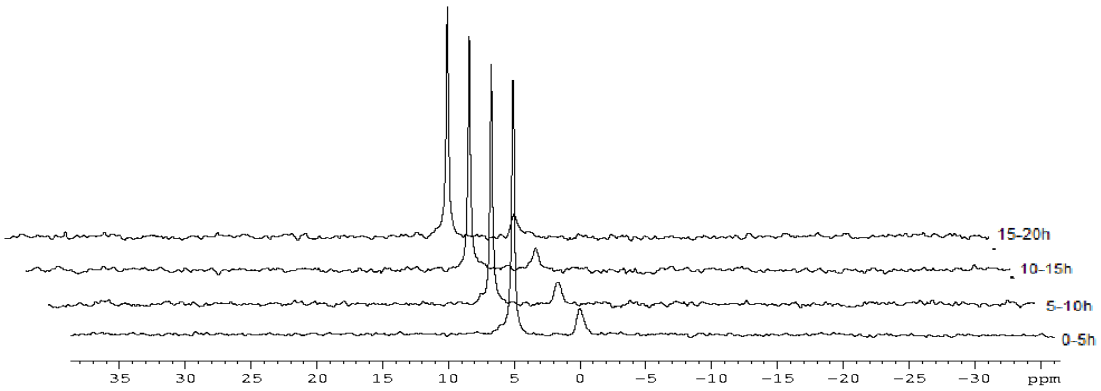


Figure 4.11 Activity assay of PSP_{D13N} monitored by ^{31}P NMR. Time course began after 1 mM PSP_{D13N} was added to 5 mM PLS solution in the buffer of 20 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM EDTA, 10 mM $MgCl_2$, and 1 mM NaN_3. The resonance at 5 ppm is PLS, and the small peak at 0 ppm is protein-bound inorganic phosphate.

Several causes can affect the activity of PSP_{D13N}. First substrate binding might be impaired. In the crystal structure of the D11N mutant PSP_{D11N}-PLS (PDB: 1L7P), the side chain carbonyl of Asn 11 has reduced electron density to make it bind to the catalytic magnesium more weakly. The protein has to compromise the conformation of Asn 11 to be able to bind PLS in the same way as wild type PSP. The net result of this mutation is Asn 11 flips away from the phosphate in PLS and there is no catalytic Mg^{2+} bound. In PSP_{D13N} there is no such problem. It still has Asp 11 to bind the catalytic magnesium and we can also infer from the TSA structure that the carbonyl should be held steadily by a 2.9 Å backbone amide and 2.7 Å side chain hydroxyl of Thr 21. This keeps the side chain amide of Asn 13 still in place to form a hydrogen bond with the βO in PLS but not be able to protonate it. Second, the β-oxygen in PLS has a pK_a ~14 so it is a bad leaving group and the attacking carboxylate $–COO^-$ of Asp 11 has a pK_a ~5, thus the reaction can happen only when the protonated carboxylic acid Asp 13 donates its proton to the β-oxygen in TS for L-serine which can then depart. In the PSP_{D13N} mutant, the pK_a of side chain amide of Asn 13 is well above 15, so it cannot be a GA. Therefore the reaction cannot proceed.

4.3.2 PSP_{D13N}-MgF_3^--H_2O and PSP_{D13N}-AlF_4^--H_2O TSA complexes

To make a PSP_{D13N}-MgF_3^--H_2O TSA complex for ^{19}F NMR study, 20 mM NH_4F was added to 0.5 mM PSP_{D13N} in 20 mM Tris-HCl, 300 mM NaCl, 1 mM EDTA, 10 mM $MgCl_2$, and 1 mM NaN_3 at pH 7.5. The complex shows three peaks: F_A = –142 ppm, F_B = –146 ppm, and F_C = –175 ppm having equal integration. A PSP_{D13N}-AlF_4^--H_2O TSA complex was formed by adding 2 mM $AlCl_3$ to the MgF_3^- complex sample and it gives four resolved peaks: F_A = –131 ppm, F_B = –137 ppm, F_C = –140 ppm, and F_D = –153 ppm (**Figure 4.12**).

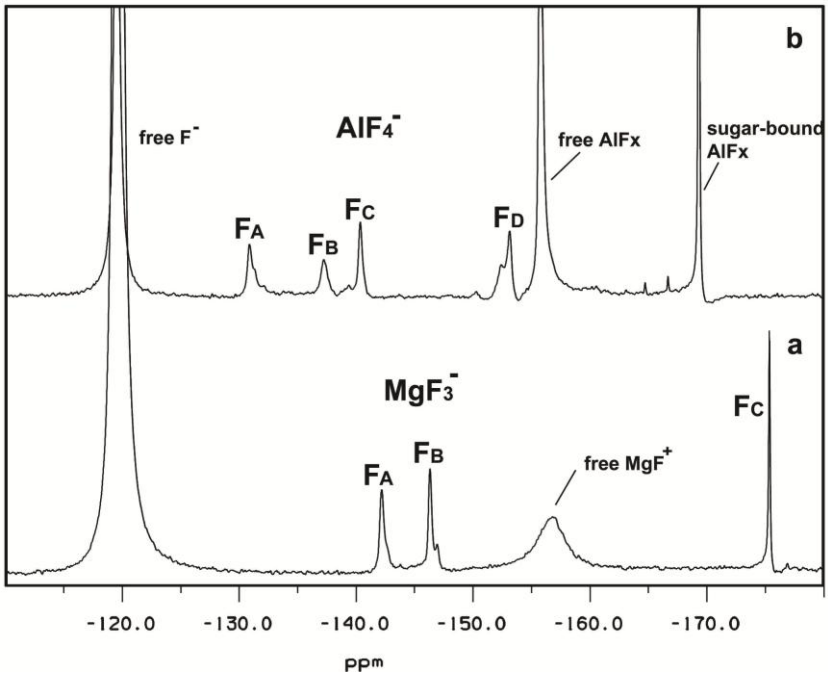


Figure 4.12 ^{19}F NMR spectra of: (**a**) PSP_{D13N}-MgF_3^--H_2O TSA complex in 20 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM EDTA, 10 mM $MgCl_2$, 20 mM NH_4F and 1 mM NaN_3; The protein-bound MgF_3^- TSA signals are F_A= –142 ppm, F_B= –146 ppm, and F_C= –175 ppm; (**b**) PSP_{D13N}-AlF_4^--H_2O TSA complex in the same conditions as **(a)** with supplement of 2 mM $AlCl_3$; The protein-bound AlF_4^- TSA signals are F_A= –131 ppm, F_B= –137 ppm, F_C= –140 ppm, and F_D= –153 ppm.

4.3.2.1 Binding of L-serine in both TSA complexes

Related work (Griffin, PhD Thesis, 2011) has shown that L-serine binds weakly to the trifluoromagnesate and tetrafluoroaluminate TSA complexes in PSP_{wt}. For wild type PSP, 10 mM L-serine cannot out-complete all the H_2O to turn PSP_{wt}-MgF_3^--H_2O complex into a PSP_{wt}-MgF_3^--Ser complex. For PSP_{wt}-AlF_4^--Ser TSA complex, 20 mM L-serine only replaces half of the water ligand in the PSP_{wt}-AlF_4^--H_2O TSA complex (K_d =13.17 ± 2.1 mM). For PSP_{D13N}, the binding of L-serine is even weaker. It can be deduced from the ^{19}F NMR spectra that when 20 mM L-serine is added to the sample of PSP_{wt}-MF_x^--H_2O TSA complexes there is no change of the spectra. That means it is still the water molecule not L-serine that binds at the axial position to form PSP_{D13N}-MgF_3^--H_2O and PSP_{D13N}-AlF_4^--H_2O TSA complexes. This shows Asp 13 is critical as it deprotonates the L-serine hydroxyl to stabilise its bonding to the central metal (**Figure 4.13**).

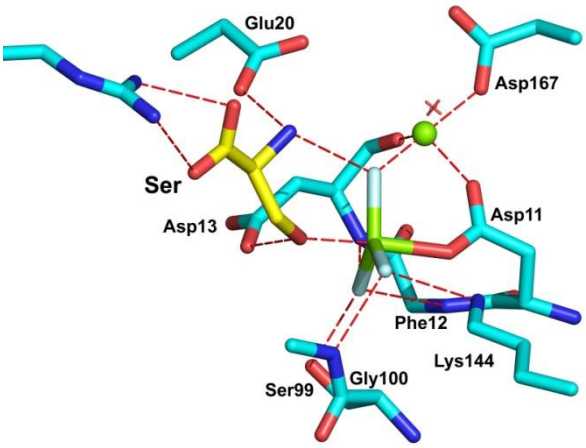


Figure 4.13 Structure of PSP_{wt}-MgF_3^--Ser TSA complex (unpublished). The protein residues are cyan sticks, catalytic magnesiums is represented as kelly sphere, magnesium trifluoride moiety is kelly and light blue sticks, and serine is yellow sticks.

As described previously, PSP_{D13N} has Pi bound to the active site after purification. However, after fluoride is added, the PSP_{D13N}-MgF_3^--H_2O complex forms in the active site with almost complete occupancy. This simply means the binding of trifluoromagnesate moiety is much stronger than that of inorganic phosphate. This is because the enzyme binds the transition state, and hence its analogue, more strongly than it binds product.

4.3.2.2 PSP_{D13N}-MgF_3^--H_2O TSA complex

The differences in the ^{19}F chemical shifts between PSP_{wt} complexes and PSP_{D13N} complexes on MgF_3^--H_2O TSA complex can be compared (**Table 4.4**)

Table 4.4 ^{19}F chemical shift (ppm) and SIIS (ppm) of PSP_{wt}-MgF_3^--H_2O complex and PSP_{D13N}-MgF_3^--H_2O complex.

Complex	F_A	F_B	F_C
PSP_{wt}-MgF_3^--H_2O	–141	–144	–176
SIIS	1.6	1.5	0.4
PSP_{D13N}-MgF_3^--H_2O	–142	–146	–175
SIIS	1.4	1.4	0.3

From **Figure 4.12** and **Table 4.4**, it is clear that the PSP_{D13N}-MgF_3^--H_2O complex isn't much different from the PSP_{wt}-MgF_3^--H_2O TSA complex. The subtle difference may be from the different electron density on the attacking water according to the charge which corresponds to the mutated residue Asp 13→Asn 13. In the case of wild type PSP, Asp 13 acts as a GB to the water to make it more nucleophilic (**Figure 4.14**, left).

Asn 13 instead is only able to accept a lone pair of electrons from water to form a hydrogen bond, as a result the attacking water is *less* nucleophilic (**Figure 4.14**, right). Therefore, if the conformation in the active site of PSP_{D13N} stayed the same as wild type, we might expect that the electron density in the MgF_3^- moiety should decrease and the average chemical shifts of the fluorine move downfield. However, comparing to PSP_{wt} only F_C moves 1 ppm downfield, conversely F_A moves 1 ppm upfield and 2 ppm for F_B.

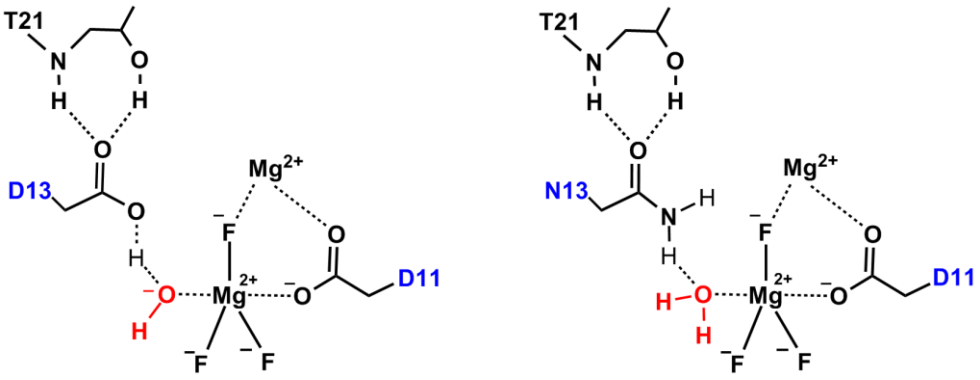


Figure 4.14 Proposed PSP_{D13N}-MgF_3^--H_2O TSA structure.

A detailed explanation is not possible in the absence of a structure of this complex. The SIIS values are very similar to those in the wild type enzyme, suggesting the trifluoromagnesate may maintain its tbp geometry and similar interactions with the rest of the active site. However, on comparing PSP_{D13N} to the equivalent mutant βPGM_{D10N}, such a small change of electron density as indicated by the small changes of chemical shift is unexpected!

4.3.2.3 PSP_{D13N}-AlF_4^--H_2O TSA complex

The differences in the ^{19}F chemical shifts between PSP_{wt} complexes and PSP_{D13N} complexes on AlF_4^--H_2O TSA complex are compared (**Table 4.5**) From **Table 4.5**, it is obvious that in the PSP_{D13N}-AlF_4^--H_2O TSA complex the chemical shift of F_A and F_C move downfield by 2 and 1 ppm compared to those of wild type PSP. Conversely, F_B and F_D move upfield 3 and 4 ppm, respectively, which indicates that F_B and F_D in the PSP_{D13N} mutant become more shielded. The bigger chemical shift changes in PSP_{D13N}-AlF_4^--H_2O complex relative to those of PSP_{D13N}-MgF_3^--H_2O complex and the wild type enzyme (**Table 4.5**) suggest that the conformational change resulting from the D10N mutation for tetrafluoroaluminate is bigger than that for trifluoromagnesate. It is known that octahedral AlF_4^- has a greater geometrical flexibility than trigonal MgF_3^-, but the

conflict between the larger isotope shift for F_B, indicating more hydrogen bonding interactions with other hydrogen donors, and its chemical shift moving upfield, suggesting it is more shielded, cannot yet be resolved. So the mismatch between the chemical shift data and SIIS data might be caused by some major conformational changes in the active site of this mutant complex corresponding to the octahedral tetrafluoroaluminate moiety. It is difficult to deconvolute the chemical shift data by simply analysing the total change of the electron density in the fluorometallic moiety, as the chemical shift of fluorine is so sensitive that it can reflect very subtle variations of the surrounding chemical environment.

Table 4.5 ^{19}F chemical shift (ppm) and SIIS (ppm) of PSP_{D13N}-AlF_4^--H_2O TSA complex.

Complex	**F_A**	**F_B**	**F_C**	**F_D**
PSP_{wt}-AlF_4^--H_2O	–133	–134	–141	–149
SIIS	1.0	0.6	1.0	0.3
PSP_{D13N} -AlF_4^--H_2O	–131	–137	–140	–153
SIIS	1.2	1.2	0.9	0.1

We tried to use structural information from the PSP_{wt}-AlF_x-H_2O complex that might be helpful to us in explaining this mismatch between chemical shift and SIIS data. However the only available structure of PSP_{wt}-AlF_3/AlF_4-H_2O complex is for a dual occupancy mixture with a 3:2 ratio between octahedral and tbp complexes (PDB: 1L7N, **Figure 4.15**). The positions of the metal fluoride moieties in this crystal structure are not reliable as it appears that the simultaneous fitting of the different species has resulted in some distortion. The equatorial plane is skewed from being perpendicular to the O–Mg–O axis as a result of the misfitting of the O–Mg bond length by the shorter O–Al bond length. Therefore the crystal structures cannot be used to rationalise the observed deuterium isotope shifts or chemical shifts, even for wild type protein. To avoid over-interpreting our limited NMR data, it is critical to have a crystal structure of this mutant to be able to interpret the NMR data correctly.

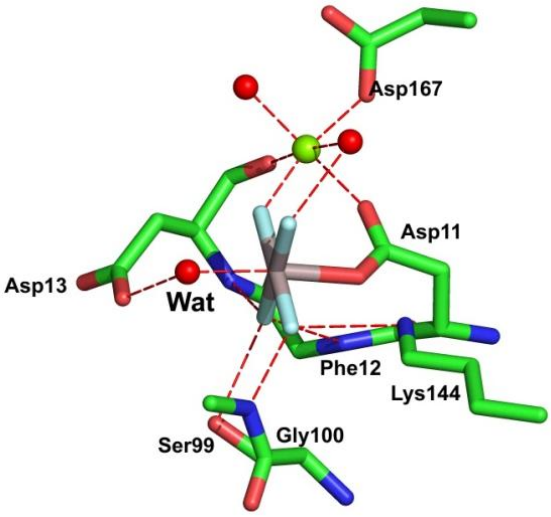


Figure 4.15 Structure of PSP_{wt}-AlF_4^--H_2O TSA complex extracted from a structure of PSP_{wt}-AlF_3/AlF_4-H_2O complex 60% occupancy of the octahedral complex (PDB: 1L7N). The protein residues are represented as green sticks, catalytic magnesiums is kelly sphere, aluminium tetrafluoride moiety is grey and light blue sticks, and water molecules are red sphere.

4.3.3 Attempts to make a PSP_{D13N}-BeF_3^--H_2O GSA complex

Since βPGM_{D10N} is able to form a βPGM_{D10N}-BeF_3^--H_2O complex as described in **Chapter** 3, PSP_{D13N}, being comparable to βPGM_{D10N}, was used to make a BeF_3^- complex with or without serine present. The provisional results suggest that a BeF_3^- complex is formed under experimental conditions of 20 mM Tris-HCl (pH 7.5), 300 mM NaCl, 10 mM DTT, 1 mM NaN_3, 1 mM EDTA, 5 mM $MgCl_2$, with 0.2 mM PSP_{D13N}, 5 mM $BeCl_2$ and 20 mM NH_4F. Because of the low protein concentration and small number of scans, the final spectrum was not good enough to measure chemical shifts and integration. Therefore its details are not discussed here. Serine (20 mM) was added to above sample to make a PSP_{D13N}-BeF_3^--Ser, but ^{19}F NMR did not seem to change; therefore we conclude the PSP_{D13N}-BeF_3^--Ser complex cannot form at this serine concentration.

4.4 Conclusion from Comparisons between wild type PSP and βPGM and their mutants

Wild type PSP uses the same mechanism as βPGM to catalyse the transfer of a phosphoryl group from an R–OH to a carboxylate. Both of them employ:

- Nucleophilic aspartate (10 *vs* 8)
- Electrophilic/catalytic magnesium
- General acid/base aspartate (13 *vs* 10)

- In-line phosphoryl transfer from substrate to enzyme
- In-line phosphoryl transfer from enzyme to water

The difference is that βPGM is mainly a mutase and has only minor activity as a hydrolase. Once it has been phosphorylated, it acts as a mutase and it catalyses the phosphorylation between C(1)-OH and C(6)-OH while it must allow the hexose to switch orientation within the active site without hydrolysis of the phosphoenzyme. By contrast, PSP is a hydrolase and it deals with the dephosphorylation of the phosphoenzyme with the most abundant nucleophile, water. To deal with alternative substrate complexes and achieve its two catalytic functions, PSP makes dual use of the ionic phosphate of PLS firstly to bind the substrate in the active site and secondly to orientate the phosphoryl group for activation to achieve the transition state for transfer to Asp 11. By contrast, in the priming step of its catalytic process, βPGM uses the 6-phosphate of βG16BP to bind in the substrate passive site while the 1-phosphate is orientated for activation in the catalytic site to achieve the transition state for transfer to Asp 8.

Bearing in mind that all metal fluoride complexes that simulate phosphoryl transfer have to be 'synthesised' by the enzyme, whether a metal fluoride TSA complex can be detected or not entirely depends on whether it can be stabilised by wild type or by mutant enzyme. Therefore, these analogues only represent the nature of the interactions but do not reflect the free energy of the real phosphates or phosphoryl group.

Table 4.6 Comparison of the general acid-base role of aspartate residues revealed in complexes formed with PSP_{D13N} and βPGM_{D10N} mutants and wt enzymes.

States	**PSP_{wt}**		**PSP_{D13N}**		**βPGM_{wt}**		**βPGM_{D10N}**	
GSA of phosphorylation of enzyme	PSP_{D11N}-PLS	√	PSP_{D13N}-PLS	?	βPGM_{wt}-G6P-BeF_3	×	βPGM_{D10N}-βG16BP	√
TSA of phosphorylation of enzyme	PSP_{wt}-MgF_3-Ser	√	PSP_{D13N}-MgF_3-Ser	×	βPGM_{wt}-MgF_3-G6P	√	βPGM_{D10N}-MgF_3-G6P	×
					βPGM_{wt}-MgF_3-βG1P	c	βPGM_{D10N}-MgF_3-βG1P	×
	PSP_{wt}-AlF_4-Ser	√	PSP_{D13N}-AlF_4-Ser	×	βPGM_{wt}-AlF_4-G6P	√	βPGM_{D10N}-AlF_4-G6P	√
					βPGM_{wt}-AlF_4-βG1P	√	βPGM_{D10N}-AlF_4-βG1P	√
PSA of phosphorylation of enzyme	PSP_{wt}-BeF_3-Ser	×	PSP_{D13N}-BeF_3-Ser	×	βPGM_{wt}-BeF_3-G6P	√	βPGM_{D10N}-BeF_3-G6P	√
					βPGM_{wt}-BeF_3-βG1P	√	βPGM_{D10N}-BeF_3-βG1P	n/
GSA of dephosphorylation of enzyme[p]	PSP_{wt}-BeF_3-H_2O	√	PSP_{D13N}-BeF_3-H_2O	√	βPGM_{wt}-BeF_3-H_2O	√	βPGM_{D10N}-BeF_3-H_2O	√
TSA of dephosphorylation of enzyme[p]	PSP_{wt}-MgF_3-H_2O	√	PSP_{D13N}-MgF_3-H_2O	√	βPGM_{wt}-MgF_3-H_2O	×	βPGM_{D10N}-MgF_3-H_2O	×
	PSP_{wt}-AlF_4-H_2O	√	PSP_{D13N}-AlF_4-H_2O	√	βPGM_{wt}-AlF_4-H_2O	×	βPGM_{D10N}-AlF_4-H_2O	×

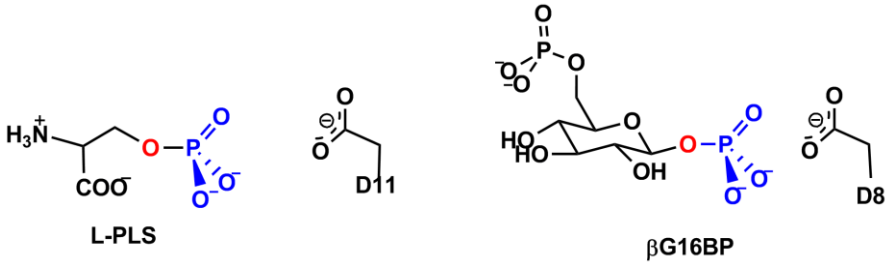
GSA: Ground state analogue. TSA: Transition state analogue. PSA: Product state analogue

c: contamination in the sample so it is uncertain if the complex can form.

?: the observation of a binding peak in ^{31}P NMR is questionable.

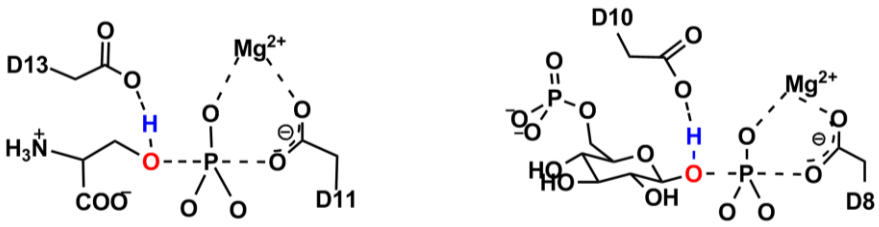
n/a: not attempted

4.4.1 Ground state analogue complexes for enzyme phosphorylation



Both GS for PSP and βPGM phosphorylation are concerned with transferring a phosphate from the oxygen of a substrate hydroxyl group onto a nucleophilic aspartate of the enzyme. As far as we know, tetrahedral BeF_3^- complex could only attach on aspartate, on glutamate or, rarely, on an imidazole ring. So if fluoroberyllate acts as a mimic of phosphate in the substrate, it always ends up staying on the anionic aspartate, with a low pK_a, but not on the hydroxyl of serine or G6P. Therefore, the GS for enzyme phosphorylation can only be simulated in the mutant with the original substrate bound but with the disadvantage of distorted Asn 11 in PSP_{D11N}-PLS complex or with Asn 10 in βPGM_{D10N}-βG16BP (see **Table 4.6**, row 2).

4.4.2 Transition state analogue complexes for enzyme phosphorylation

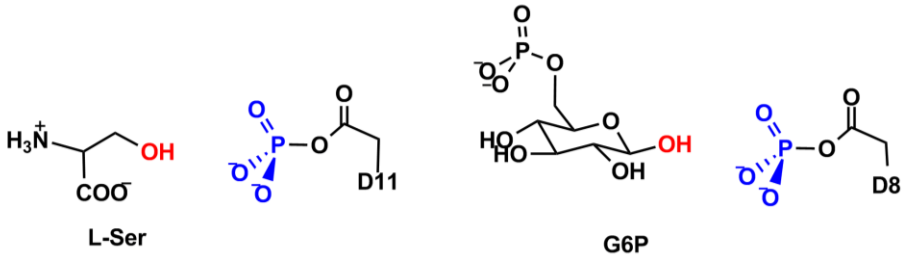


It is obvious that both wild type enzymes can stabilise the trifluoromagnesate and tetrafluoroaluminate TSAs for enzyme phosphorylation to give PSP_{wt}-MF_x^--Ser and βPGM_{wt}-MF_x^--G6P complexes (**Table 4.6**, columns 2 and 4). But different from βPGM_{D10N}, PSP_{D13N} fails to form any TSA complex (see **Section 4.3.2.1**) because the binding of L-serine is unfavourable.

The metal fluoride moiety must mimic the phosphate in the PLS TS and act as an anchor for substrate binding. But the side chain oxygen of L-serine cannot form a strong interaction with the central metal in the D13N mutant as it is not hydrogen bonded to a general base to make it more electronegative. Yet βPGM_{D10N} successfully forms a tetrafluoroaluminate complex with G6P and βG1P and this shows the substrate

phosphate is very beneficial for binding in the passive phosphate site for these two reactants (see **Section 3.3.4**). This can be contrasted with an experiment in which glucose and metal fluoride were used with wild type βPGM to try to form a TSA for the hydrolysis step of G6P. The use of glucose as a ligand failed to demonstrate formation of a TSA complex, which shows how important the passive phosphate is as a binding and orientation facility for the substrate hexose phosphate (Griffin, PhD thesis, 2011).

4.4.3 Product state analogue complexes for enzyme phosphorylation



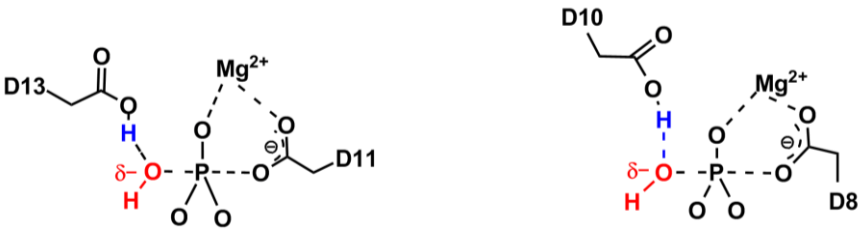
The failure to observe L-serine binding in a product state analogue (PSA) for PSP and its mutant (**Table 4.6**, column 2) once again suggests that interactions between PSP and L-serine are weak to favour product release (see **Section 4.3.2.1**). By contrast, in both βPGM_{wt} and βPGM_{D10N}, this limitation is overcome and reflected by the formation of BeF_3^--G6P PSA complex (**Table 4.6**, column 4, also see **Section 3.3.6**) because of the strong interaction of the phosphate of G6P with the arginine clamp in the passive phosphate binding site. In this step, the protein adapts to an open conformation and there is no participation of Asp 10 as catalytic base, so the βPGM_{D10N}-BeF_3^--G6P PSA complex still readily forms and shows similar ^{19}F chemical shifts.

4.4.4 Ground state analogue complexes for enzyme dephosphorylation

After the first product has been released, for the second phosphoryl transfer step both of these wild type enzymes stabilise GSA complexes, PSP_{wt}-BeF_3^--H_2O and βPGM_{wt}-BeF_3^--H_2O, but with different architectures (**Table 4.6**, columns 2 and 4, also see **Section 3.3.6**). In PSP_{wt}-BeF_3^--H_2O, Asp 13 is positively employed to interact and activate the attacking water in an in-line position. But in βPGM_{wt}-BeF_3^--H_2O the water

molecule is passively placed in a non-in-line position by one of the fluorines in the BeF_3^- moiety through an ionic bond (seen in PDB: 2WFA). The difference in the strategic positions for the two waters reveals the different efficiency of the enzyme dephosphorylation step in these two enzymes. It shows that PSP is geared towards hydrolysis to complete the turnover cycle while βPGM is protective in the active site for hexose phosphate. The fact that PSP_{D13N}-BeF_3^--H_2O and $βPGM_{D10N}$-BeF_3^--H_2O complexes both exist reflects that Asp 13 and Asp 10 as GABCs participate less in the GS than they do in TS.

4.4.5 Transition state analogue complexes for enzyme dephosphorylation



The acceptance of water as an intended nucleophile is enhanced in the second TS for PSP as shown by the formation of a PSP_{wt}-MF_x^--H_2O TSA complex, while there is no corresponding $βPGM_{wt}$-MF_x^--H_2O TSA complex (**Table 4.6**, columns 2 and 4). The existence of PSP_{D13N}-MF_x^--H_2O for the D13N mutant shows that presence of a GB to activate an H_2O nucleophile is not demanded as strongly as in the first TS for dphosphorylation of L-serine (see **Section 4.3.2.2** and **4.3.2.3**). By contrast for βPGM, even though H_2O is much more abundant than hexose phosphate, because there is no other function in the active site of βPGM when Asp 10 is in the 'out' conformation for such water stabilisation, a TSA complex for βPGM enzyme dephosphorylation is not stable and does not form (**Table 4.6**, columns 2 and 4).

4.4.6 What we have learned from metal fluoride studies on PSP_{D13N} and $βPGM_{D10N}$

All the evidence obtained here shows that phosphoryl transfer requires apical oxygens that either are anionic or are hydrogen bonded to a catalytic base. This is accurately reflected in the stability of formation of the corresponding trifluoromagnesate complexes. The failure to make TSA complexes PSP_{D13N}-MgF_3^--Ser and $βPGM_{D10N}$-MgF_3^--G6P as well as $βPGM_{D10N}$-MgF_3^--βG1P supports this (**Table 4.6**, columns 3 and 5). By contrast, fluoroaluminate appears to make complexes that are more stable and so

can overcome the elevated TS energy for the asparagines mutant to form TSA complexes βPGM$_{D10N}$-AlF$_4^-$-G6P and βPGM$_{D10N}$-AlF$_4^-$-βG1P (**Table 4.6**, column 5, and also see **Section 3.3.4**, **Section 3.3.5.1**, and **Section 4.3.2.1**). But this cannot be achieved in the case of PSP$_{D13N}$ to give a Ser-AlF$_4^-$ TSA complex because ligand binding also plays a decisive role on the success for complex formation and the product exhibits the weakest binding in the turnover cycle of this enzyme (see **Section 4.3.2.1**). Fluoroberyllate forms either a GSA or a PSA according to the relative nucleophilic or anionic character of Asp and the substrate OH group (see **Section 3.3.6** and **Section 4.3.3)**. This suggests that GABC is not as involved in the GS for phosphoryl transfer catalysis as it is in the TS. In a word, the D13N mutant of PSP and D10N mutant of βPGM establish the role of Asp 13 and Asp 10 as GABCs that are essential for phosphoryl transfer in both enzymes.

Chapter 5. cAMP-Dependent Protein Kinase (cAPK)

5.1 Introduction

Protein kinases catalyse the transfer of the γ-phosphate of ATP to the hydroxyl group of a serine, threonine, or tyrosine residue in the target protein. The substrate specificity for phosphorylation of these residues is vital to the accuracy of cellular function and protein kinases are critically responsible for accurate regulation of many cellular functions. Kinase activity in the wrong place or at the wrong time can have disastrous consequences, often leading to cell transformation and cancer.(Huse *et al.*, 2002)

The cAMP-dependent protein kinase (cAPK; also known as protein kinase A, PKA) is a protein kinase primarily regulated by the second messenger cAMP in mammalian cells. Its substrate sequence is RRX**S/T***Z*, where *Z* is usually a hydrophobic residue. cAPK was the first protein kinase structure to be solved (Knighton *et al.*, 1991) and, in ongoing studies, Susan Taylor and her co-workers have provided fascinating detail of its composition, subunit interactions, and dynamics. Along with analysis of its structure-function relationships and kinetic properties, cAPK has been the most studied and best understood of 2000 eukaryotic protein kinases and has become the structural paradigm for all protein kinases, providing valuable mechanistic information. This includes in-line transfer of a γ-phosphoryl group from ATP to an acceptor serine residue and GABC from Asp 166 in a mainly associative process (**Figure 5.1**).(Herberg *et al.*, 1993; Madhusudan *et al.*, 1994; Madhusudan *et al.*, 2002; Akamine *et al.*, 2003; Yang *et al.*, 2004; Das *et al.*, 2007; Masterson *et al.*, 2008; Kennedy *et al.*, 2009; Yang *et al.*, 2009; Masterson *et al.*, 2010; Taylor *et al.*, 2011)

In cAPK, the peptide substrate SP20 (TTYADFIASGRTGRRA**S**IHD) binds in an extended conformation across the front face of the nucleotide binding pocket of cAPK, close to the γ-phosphate of ATP and with Asp 166 hydrogen bonding to the target serine hydroxyl. The absolutely conserved residues Asn 171 and Asp 184 coordinate the two magnesium cations involved in nucleotide recognition. As a result, ATP is oriented for in-line attack by the substrate serine hydroxyl (**Figure 5.2**). Water is excluded by hydrophobic residues that cluster around the substrate phosphorylation site to ensure that the γ-phosphate group is not subject to competitive hydrolysis.(Madhusudan *et al.*, 2002; Yang *et al.*, 2004)

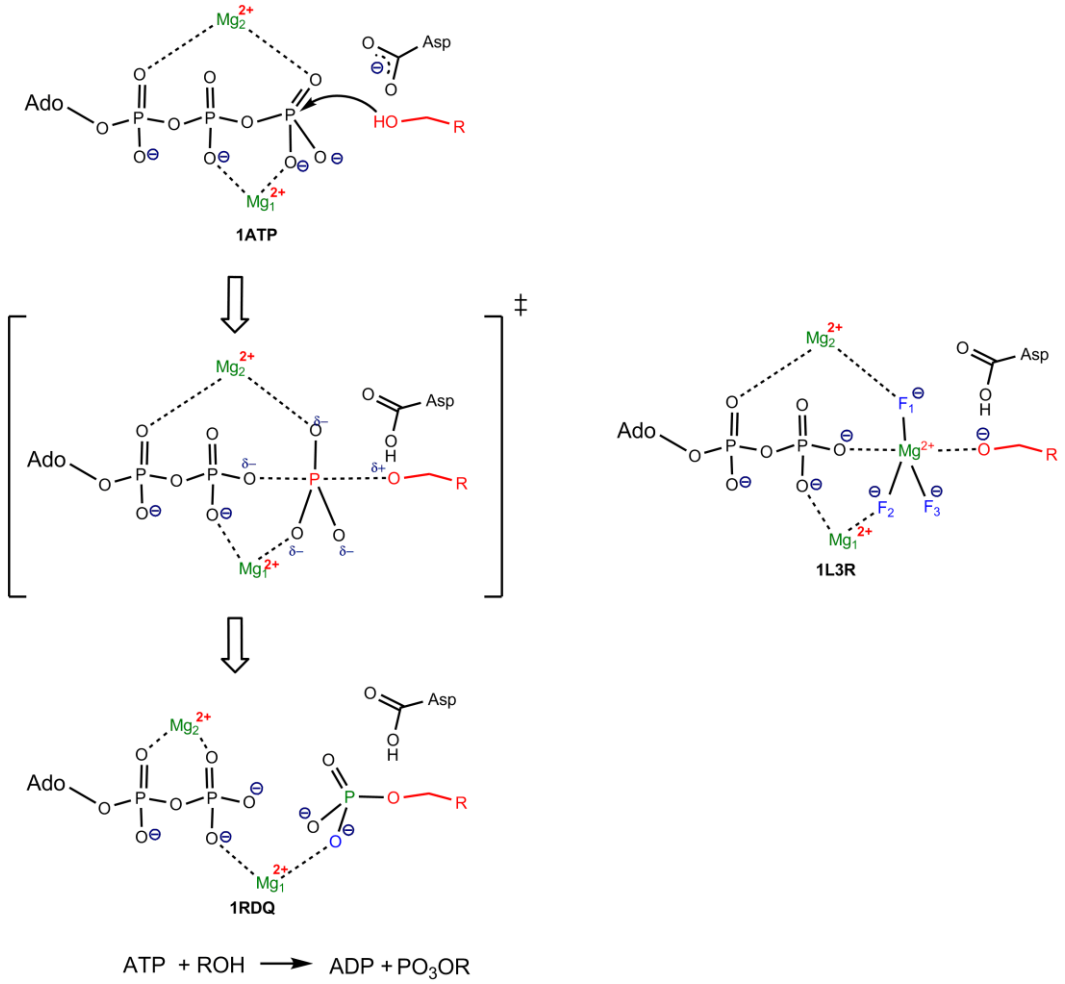


Figure 5.1 Catalytic reaction for phosphorylation of serine by ATP in cAPK. The three schemes are derived from three structures of cAPK PDB: 1ATP, 1L3R and 1RDQ.

Kinases operating on small substrates, *e.g.* PGK, UMP/CMP kinase and G proteins, generally use a single magnesium ion which binds to the β- and γ-phosphate non-bridging oxygens. They also employ a cationic lysine or arginine to bridge α- and γ-oxygens in the nucleotide triphosphate substrate. Protein kinases, on the other hand, employ two catalytic magnesium ions. Mg^1 coordinates the α- and γ-phosphate non-bridging oxygens as well as to the β,γ-bridging oxygen while Mg^2 coordinates the non-bridging β- and γ-phosphate oxygens. These two magnesiums provide a stronger scaffold for locating the adenine nucleotide that result in a smaller conformational fluctuation when the protein kinase releases its product than is the case for single-magnesium kinases.(Cheng *et al.*, 2005)

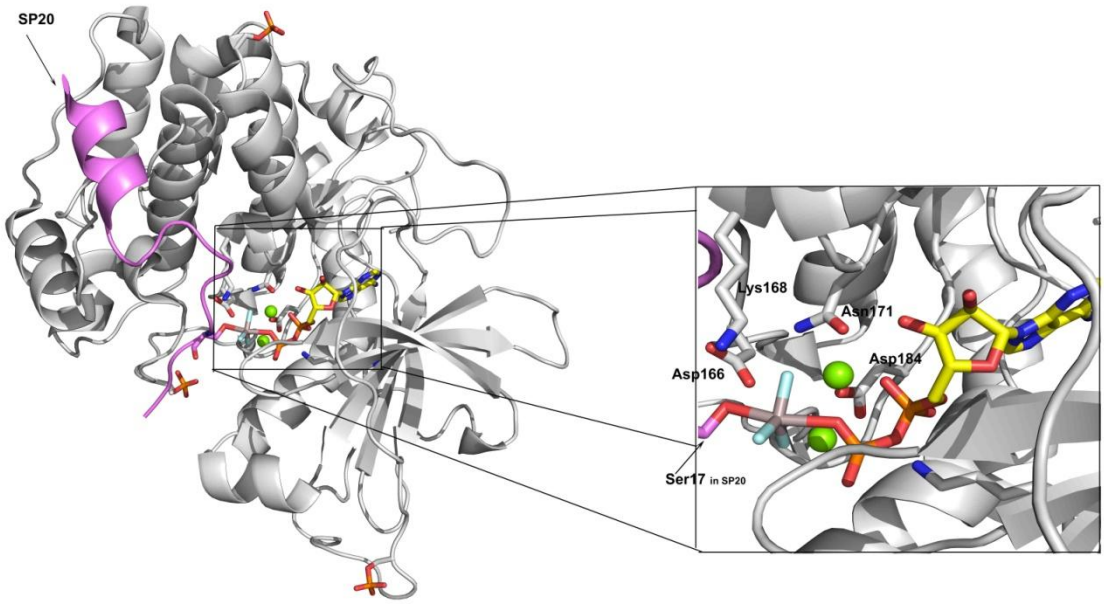


Figure 5.2 A cartoon TSA structure of cAPK (grey) with SP20 (magenta), AlF_3 (grey and cyan sticks), and ADP (yellow sticks) bound.(PDB: 1L3R)

However, the formation of an AlF_3^0 TSA complex by the catalytic subunit of cAPK with ADP and a serine-substrate peptide (SP20) (PDB: 1L3R) constitutes an clear exception to the Charge Balance Hypothesis (CBH), which has been developed by analysis of many single-Mg^{2+} enzymes examined by PIXE and ^{19}F NMR as described in **Section 1.9**. That is because this AlF_3^0 TSA complex has one less negative charge than the transferring PO_3^- group.(Graham *et al.*, 2002; Cliff *et al.*, 2010) By contrast, an octahedral AlF_4^- TSA complex for another major signalling tyrosine/threonine protein kinase, MEK6, with ADP and its substrate protein, p38, has recently been analysed by the Sheffield group using ^{19}F NMR.(Xiaoxia *et al.*, 2011) It follows that MEK6 does indeed obey the CBH and detailed analysis shows that MEK6 behaves as do other metabolic, small molecule phosphoryl transfer enzymes with respect to transition state stabilisation. However, because MEK6 is a dual specificity protein kinase, phosphorylating both Thr and Tyr hydroxyl groups, and in view of the large number of protein kinases and their diversity of their signalling pathways, it cannot be assumed that they all employ the same chemistry. It is therefore timely and significant to determine whether cAPK is an exception to the CBH concept and so constitutes a different kinase catalytic mechanism from that which we have analysed for MEK6.

The need to resolve this anomaly has become more acute as a result of the recent publication of an analysis of the tbp TSA complex for the protein kinase, CDK2. This

has been assigned (Bao *et al.*, 2011) as containing MgF_3^-. CDK2, like MEK6, is a two-magnesium kinase that phosphorylates proximate Thr and Tyr hydroxyls and hence is not a direct model for cAPK. It is here established by ^{19}F NMR that cAPK follows the precedence of MEK6 in forming an AlF_4^- TSA complex in solution while it also forms an MgF_3^- TSA complex as does CDK2, thereby establishing that cAPK conforms to the CBH. Re-analysis of the deposited coordinates for the cAPK-ADP-AlF_3^0-SP20 crystal structure (PDB: 1L3R) not only shows the tbp moiety to have only 70% of the occupancy with the other 30% being an octahedral complex, but it also confirms our ^{19}F NMR analysis of the priority of charge over geometry for TSA. This is also another case of a mixed occupancy metal fluoride moiety having been misassigned as a tbp AlF_3^0 at high pH.

We also establish that cAPK shows unexpected behaviour. In MEK6 and in the numerous metabolic phosphoryl transfer enzymes studied to date tetrafluoroaluminates predominate very strongly over trifluoromagnesates. But in cAPK the stronger competitiveness of MgF_3^- with AlF_4^- than other proteins from the results described in this thesis raises an inhibitory significance of MgF_3^- in physiology.

5.2 Materials and Experiments

5.2.1 Protein expression and purification

The expression construct for the catalytic subunit of cAPK from mouse was kindly provided by Dr Peter Rellos, Structural Genomics Consortium, Oxford, UK. The protein, having an uncleavable N-terminal His_6 tag encoded by pACYCDuet-1 vector, was expressed in LB medium at 18°C overnight (or 25°C for 6 h) in *E. coli* strain BL21-GOLD (DE3) with 1 mM IPTG. Cells were lysed by sonication in lysis buffer (50 mM HEPES pH 7.5, 300 mM NaCl, 20 mM imidazole) and the protein in the clarified supernatant was bound onto an Ni-NTA Sepharose affinity column (10 mL, Qiagen) followed by washing with lysis buffer supplemented with 1 M NaCl. Protein was eluted with lysis buffer supplemented with 150 mM imidazole. A final purification step involved size-exclusion chromatography using Superdex^{TM} G75 Gel Filtration column (GE Healthcare) eluted with a buffer containing 50 mM HEPES pH 7.5, 300 mM NaCl and 5 mM DTT. ^1H NMR showed that the protein was folded and ^{31}P NMR showed that the protein was phosphorylated at four sites, corresponding to the isoforms reported previously.(Yonemoto *et al.*, 1997; Madhusudan *et al.*, 2002)

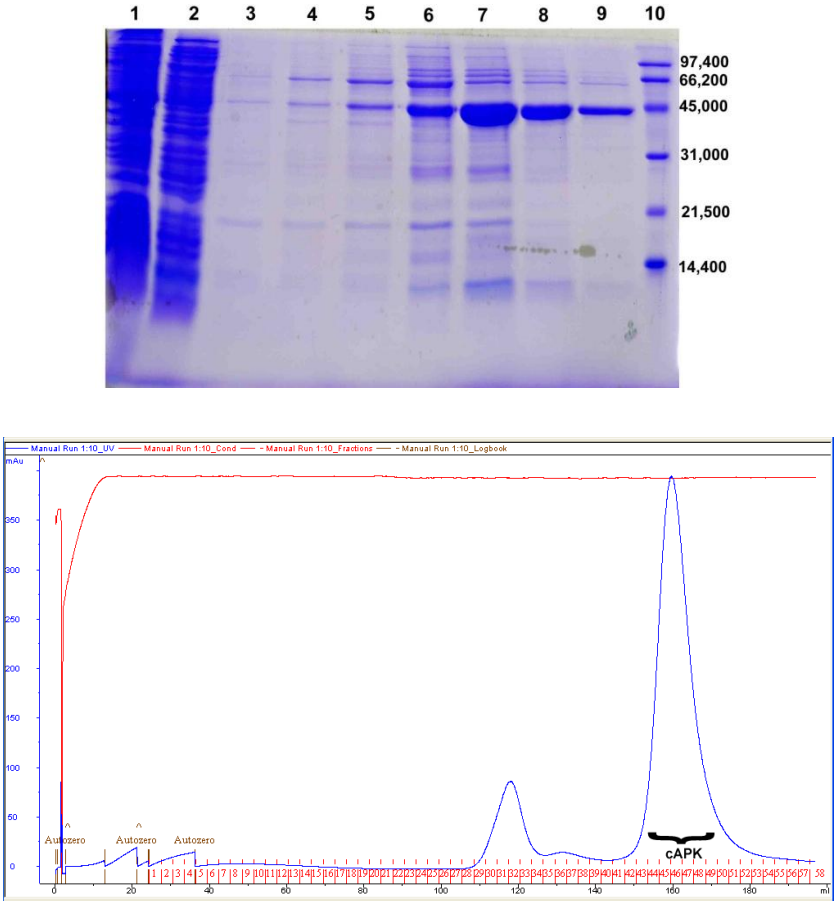


Figure 5.3 Elution of cAPK during Ni-NTA affinity chromatography shown on SDS-PAGE. lane 1 − lysate, lane 2 - flow through after Ni-NTA column, lanes 3~9 − fractions of the elution, lane 10 − low range marker. UV trace is for gel filtration chromatography in the second step.

5.2.2 Preparation of NMR samples of cAPK-ADP-MgF_3^--SP20 and cAPK-ADP-AlF_4^--SP20 TSA complexes

Synthetic substrate peptide SP20 was provided by PeptideSynthetics as its trifluoroacetate (TFA) salt. The sample of the cAPK-ADP-MgF_3^--SP20 TSA complex contained 600 μM His_6-cAPK, 300 mM NaCl, 5 mM DTT, 20 mM $MgCl_2$, 40 mM NH_4F, 5 mM ADP, and 1 mM SP20 in 50 mM HEPES buffer (pH 7.2). A 5 mm Teflon liner was inserted into the NMR tube to prevent fluoride-mediated leaching of aluminium from the glass and 10% v/v D_2O was employed for the deuterium lock. The sample of the cAPK-ADP-AlF_4^--SP20 TSA complex contained all the same ingredients as above supplemented with $AlCl_3$ (up to 3.4 mM) at pH 7.4.

5.2.3 NMR methods

For the 180^o inversion experiment, the delay for exchange was set at 2.5 ms. Integration of peaks was performed using FELIX (Felix NMR Inc, San Diego, CA),

with phase adjustment and baseline correction optimized for the chosen peaks. ^{31}P NMR spectra were accumulations of 128k scans at 298 K on a Bruker Avance 500 MHz spectrometer equipped with a 5 mm multinuclear probe tuned to ^{31}P (202.5 MHz).

5.3 Results and Discussion

5.3.1 Re-analysis of the RCSB PDB deposited coordinates for the cAPK-ADP-AlF_3^0-SP20 crystal structure [PDB: 1L3R]

A reanalysis of structure PDB:1L3R was made by Dr. Matthew Bowler from the Macromolecular Crystallography Group, ESRF, Grenoble, after all the NMR work has been completed and is discussed later in this chapter.(**Appendix II**) The reason it is described here initially is to make a clear presentation of the two complexes cAPK-ADP-MgF_3^--SP20 and cAPK-ADP-AlF_4^--SP20 TSA, to which the NMR analysis will refer later.

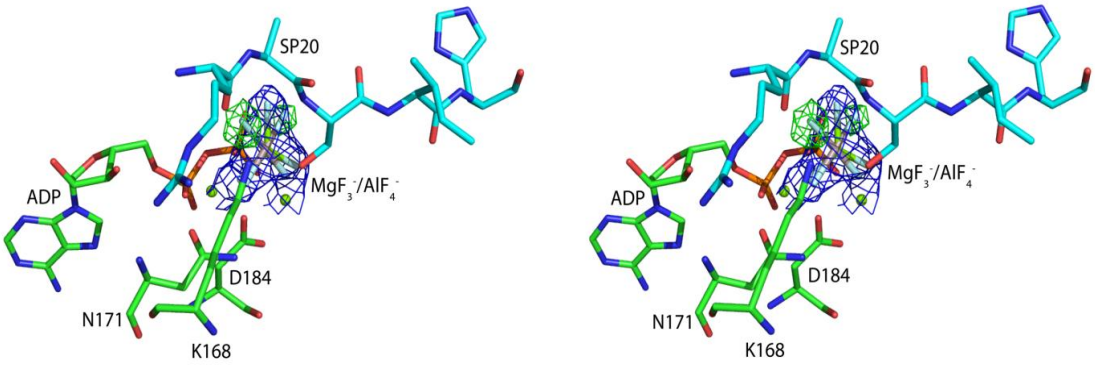


Figure 5.4 Density for metal fluoride complexes in the active site of PDB: 1L3R, the view is in stereo. The $2F_o$-F_c electron density (blue mesh, contoured at 1.5σ) and difference electron density (F_o-F_c, shown as a green mesh contoured at 3σ) is shown after the unrestrained refinement of an MgF_3^- species instead of AlF_3^0. Two difference map peaks (heights 4.5σ) are visible for the fluorine atoms of an octahedral AlF_4^- species before its inclusion in the model. The refined atomic positions of the AlF_4^- species are also shown.

The resulting electron density maps are shown in **Figure 5.4**. By comparing atomic *B*-factor values for the metal fluoride moieties with the bound nucleotide, the occupancy of the active site was judged to contain 70% tbp MgF_3^- to 30% octahedral AlF_4^- (**Figure 5.5**)

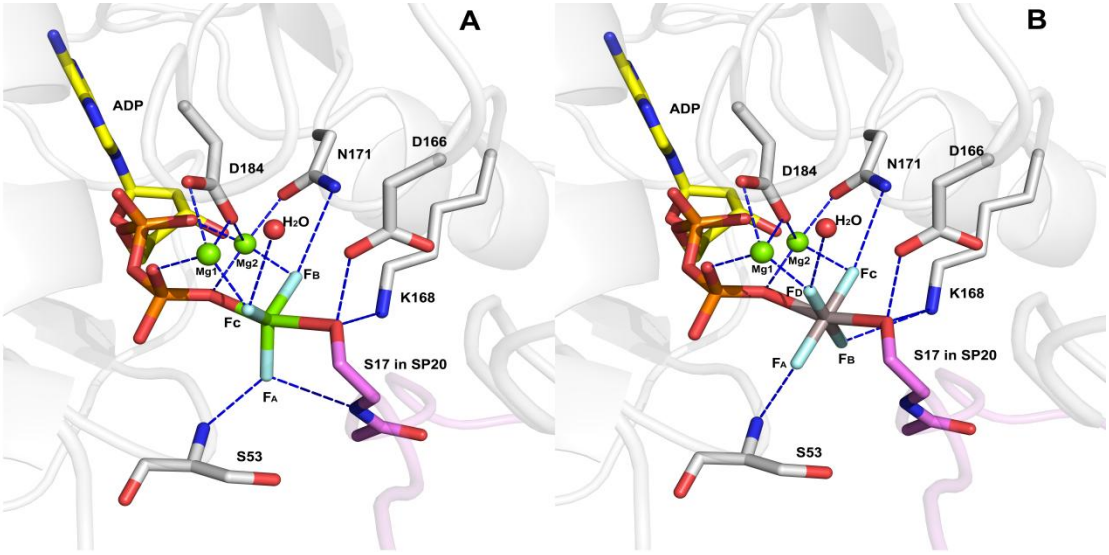


Figure 5.5 Structure of the cAPK-ADP-MgF_3^--SP20 and cAPK-ADP-AlF_4^--SP20 TSA complexes derived from a reanalysis of the crystal structure (PDB: 1L3R). (**A**) The AlF_3^0 moiety re-assigned as MgF_3^- with atom numbering F_1, F_2 and F_3 in 1L3R corresponding to F_B, F_C and F_A, respectively in the ^{19}F NMR spectrum of the cAPK-ADP-MgF_3^--SP20 TSA complex. (**B**) The AlF_4^- moiety assigned on the basis of additional positive density in the active site in 1L3R. The density corresponds to an octahedral complex at 30% occupancy derived by atomic *B*-factor analysis. The fluorine atoms are labelled on the basis of ^{19}F NMR chemical shifts and SIIS data.

5.3.2 cAPK-ADP-MgF_3^--SP20 and cAPK-ADP-AlF_4^--SP20 TSA complexes

5.3.2.1 Chemical shift data

cAPK readily forms a cAPK-ADP-MgF_3^--SP20 TSA complex in the presence of magnesium, fluoride, ADP, and the substrate peptide SP20. The ^{19}F NMR spectrum of the complex shows two well-resolved resonances of equal intensity (F_B = −185 and F_C = –191 ppm, **Figure 5.6a**), clearly upfield of the peak for free fluoride (–119 ppm) and these are absent in the spectra of buffer controls. A third peak, F_A (at –156.5 ppm) is obscured by the broad peak of free MgF^+ (**Figure 5.6a**) but could be observed by two methods. The first, shown in **Figure 5.6c**, is through summation of a standard spectrum (**Figure 5.6a**) and one in which the free F^- signal at –119 ppm has been inverted (**Figure 5.6b**), which results in inversion of the resonance of free MgF^+ species through fast exchange.

The second method is to subtract the spectrum of an SP20-free sample from the equivalent spectrum of the corresponding TSA complex. This also reveals the F_A peak at the same frequency (data not shown). The base line and the peaks intensities of F_A are slightly different from these two processing methods. Presaturation cannot be used

in this case because the life-time of the MgF_3^- complex is shorter than the 1 s saturation time.

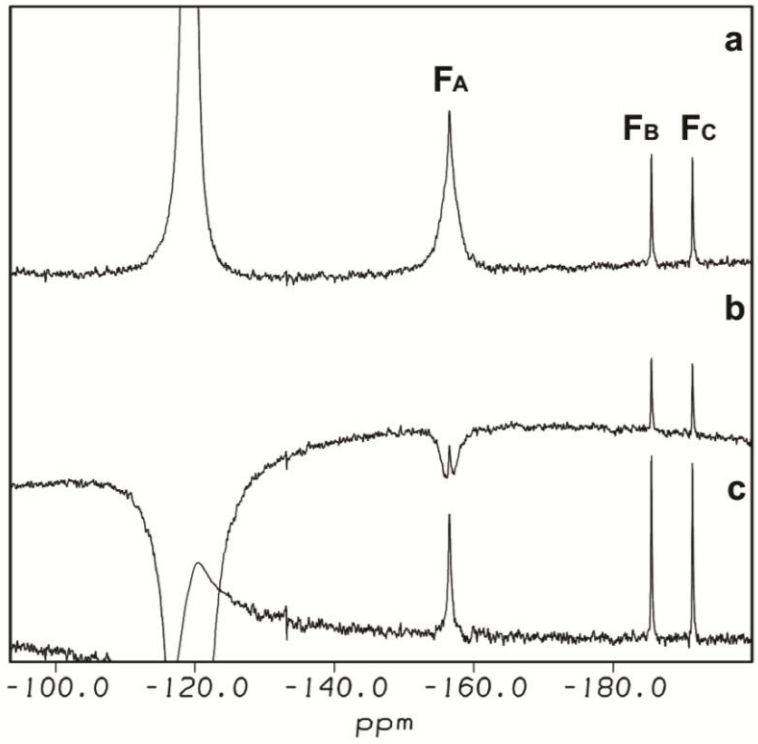


Figure 5.6 ^{19}F NMR spectra of the cAPK-ADP-MgF_3^--SP20 TSA complex recorded in 50 mM HEPES pH 7.5 containing 300 mM NaCl, 5 mM DTT, 5 mM ADP, 20 mM $MgCl_2$, 40 mM NH_4F, 220 μM cAPK, and 1 mM SP20. Spectrum (**c**) was generated by the summation of (**a**) the spectrum recorded under standard conditions and (**b**) an equivalent spectrum recorded under identical conditions but after 180° pulse inversion of peaks for free fluoride (–119 ppm) and MgF^+ (–156 ppm). F_A = –157, F_B = –185, and F_C = –191 ppm.

The ^{19}F chemical shifts of this MgF_3^- TSA complex are on average some 15 ppm upfield relative to those observed for MgF_3^- species in single-magnesium phosphoryl transfer enzymes, *e.g.* hPGK, PSP, and β-PGM.(Baxter *et al.*, 2006; Baxter *et al.*, 2009; Baxter *et al.*, 2010; Cliff *et al.*, 2010) This shows that the fluorines in the cAPK MgF_3^- TSA complex are relatively less coordinated by hydrogen bonding and more shielded through increased coordination to the two catalytic magnesiums in the active site.

In the search for a cAPK tetrafluoroaluminate TSA complex by ^{19}F NMR, we first employed solution conditions that had been successful for PSP, PGM, MEK6/p38, and UMP/CMP kinase.(Baxter *et al.*, 2008; Cliff *et al.*, 2010; Xiaoxia *et al.*, 2011) We then used the conditions employed for crystallisation of an aluminium fluoride TSA complex of cAPK.(Madhusudan *et al.*, 2002) Neither of these gave any evidence of an aluminium fluoride TSA complex. We were eventually successful only on increasing

fluoride to 40 mM and aluminium above 3.4 mM (**Figure 5.7c**). Four peaks corresponding to a cAPK-ADP-AlF_4^--SP20 TSA complex have equal intensity (F_A = –146, F_B= –148, F_C= –161, F_D= –168 ppm). In common with other AlF_4^- TSA complexes, the four resolved resonances indicate the fluorides are each bound in a distinct site and do not exchange environments on the NMR timescale. When aluminium was excluded from the NMR sample by the use of deferroxamine, the ^{19}F NMR spectrum reverted to the three resonances typical of MgF_3^- complexes. Therefore we are confident that these four peaks are indeed from the AlF_4^- TSA complex. To form a fully occupied AlF_4^- TSA complex, we used concentrations of both aluminium and fluoride that are significantly higher than the concentrations employed for tetrafluoroaluminate complexes of PGK *etc*. (10 mM fluoride and 2 mM aluminium). At pH 7.4, 20 mM Mg^{2+} and 1.8 mM Al^{3+} seven fluoride signals were observed, in approximately equal strengths, corresponding to equal populations of MgF_3^- and AlF_4^- TSAs (**Figure 5.7b**). This shows the binding affinity of MgF_3^- moiety is only about ten times lower than AlF_4^-. This special feature will be discussed in details later in **Section 5.3.3**.

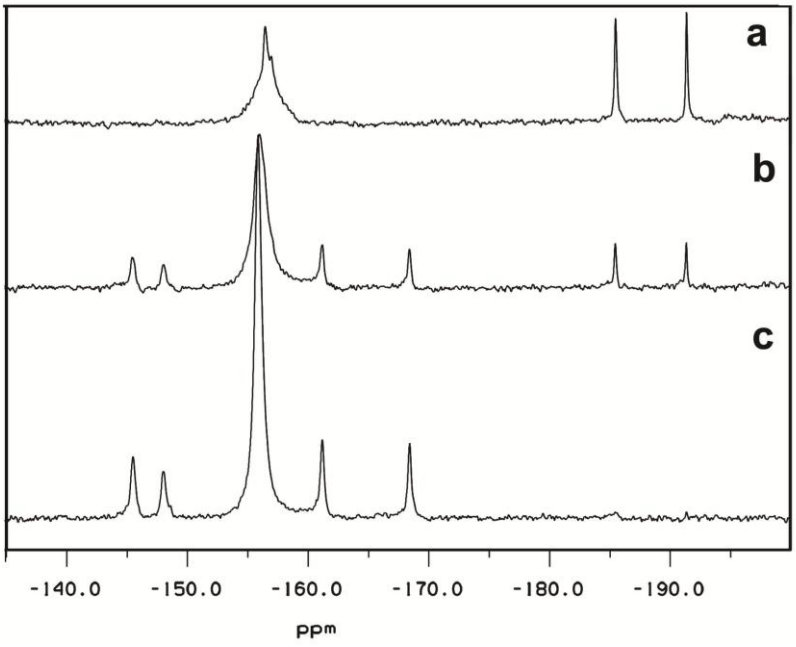


Figure 5.7 ^{19}F NMR spectra of: (**a**) cAPK-ADP-MgF_3^--SP20 TSA complex recorded in 50 mM HEPES buffer at pH 7.5 containing 300 mM NaCl, 5 mM DTT, 5 mM ADP, 20 mM $MgCl_2$, 40 mM NH_4F, 600 μM cAPK and 1 mM SP20 at 25°C; (**b**) the addition of 1.8 mM $AlCl_3$ to the cAPK-ADP-MgF_3^--SP20 TSA complex; (**c**) the addition of 3.4 mM $AlCl_3$ to the cAPK-ADP-MgF_3^--SP20 TSA complex.

5.3.2.2 SIIS data

In the cAPK-ADP-MgF_3^--SP20 TSA complex, F_A (F_3 in PDB: 1L3R) has the largest SIIS (1.4 ppm) and correlates with a 2.8 Å hydrogen bond to the Ser 17 N–H in SP20 and a 2.7 Å hydrogen bond to Ser 53 N–H of cAPK (**Figure 5.5a**). The smaller SIIS values for F_B (F_1 in 1L3R, 0.3 ppm) and F_C (F_2 in 1L3R, 0.2 ppm) result from their relatively low surrounding proton densities because of their primary coordination to Mg^2 (2.0 Å) and Mg^1 (2.0 Å) respectively (**Figure 5.5a**). Fluoride F_1 is likely to correspond to resonance F_B because of the additional small, downfield chemical shift difference with respect to F_2 and slightly enhanced SIIS induced by a long hydrogen bond (3.1 Å) to Asn 171 (**Figure 5.5a**).

In the cAPK-ADP-AlF_4^--SP20 TSA complex, F_A has the biggest SIIS (0.8 ppm) correlating with a strong hydrogen bond (2.7 Å) to Ser 53 N–H of cAPK It is nonetheless larger than that for F_B (0.4 ppm) from its weak hydrogen bonding interaction (3.0 Å) with Lys 168 side chain. F_D (0.6 ppm) forms hydrogen bonds with Mg^1 and a water molecule (2.7 Å), and F_C (0.1 ppm) having coordination with Mg^2 and a weak hydrogen bond (3.0 Å) to side chain amide of Asn 171 (**Figure 5.5b)**. Given the low occupancy (30%) and the low resolution (2.0 Å) of this AlF_4^- TSA complex, the interpretation of the SIIS according to the corresponding hydrogen bond lengths and directions may not be unique, particularly with respect to F_B and F_C.

Table 5.1 ^{19}F chemical shifts (ppm) and SIIS (ppm) of the cAPK-ADP-MgF_3^--SP20 and cAPK-ADP-AlF_4^--SP20 TSA complexes. The corresponding data for the MEK6-ADP-AlF_4^--p38α TSA complex are provided for comparison.

Complex	F_A	F_B	F_C	F_D
cAPK-ADP-MgF_3^--SP20	–157*	–185	–191	–
SIIS /ppm	1.4	0.3	0.2	–
cAPK-ADP-AlF_4^--SP20	–146	–148	–161	–168
SIIS /ppm	0.8	0.4	0.1	0.6
MEK6-ADP-AlF_4^--p38α	–142	–148	–154	–167

*Signals for MgF^+ ($MgF(H_2O)_5^+$) and F_A are coincident but are resolved by 180° inversion of the free F^- signal.

5.3.2.3 Line width analysis

In the cAPK-ADP-MgF_3^--SP20 TSA complex, F_B and F_C generally have narrower line widths and upfield chemical shifts than those of F_A because these two fluorides coordinate the catalytic Mg^{2+} ions and are less affected by relaxation from protons

coordinating to them. In the cAPK-ADP-AlF_4^--SP20 TSA, similar to that in MgF_3^- complex, the two downfield peaks F_C and F_D have a smaller line width for the same reason. But on average, the AlF_4^- complex has a bigger line width than MgF_3^- complex which indirectly reflects the difference between their binding affinities.

5.3.2.4 Comparison of the binding constant of Kemptide and SP20 for MgF_3^- TSA complexes

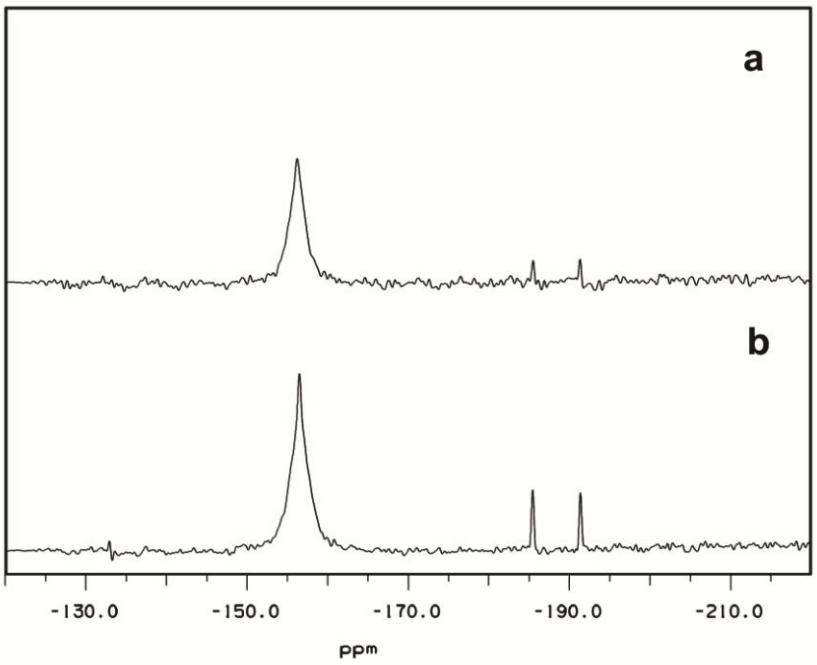


Figure 5.8 ^{19}F NMR spectra of: (**a**) cAPK-ADP-MgF_3^--Kemptide TSA complex recorded in 50 mM HEPES buffer at pH 7.5 containing 300 mM NaCl, 5 mM DTT, 5 mM ADP, 20 mM $MgCl_2$, 40 mM NH_4F, 200 μM cAPK and 1.5 mM Kemptide at 25°C; (**b**) cAPK-ADP-MgF_3^--SP20 TSA complex in the same buffer with 1.5 mM SP20 replacing Kemptide.

The literature binding constant K_d of the heptapeptide substrate Kemptide (LRRA**S**LG) is between 300 μM ~ 1 mM, depending on the order of addition of the two substrates (measured using AMPPNP as ATP analogue), and its K_M is 16 ~ 28 μM.(Knighton *et al.*, 1991; Madhusudan *et al.*, 1994; Masterson *et al.*, 2008) By comparison, the K_M of peptide substrate SP20 is less than 5 μM.(Madhusudan *et al.*, 1994) Our observations confirm the binding constants between the two peptides substrate are different in the formation of MgF_3^- complexes. The intensity of the peaks from an MgF_3^- complex were compared for the same concentrations of Kemptide and SP20 and in the same buffer conditions.

The ^{19}F NMR spectra show the two complexes have the same signals at –185 and –191 ppm (**Figure 5.8**). This indicates the MgF_3^- TSA complexes from Kemptide and SP20 have exactly the same conformation and chemical environment. The binding constant of the MgF_3^- TSA complex with SP20 is four-fold bigger than with Kemptide as gauged by the relatvie signal intensities.

5.3.3 MgF_3^- is more competitive with AlF_4^- for cAPK than for other proteins

A striking feature of the cAPK-ADP-SP20 metal-fluoride TSA complexes in solution is that the aluminium concentration has to exceed that of the protein. This scenario has not been observed with small molecule phosphoryl transferases or with the dual specificity protein kinases studied to date. In βPGM, for example, a PGM-MgF_3^--G6P TSA complex cannot be detected in the presence of Al^{3+} at stoichiometric levels with the protein, even when Mg^{2+} is in a molar excess of 500.(Baxter *et al.*, 2008) The formation of the AlF_4^- complex for MEK6/p38 (a two catalytic Mg^{2+} protein complex) only needs a concentration of 10 mM fluoride and stoichiometric aluminium (210 μM) at pH 7.5.(Xiaoxia *et al.*, 2011) In the aluminium titration experiment for cAPK (**Figure 5.9**), with conditions pH 7.4, 20 mM magnesium and 1.8 mM aluminium, there is an equal population of the trifluoromagnesate and tetrafluoroaluminate TSA complexes. This suggests that the binding constant of MgF_3^- TSA is only ~10 fold weaker than for the AlF_4^- TSA. Taking account the fact that AlF_4^- has very much greater intrinsic stability in solution than MgF_3^-, this observation shows that cAPK has a much higher affinity for the trifluoromagnesate TS complex than for the non-isosteric tetrafluoroaluminate complex and it seems likely that this feature arises from the combined coordination specificity of the two catalytic magnesium ions. From the crystal structure of the cAPK TSA complex, it is clear that the tbp metal trifluoride has the same regular geometry as the real TS, especially in the equatorial plane where the angle between F_C and F_B is close to 120° as a result of tight coordination of each fluorine to its respective catalytic magnesium. The switch to an octahedral AlF_4^- moiety in the same locus imposes steric demands on the organisation of the active site that are expected inevitably to weaken the binding of the AlF_4^- moiety relative to MgF_3^-. That observation is also supported by the much larger upfield shifts for F_C and F_D in the cAPK complex than those in the single magnesium kinase complexes.

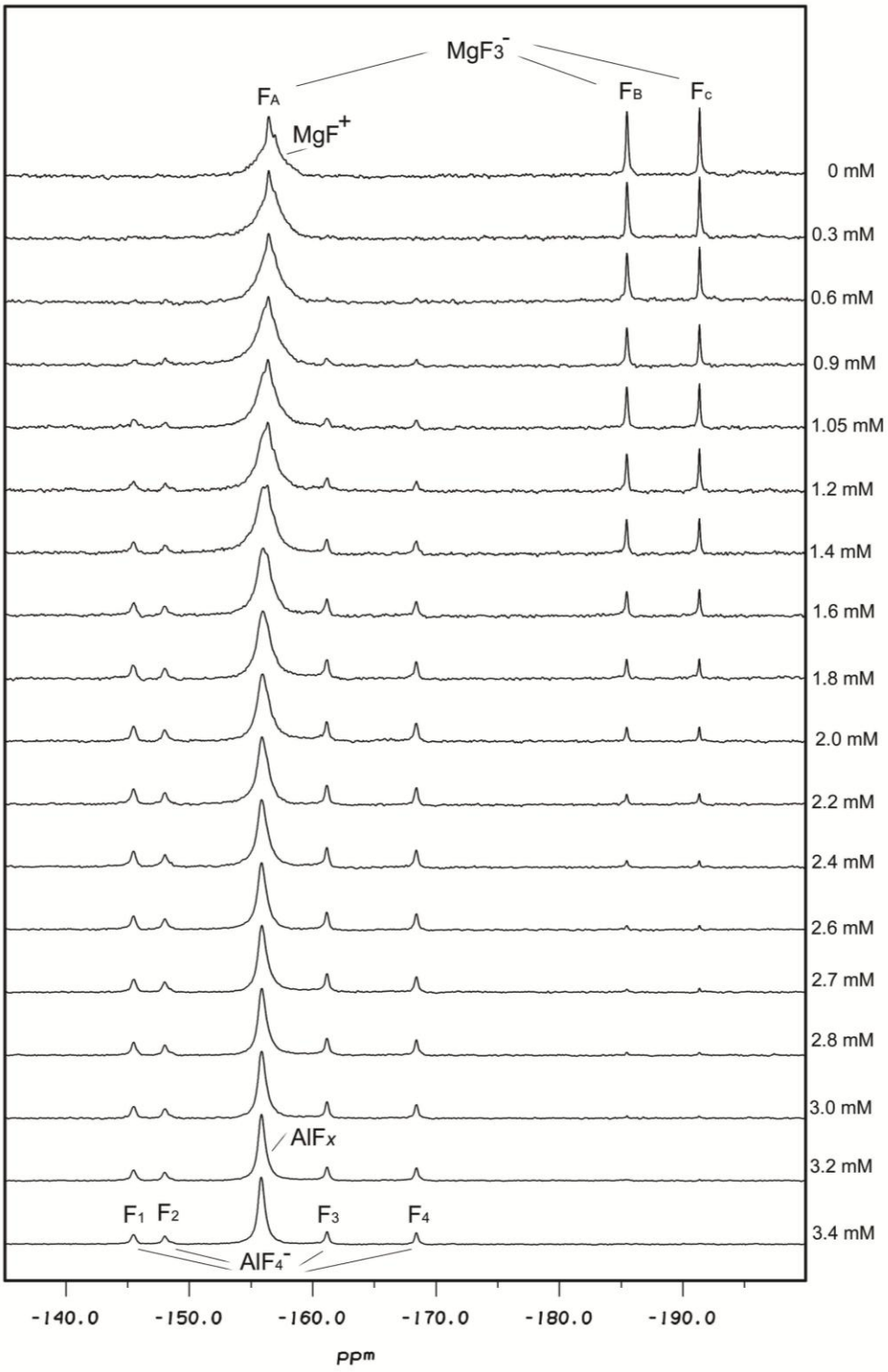


Figure 5.9 Aluminium titration of cAPK-ADP-MF-SP20 complexes, in 50 mM HEPES buffer at pH 7.5 containing 300 mM NaCl, 5 mM DTT, 5 mM ADP, 20 mM $MgCl_2$, 40 mM NH_4F, 600 μM cAPK and 1.5 mM SP20 at 25°C.

Based on the aluminium titration data, the binding constant of MgF_3^- and AlF_4^- have been calculated by Dr. Matthew Cliff (Sheffield) using peak intensities measured after manual baseline correction for each peak. The simultaneous equations were solved

using Newton's method. The fitting was done as non-linear least squares minimisation using the Levenberg-Marquardt algorithm (Appendix I).(Moré, 1978)

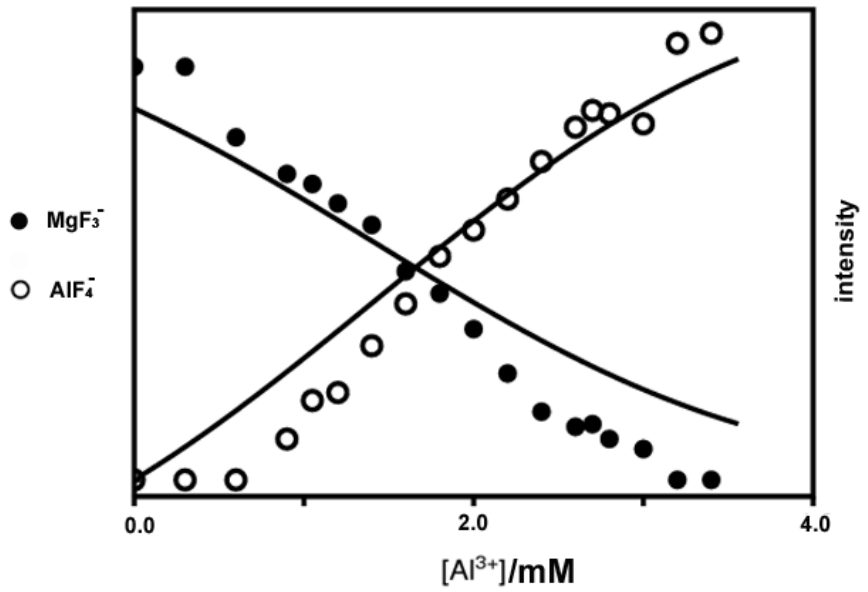


Figure 5.10 Data fitting of the intensity of the resonances for AlF_4^- and MgF_3^- complexes from aluminium titration. [•] Averaged integration of the four peaks from AlF_4^-, and [◦] averaged integration of two peaks F_B and F_C from MgF_3^-.

The fitting results give the value of dissociation constant of $K_d(MgF_3^-) \approx 3.7$ fM of $K_d(AlF_4^-) \approx 840$ fM.

5.3.4 pH titration

It has been shown that the pH-switch in fluoride coordination in a metallofluoride TSA complex does not derive from an AlF_4^- moiety converting into AlF_3^0. Rather, as the pH increases AlF_4^- is progressively replaced by MgF_3^-.(Baxter *et al.*, 2008) To confirm our conclusion that TSA complex prioritizes anionic charge character over a very broad range of pH at the expense of geometry (tbp versus octahedral), a pH titration was performed to show cAPK is not an exception.

At pH 8.0, which is the pH used for crystallisation of the complex giving structure 1L3R, the occupancy of AlF_4^- and MgF_3^- is about 20:80 judged by fluorine signal integration (**Figure 5.11**). This is significantly different from the corresponding ratio seen in the reanalysis of the electron density, where the double occupancy ratio is AlF_4^-: $MgF_3^- = 70:30$. This difference is very likely due to a pH change in the diffusion

process during crystallisation or to a different thermodynamic equilibrium in crystal and solution for AlF_4^- and MgF_3^- complexes. More significantly, there is no sign of any AlF_3^0 species over the full range of pH from 7.0 to 9.0.

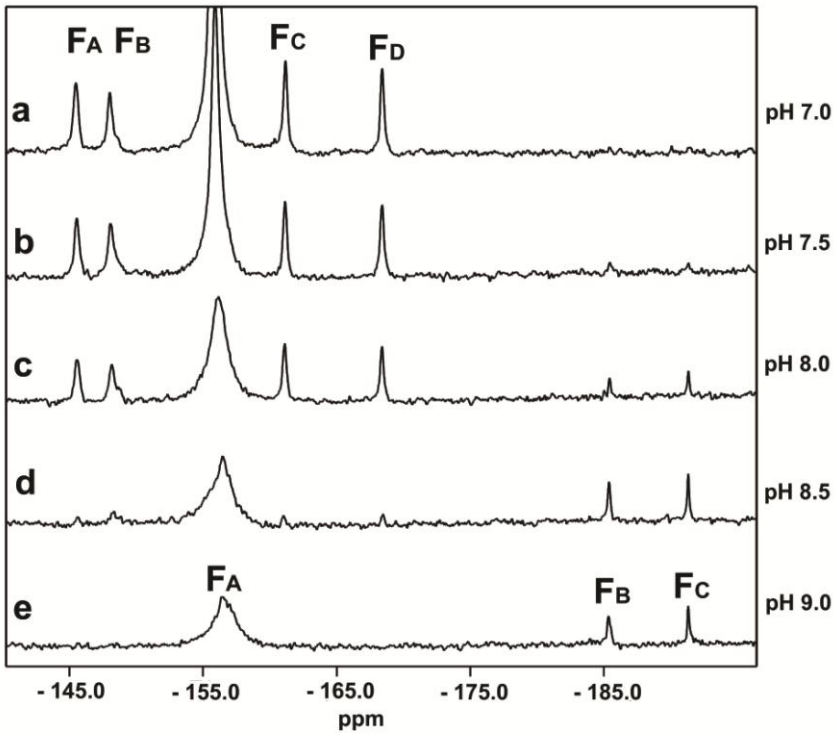


Figure 5.11 pH Titration of the cAPK-ADP-AlF_4^--SP20 TSA complex from pH 7.0 to pH 9.0 monitored by ^{19}F NMR: (**a**) pH 7.0, F_A = –146 ppm, F_B = –148 ppm, F_C = –161 ppm and F_D = –168 ppm; (**b**) pH 7.5; (**c**) pH 8.0; (**d**) pH 8.5; (**e**) pH 9.0, F_A = –157 ppm, F_B = –185 ppm and F_C = –191 ppm, which corresponds exactly to the cAPK-ADP-MgF_3^--SP20 TSA complex.

5.3.5 Physiological significance of MgF_3^- inhibition of cAPK

The unusually high preference of cAPK for MgF_3^- versus AlF_4^- has potentially significant physiological consequences. While ingestion of small amounts of fluoride can help prevent dental caries and strengthen bones,(Caverzasio *et al.*, 1996; Caverzasio *et al.*, 1998) adverse affects on human health can result from chronic ingestion at high fluoride doses.(Kaur *et al.*, 2009) The molecular basis of these effects remains unresolved, but links have been found to the ability of fluoride to interact with phosphoryl transfer enzymes. Fluoride has a strong inhibitory action on a large number of such enzymes *in vitro* which, in some examples, has been related directly to the formation of such magnesium or aluminium complexes.(Gazzano *et al.*, 2010) This raises the question whether such enzyme inhibition is responsible for some of the adverse consequences for life. Moreover, surveys of the effects of fluoride in drinking water have repeatedly drawn attention to the symbiotic potentiation of fluoride risk by

aluminium, particularly in relation to bone and neurological problems.(Toxicology, 2006) Aluminium has been shown to potentiate the effect of fluoride on tyrosine phosphorylation and osteoblast replication *in vitro* and bone mass *in vivo*, specifically linked to G-protein activation and tyrosine phosphorylation.(Caverzasio *et al.*, 1996; Kaur *et al.*, 2009; Gazzano *et al.*, 2010) The binding of fluoroaluminate to G-proteins, which makes them constitutively active, and the inhibition of phosphotyrosine phosphatases leads to an intracellular increase of tyrosine phosphorylation and activation of the mitogen-activated protein kinase pathway.(Gazzano *et al.*, 2010) Both these activation effects would be mitigated by any inhibitory effects of fluoride on the downstream kinases, *e.g.* MEK6.(Kaur *et al.*, 2009)

The present study shows that a focus on inhibition by fluoroaluminate may be too narrow. The previously reported overwhelming dominance of AlF_4^- complexes over MgF_3^- complexes in some enzymes *in vitro* supported the proposal that it was unlikely that the detrimental physiological effects of elevated fluoride levels resulted from enzyme inhibition by MgF_3^- moieties.(Baxter *et al.*, 2006) However, under physiological conditions, where levels of Mg^{2+} exceed those of Al^{3+} typically by some three orders of magnitude, any fluoride inhibition of cAPK, for example, would be dominated by an MgF_3^- complex rather than by an AlF_4^- complex. Extrapolating this behaviour to the many related protein kinases, which are active components in the signalling pathways that regulate metabolism, transcription, cell-cycle progression, differentiation, apoptosis, and many other key cellular functions,(Johnson, 2009) fluoromagnesate inhibition should not be discounted as a source of physiological toxicity.

5.3.6 Charge balance calculation of cAPK

On the basis of our ^{19}F NMR studies, we have found the binding of a monoanionic moiety, such as MgF_3^- or AlF_4^-, mimics the transferring PO_3^- phosphoryl group and sustains a close balance between the positive and negative charges within a sphere of radius up to 15 Å. Such charge balance predominates over geometry in the prioritization of the components of the TSA.(Blackburn *et al.*, 2003; Baxter *et al.*, 2006; Baxter *et al.*, 2008; Cliff *et al.*, 2010)

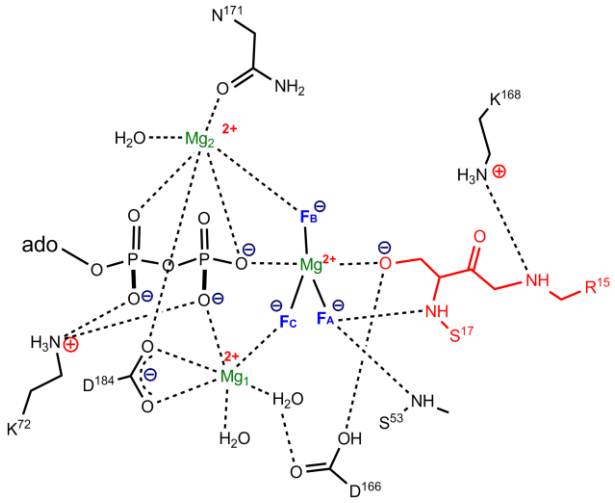


Figure 5.12 Scheme showing charge balance details in the active site of cAPK.

The trifluoromagnesate and tetrafluoroaluminate TSA complexes for cAPK, here established by ^{19}F NMR data, fully support the charge balance hypothesis (CBH) for phosphoryl transfer in this protein (**Figure 5.13**, black line, **Figure 5.14**). By contrast, analysis of the cAPK tbp crystal structure assigned as an AlF_3^0 complex shows clear departure from charge balance (**Figure 5.13**, red line, **Figure 5.14**). It is therefore important to address this apparent inconsistency to confirm the generality of the concept of charge balance.

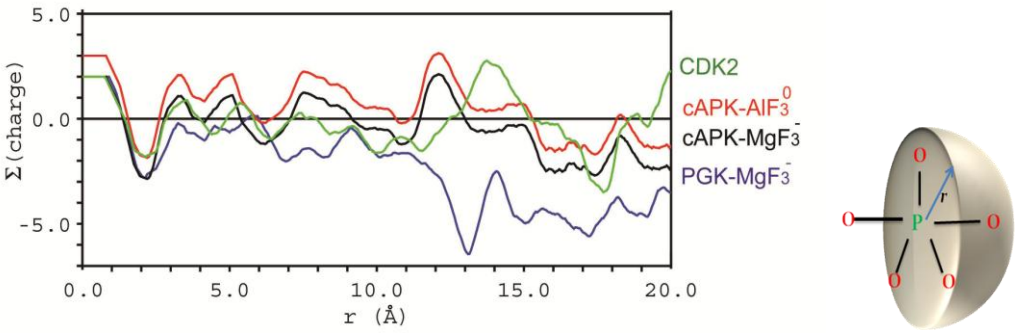


Figure 5.13 Charge balance for cAPK, human PGK, and CDK2. Net charge is calculated for atoms within a sphere of radius *r* (right) centred on the core Mg atom of the MgF_3^- TSA complex (cAPK, black curve; and for CDK2, green curves) and for the same structure analysed as an tbp AlF_3^0 complex (red curve). The charge balance calculation has included phospho-Ser and phospho-Thr residues as appropriate. The equivalent analysis for human PGK is included for comparison (purple curve).

Indeed, charge neutrality appears to hold nearly to 15 Å from the transferring phosphoryl group, which happens to be the distance to the regulatory Thr 197 phosphate residue (**Figure 5.15**). It is especially significant that eight of the amino acids of the

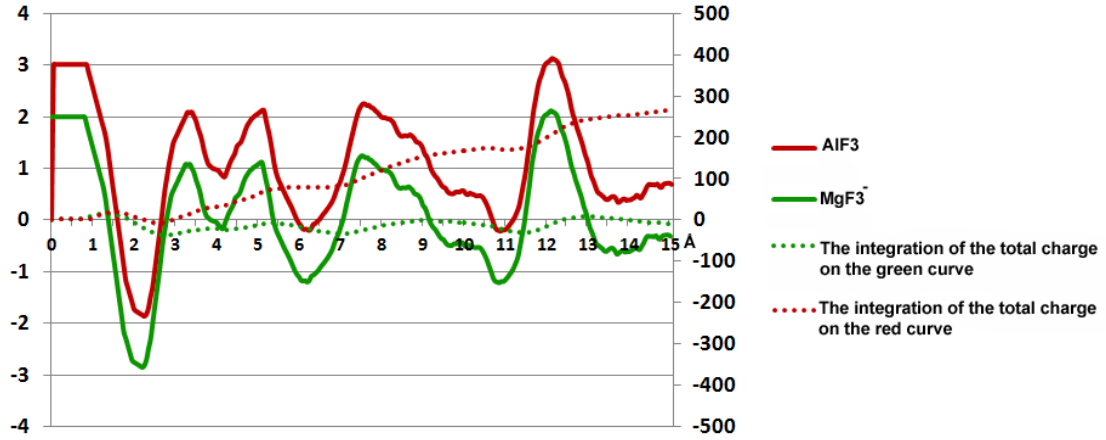
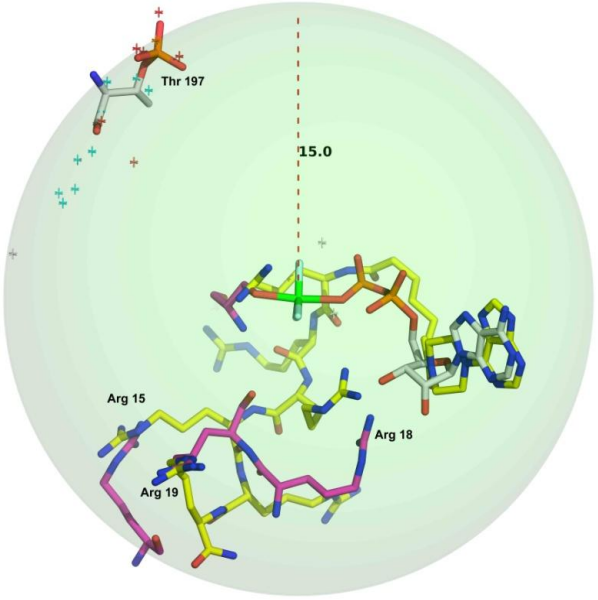


Figure 5.14 The dotted line shows the integration of total charge on the charge balance curve (solid line) out to 15 Å. It is a measure of the uniformity of the balance of total charge in a sphere of radius *r* for cPKA calculated as AlF_3^0 (red curves) and as MgF_3^- (green curves). The deviation from zero of the dotted red line shows the red curve deviates towards positive net charge.

SP20 substrate peptide, including the four cationic residues: Arg 15, Arg 18, Arg 19, and His 23, are located within a 15 Å shell around the transferring phosphoryl group (**Figure 5.15**). This shows that charge balance for protein kinases, and maybe also for protein phosphatases, is achieved by involving residues in the substrate as well as in the catalytic protein.

Figure 5.15 Alignment of cAPK-ADP-MgF_3^--SP20 TSA complex (part of SP20 in magenta sticks) and a complex of cAPK with its bisubstrate AR670 (AR670 in yellow sticks). The green sphere has a radius of 15 Å and the three positively charged arginines: Arg 15, Arg 18, and Arg 19 in SP20 are labelled. Phospho-Thr 197 is also highlighted which is on the surface of the sphere.



Furthermore, a recent structural analysis of a complex between cAPK and a rationally-designed, synthetic bisubstrate analogue having adenosine linked to a hexa-arginine

chain shows that the inhibitor AR670 occupies both the adenosine and peptide sites of the cAPK complex. When that structure is aligned with the metal fluoride TSA complex, it can be seen that all six of the arginine cations are located within a 15 Å sphere centred on the transferring phosphoryl group locus (**Figure 5.15**). It can be concluded that CBH may contribute to the future design of competitive inhibitors of protein kinases.

5.3.7 cAPK-ADP-BeF_3^--SP20 GSA complex

So far, no BeF_3^- GSA complex structure has been reported in the PDB for cAPK and it is very important for us to establish a GSA complex using wild type cAPK to compose the whole reaction coordination for cAPK with the transition state together with a product state structure. A pioneering BeF_3^- complex would be able to indicate any conformational change of a protein in the ground state compared to the transition state. It could also facilitate our dynamic dispersion measurements in NMR in future.

The only published near-ground state structure (PDB: 1ATP) in the PDB consists of cAPK, ATP and a 20-residue peptide inhibitor from PKI (5~24, TTYADFIASGRTGRRN**A**IHD), which was grown in low concentrations of Mg^{2+}, then soaked in Mn^{2+}. One Mn^{2+} interacts with the side chains of Asn 171 and Asp l84 as well as with a H_2O molecule. The second Mn^{2+} bridges the γ- and β-phosphates and coordinates to Asp 184 and two H_2O molecules. When this structure was overlaid with 1L3R, the TSA complex structure, the two metals align very well without many other changes to the key residues, except that the attacking hydroxyl from SP20 Ser 17 is within 2.6 Å of F_C and F_B.

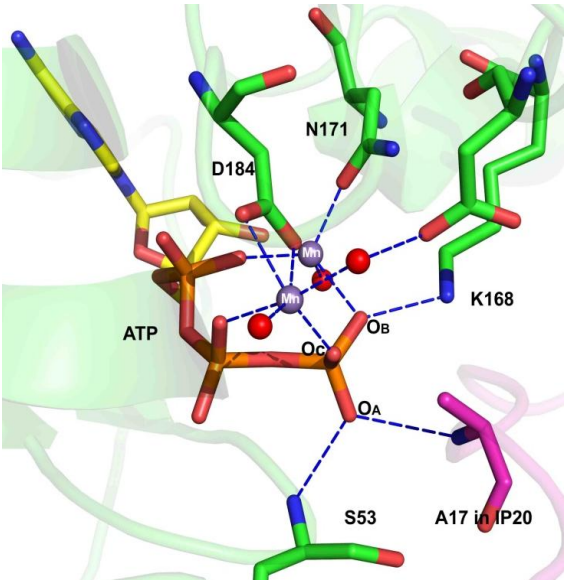


Figure 5.16 Ground state structure (PDB: 1ATP) of cAPK complex (green sticks) with IP20 substrate peptide (magenta sticks) and ATP (yellow sticks). Manganese ions are in purple spheres and water molecules are in red spheres. The designated oxygens in γ-phosphate of ATP correspond to the assignment of the fluorines in ^{19}F NMR spectrum of cAPK-ADP-BeF_3^--SP20 GSA complex in **Figure 5.17**.

cAPK readily forms a cAPK-ADP-BeF_3^--SP20 TSA complex in the presence of magnesium, beryllium, fluoride, ADP, and the substrate peptide SP20. Complete conversion of the MgF_3^- complex into a BeF_3^- complex by adding 5 mM Be^{2+} was observed by ^{19}F NMR which indicates that the affinity of beryllium fluoride for the substrate complex is at least 4 times bigger than that of MgF_3^-.

The two upfield resonances in cAPK-ADP-BeF_3^--SP20 GSA complex are for F_B at –181.8 ppm and F_C at –186.5 ppm (**Figure 5.17**). They are assigned as labelled in the crystal structure in **Figure 5.16**. Similar to the chemical shifts in MgF_3^- TSA complex, these two fluorine chemical shifts are more upfield than for other single magnesium kinases and phosphotransferases we have yet studied. In $UMPK_{dicty}$, the chemical shifts of the fluorines in ADP-BeF_2-UDP complex are –155.4 and –174.1 ppm while in the βPGM_{D10N}-BeF_3^--G6P complex these values are –148.3, –153.1, and –178.2 ppm. The reason for the upfield shift for cAPK is that the deshielding effect on fluorine from the second magnesium is not as strong as hydrogen bonding interactions. So the fluorine nucleus which interacts with Mg^{2+} is more shielded than the one interacting with hydrogens. F_A raises at –165.6 ppm which has been identified by beryllium titration (data not shown) to avoid the misassignment as the chemical shift of F_A is very close to other free BeF_x species in solution.

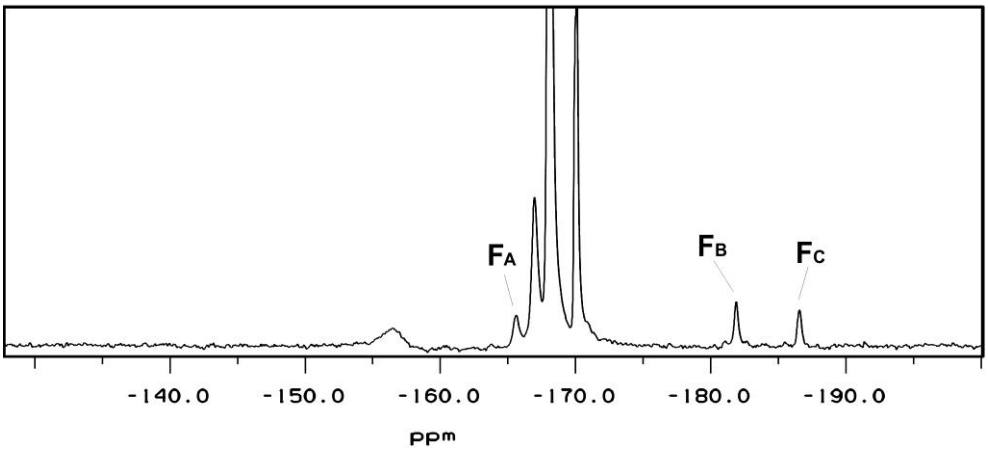


Figure 5.17 ^{19}F NMR spectra of cAPK-ADP-BeF_3^--SP20 GSA complex in 50 mM HEPES buffer at pH 6.5 containing 300 mM NaCl, 5 mM DTT, 5 mM ADP, 20 mM $MgCl_2$, 5 mM $BeCl_2$, 40 mM NH_4F, 0.5 μM cAPK, and 1.2 mM substrate peptide SP20 at 25°C. The resonances from protein-bound BeF_3^- GSA complex are F_A= –165.6 ppm, F_B = –181.8 ppm and F_C = –186.5 ppm.

5.4 Conclusions

cAPK has been successfully expressed in its active form under rigorous conditions. Our ^{19}F NMR results establish the tbp moiety previously identified in the active site of cAPK with the substrate peptide SP20 is indeed trifluoromagnesate (**Section 5.3.2**), thus correcting its misassignment as an 'AlF_3^0' TSA in the published structure (PDB: 1L3R). That result establishes that cAPK conforms to the CBH as do other kinases. An equivalent complex can be produced by the use of the heptapeptide Kemptide (**Section** 5.3.2.4), which has not yet been cocrystallised with cAPK. The pH titration analysis confirms that there is no AlF_3^0 species over the whole pH range from 7.0 to 9.0, particularly at pH 8.0 at which the cAPK-ADP-AlF_3^0-SP20 TSA complex was crystallised (**Section 5.3.4**). Aluminium titration exhibits the intriguing feature of the cAPK-ADP-MgF_3^--SP20 TSA complex formation that the binding constant of MgF_3^- TSA is only ~10 fold weaker than for the AlF_4^- TSA complex (**Section 5.3.3**). That contrasts sharply with the observation that in most other small molecule kinases and phosphatases, Al^{3+} out-competes Mg^{2+}, even when Mg^{2+} is present in 500 molar excess of aluminium. These data raise the possibility that, under physiological conditions where Mg^{2+} levels exceed those of Al^{3+} typically by some three orders of magnitude, fluoride inhibition of a wide range of proteins may be dominated by MgF_3^- complexes rather than by the more studied AlF_4^- complexes. Finally, the ^{19}F NMR data for the formation of a cAPK-ADP-BeF_3^--SP20 GSA solution complex points to the feasibility of forming a BeF_3^- GSA in the solid state (**Section 5.3.7**).

Chapter 6. RhoA-RhoGAP (Small G Protein)

6.1 Background

6.1.1 Introduction to small G proteins

G Proteins (guanine nucleotide-binding proteins) are a family of proteins involved in transmitting chemical signals outside the cell and causing changes inside the cell. G Proteins regulate metabolic enzymes, ion channels, transporters, and other parts of the cell machinery. They control transcription, motility, contractility, and secretion.(Neves *et al.*, 2002) G Proteins were discovered when Gilman and Rodbell investigated stimulation of cells by adrenaline. They found that when adrenaline binds to a receptor it stimulates a G protein first. This G protein can subsequently turn on adenylate cyclase, which produces the second messenger cyclic AMP. Without a G protein, there is no biological effect. For this discovery, they were awarded the 1994 Nobel Prize in Physiology or Medicine.(Nobelprize.org, 1994)

There are two distinct families of G proteins. Heterotrimeric G proteins, also called 'large' G proteins, are activated by G protein-coupled receptors and made up of α, β, and γ subunits. There are also 'small' G proteins (20 ~ 25kDa) that belong to the Ras superfamily of small GTPases. These monomeric proteins are homologous to the alpha subunit found in heterotrimers.

The Rho family is a member of the Ras homologue superfamily, which is a class of small GTPase proteins known to regulate the actin cytoskeleton and it was these that first drew attention to this family of signalling proteins. The Rho family consists of 20 members, including RhoA, RhoB, RhoC, RhoD, Rac1, Rac2, Rac3, RhoG, Cdc42, *etc*.(Boureux *et al.*, 2007) RhoA is the member A subclass of Rho and regulates the actin cytoskeleton in stress fibres.(Ridley *et al.*, 1992) Generally speaking, RhoA has 5 α-helixes and 6 β-sheets. There are two 'switch regions' that display conformational changes between GDP- and GTP-bound forms: switch I (residues 27 ~ 36) and switch II (residues 62 ~ 69, in which Gln 63 is catalytically important). A phosphate binding loop (P loop, residues 13 ~ 20) wraps around the nucleotide so strongly that GTP and GDP bind to the Rho with dissociation constants in the nanomolar range. A Mg^{2+} ion plays a key role in bringing together the functional regions of the phosphate-binding, switches I and II (**Figure 6.1**).(Wei *et al.*, 1997; Ihara *et al.*, 1998)

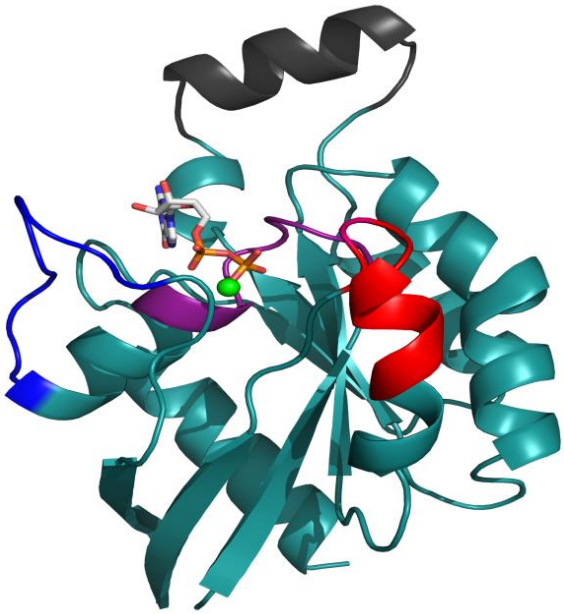


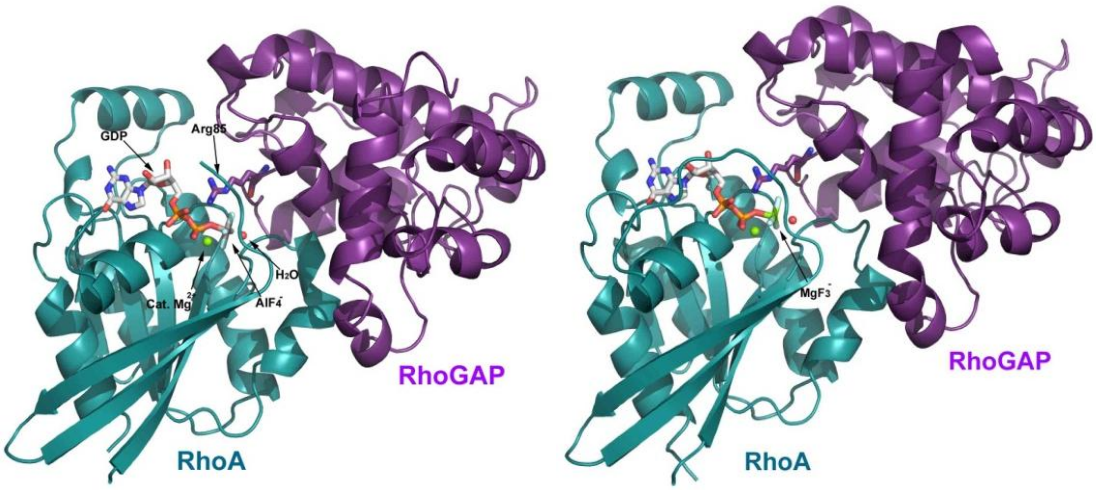
Figure 6.1 Three dimensional structure of RhoA (PDB:1OW3). Catalytic magnesium is a green sphere, GDP is in white sticks, switch I (residues 27 ~ 36) is in blue, switch II (residues 62 ~ 69) is in red, the insert (residues 124 ~ 136) is in black, and the P loop (residues 13 ~ 20) is in purple.

Small G proteins are sometimes referred as 'molecular switches'. This means they can be turned 'on' by binding GTP and turned 'off' after the GTP is hydrolysed into GDP and phosphate (P_i). The off-to-on process requires dissociation of GDP, an intrinsically slow and reversible process, and then binding of GTP. This process can be accelerated by a guanidine nucleotide exchange factor (GEF) (Bos *et al.*, 2007) and regulated by other proteins such as GDP dissociation inhibitors (GDI).(Pfeffer *et al.*, 1995) The on-to-off process is also intrinsically slow and irreversible and involves hydrolysis of GTP to GDP catalysed by small G proteins which is in fact very slow, some 10^{-4} s^{-1} at 37°C. This process is catalysed by GTPase-activation proteins (GAP).(Boguski *et al.*, 1993)

The structural and biochemical details of GAP-assisted GTP hydrolysis was first obtained from a Ras-RasGAP complex.(Scheffzek *et al.*, 1997) The function of RasGAP is to stabilise the position of Glu 61 of Ras, which orientates the attacking water in line with the γ-phosphate of GTP. In addition, Arg 789 of RasGAP is positioned into the phosphate-binding site and stabilises the TS by neutralising negative charge at the γ-phosphate (even though in that publication, the phosphate mimics was identified as AlF_3^0). This arginine was given the name 'arginine finger', because of the functional similarity to the arginine found in the helical insertion of α-subunits of heterotrimeric G proteins.(Coleman *et al.*, 1994; Sondek *et al.*, 1994)

Almost at the same time, a similar mechanism was found for RhoGAP-assisted hydrolysis for RhoA.(Rittinger *et al.*, 1997) In this RhoA-GDP-AlF_4^--RhoGAP

complex, Glu 63 of RhoA is structurally and functionally equivalent to Glu 61 of Ras. AlF_4^- is mimicking the γ-phosphate of GTP in transfer. Critically catalytic residue Arg 85 of RhoGAP plays the same role as Arg 789 in RasGAP which interacts with AlF_4^- in the TSA complex structure. Comparing the $RhoGAP_{R85A}$ mutant to wild type RhoGAP, it loses activity by 240-fold (from 5.4 s^{-1} down to 0.023 s^{-1}), but it still increases the intrinsic hydrolysis of the GTP by 160-fold (from 1.4×10^{-4} s^{-1} to 0.023 s^{-1}).(Graham *et al.*, 1999) This suggests that both the conformational changes of Glu 63 of RhoA on binding RhoGAP and the insertion of Arg 85 of RhoGAP to the active site during the transition state are both important to catalysis. The markedly depressed catalytic activity observed in the GAP mutant lacking Arg 85 is caused by its inability to neutralise the negative charge on phosphoryl group and so stabilise the transition state.



a **b**

Figure 6.2 Cartoon structures of (**a**) RhoA-GDP-AlF_4^--RhoGAP (PDB: 1TX4) and (**b**) RhoA-GDP-MgF_3^--RhoGAP (PDB: 1OW3) TSA complexes.

Thus far, all the TSAs in G proteins were interpreted as AlF_3^0 or AlF_4^-, depending on the geometry of the MF_x moiety observed in the active site until the first MgF_3^- TSA complex for RhoA-RhoGAP was published in 2002 (**Figure 6.2b**).(Graham *et al.*, 2002) The significance of this complex will be addressed later in this chapter.

6.1.2 The physiological significance of the inhibition of G proteins

The inhibition of G proteins has already attracted people's attention for quite a long time. Aluminium fluoride is believed to be one of the main causes of dementia, because aluminium and fluoride are widely present in everyday life, such as aluminium cookers,

kitchen foil, water pipes, tea, toothpaste, *etc*. These chemicals are not very harmful by themselves but, aluminium in combination with fluoride as AlF_4^- can mimic the γ-phosphate in GTP in its transition state in G proteins and is therefore able to inhibit their GTPase activity and so result in dementia.(Shad, 2006)

6.1.3 Why we are examining TSAs for RhoA-RhoGAP

The X-ray scattering properties of aluminium and magnesium are not sufficiently different to allow them to be distinguished by X-ray crystal diffraction data. Thus, MgF_3^- was first identified using PIXE in RhoA-GDP-MgF_3^--RhoGAP complex as introduced in **Section1.8**. The result is satisfactory, but it lacks any direct observation of fluorine. ^{19}F NMR is ideal for monitoring the formation of metal fluoride complexes since many other complexes have already been analysed in PSP, βPGM, and hPGK by this means.

In this chapter, it is first shown that our NMR results confirm the PIXE analysis. Next, because the RhoA-RhoGAP system is a good example to reveal the evolution of natural charge balance between a GTPase and its activating protein, the concept of charge balance in small G proteins is developed. By recuiting the R85A mutant of RhoGAP, the spectial structural feature in the active site is revealed by both ^{19}F NMR and crystallography. These results show the CB concept applied to small G proteins has to be modified from its use for phosphomutase βPGM and small molecule kinase HsPGK. To test this new CB feature, a substrate analogue GTPγF was used to examine the hydrolytic activity with or without the contribution from an arginine finger in RhoGAP. The results show the CB situation is more complex for a small G protein and our CB model needs to be refined to be able to accord with our observation.

6.2 Materials and Experiments

6.2.1 The expression of RhoA and RhoGAP

The pGEX-2T expression vectors carrying human p29RhoGAP or RhoA gene were provided by Dr. Katrin Rittinger (MRC, National Institute for Medical Research, London). This vector **(Figure 6.3)** contains a T7 promoter that is specific to T7 RNA polymerase, a *lac* operator which can block transcription, a sequence for GST domain followed by an *N*-terminal thrombin cleavage site, and an ampicillin resistance gene.

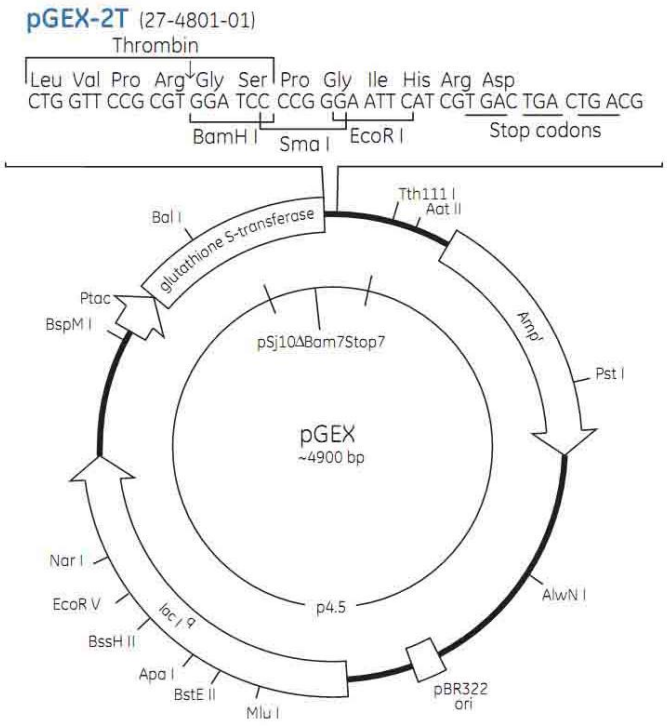


Figure 6.3 Vector map of pGEX-2T.

The full sequence of GST-RhoA is:

```
        10         20         30         40         50         60
MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID

        70         80         90        100        110        120
GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV

       130        140        150        160        170        180
DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK

       190        200        210        220        230        240
KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LVPRGSPAAI RKKLVIVGDG

       250        260        270        280        290        300
ACGKTCLLIV NSKDQFPEVY VPTVFENYVA DIEVDGKQVE LALWDTAGQE DYDRLRPLSY

       310        320        330        340        350        360
PDTDVILMCF SIDSPDSLEN IPEKWTPEVK HFCPNVPIIL VGNKKDLRND EHTRRELAKM

       370        380        390        400        410
KQEPVKPEEG RDMANRIGAF GYMECSAKTK DGVREVFEMA TRAALQARRG KKKSGCLVL
```

After cleavage, the number of amino acids in RhoA is 193; molecular weight is 21735.0; calculated pI is 5.83. The residues in bold show the sequence for thrombin cleavage.

The full sequence of GST-p29RhoGAP is:

```
        10         20         30         40         50         60
MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID
```

```
        70         80         90        100        110        120
GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV

       130        140        150        160        170        180
DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK

       190        200        210        220        230        240
KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LVPRGSPGIH VKLEQLGIPR

       250        260        270        280        290        300
QVLKYDDFLK STQKSPATAP KPMETPPRPP LPNQQFGVSL QHLQEKNPEQ EPIPIVLRET

       310        320        330        340        350        360
VAYLQAHALT TEGIFRRSAN TQVVREVQQK YNMETGLPVD FDQYNELHLP AVILKTFLRE

       370        380        390        400        410        420
LPEPLLTFDL YPHVVGFLNI DESQRVPATL QVLQTLPEEN YQVLRFLTAF LVQISAHSDQ

       430        440        450        460        470
NKMETTNTNL AVVFGPNLLW AKDAAITLKA INPINTFTKF LLDHQGELFP SPDPSGL
```

After cleavage, the number of amino acids in RhoGAP is 242; molecular weight is 27416.5; calculate pI is 6.18. The residues in bold show the sequence for thrombin cleavage.

Proteins were expressed in *E. coli.* strain BL21-DE3-Rosetta-pLyss as GST-RhoA and GST-RhoGAP. The procedure of transformation and overnight culture follow the standard procedure as described in Chapter 2. Ampicillin was used as antibiotic. After inoculation, the cells were incubated at 37°C, 250 rpm until an OD_{600} of 0.6 ~ 0.8 was reached. Expression was induced with 1 mM (final concentration) IPTG and the culture was incubated for a further 5~6 h at 30°C, 250 rpm. Cells were harvested by centrifugation at 4,200 rpm, 4°C for 20 min and the pellets were stored at –80°C.

6.2.2 Site-Directed Mutagenesis of RhoGAP to generate R85A mutant

The primers for site-directed mutagenesis to make R85A mutant had the sequence shown below, in which the bold letters are the mutated bases at Arg 85:

Forward primer 5’-GAGGGCATCTTC**GC**GAGGTCGGCC-3’

Reverse primer 5’-GGCCGACCTC**GC**GAAGATGCCCTC-3’

Tm= 71.3°C, GC%= 70.8%

Both DNA sequences of p29RhoGAP wild type and p29RhoGAP-R85A mutant were confirmed by DNA sequencing (Source Bioscience-geneservice).

6.2.3 Protein purification

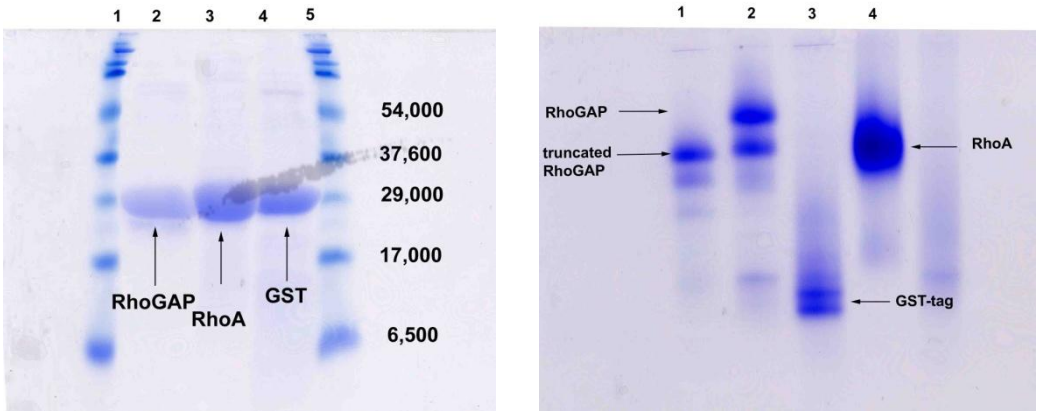
Human Rho A, RhoGAP and its mutant R85A were produced following the same procedure.

Table 6.1 Composition of buffers and solutions for RhoA-RhoGAP purification.

Lysis Buffer (1 L):	Tris-HCl	50 mM, pH 7.6
	$MgCl_2$	5 mM
	NaCl	50 mM
	DTT	1 mM
	trace $CHCl_3$ as anti-microbiotic	
Thrombin Digestion Buffer (1 L):	Tris-HCl	50 mM, pH 8.0
	$MgCl_2$	5 mM
	DTT	1 mM
	NaCl	150 mM
	trace $CHCl_3$ as anti-microbiotic	
Elution buffer for Gel Filtration (1 L):	Tris-HCl	50 mM, pH 8.0
	NaCl	150 mM
	$MgCl_2$	5 mM
	DTT	1 mM
	trace $CHCl_3$ as anti-microbiotic	

The pellet was resuspended in 40 mL lysis buffer with addition of a cocktail protease inhibitor tablet (EDTA free). The cells were lysed by sonication on ice for 4 × 20 s with 1 min gap and then centrifuged for 40 min at 24,500 rpm, 4 °C. The supernatant was filtered through a 0.45 µm syringe filter. The filtered lysate with GST-RhoA was mixed 1:1 with 40 mL Glutathione Resin (Sigma-Aldrich, Cat. No. G4510C) in a 2.5 × 50 mL glass column and incubated on a rotating wheel at 4°C for 2 h. The resin was washed in the column with 150 mL cold lysis buffer (without DTT this time) until the baseline was flat.

The following procedure was used for on-column digestion. The column was equilibrated with 100 mL thrombin digestion buffer and 40 U thrombin was added. The column was incubated on a rotating wheel at 4°C overnight. All the digested solution in the column was collected and the resin was twice washed with 20 mL digestion buffer. The digested solution and the washes were combined and filtered through a 0.2 µm syringe filter and concentrated to 4 mL. This preliminarily purified RhoA-RhoGAP solution was loaded on a pre-equilibrated (with elution buffer) Superdex^{TM} G75 Gel Filtration column (GE Healthcare) followed by elution at a flow rate 1.5 mL/min with Gel Filtration Elution buffer. An SDS-PAGE was run to check the purity (**Figure 6.4**), and then the pure fractions concentrate to a final 0.5 mM.



a **b**

Figure 6.4 (**a**) SDS-PAGE and (**b**) urea gel of digested RhoA and RhoGAP comparing to GST-tag.

6.2.4 Ion exchange of GTPγF

GTPγF (13 mg) was synthesized by PhD student Joanna Szeremeta in the lab of Prof. Dr. Barbara Nawrot, Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, Lodz (arrived June 29, 2010).

The first batch of GTPγF came as its triethylammonium salt. So before being used for activity assay it was put through a strong cation exchange column to change it into Na^+ salt. The column was packed manually using Amberlite IR120 Chromatographic-Grade Ion-Exchange Resins (H^+ form, Sigma, Cat. No.216534). The resin was first washed with 10 column volumes 1 M NaOH to change it into Na^+ form, followed by 10 volumes of distilled water to pH 7 (indicated by pH paper). The GTPγF·Et_3NH^+ water solution was loaded onto the pre-equilibrated column and the flow through was collected. The column was washed with 3 column volumes of distilled water and the washings were also collected. The combined flow through and washings were freeze dried overnight. The GTPγF·Na^+ compound was then ready for use.

6.2.5 NMR sample preparation

The sole commercially available RhoGEF is a human recombinant hDbs-6His protein, DH/PH domain (shorted for 'Dbs' in the rest of this thesis) purchased from Cytoskeleton Inc USA (Cat. Number GE01; 2 × 50 μg) as a lyophilised powder. The protein was reconstituted to give a 2mg/mL stock in 20 mM Tris-HCl, pH 7.5, 0.5 mM $MgCl_2$, 0.5% sucrose, 0.1% dextran by addition of 25 μL of ddH_2O. RhoA is in 50 mM

Tris-HCl pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT. All the controls and assays were carried out in 50 μM RhoA (with 1 eq. GDP bound), 2.5 μM RhoGEF, 50 μM RhoGAP depending on the experiment, in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT buffer.

6.3 Results and Discussion

6.3.1 RhoA-GDP-MgF_3^--RhoGAP TSA complex

To make a RhoA-GDP-MgF_3^--RhoGAP TSA complex, 10 mM NH_4F is first added to 0.5 mM RhoA with 1 eq. GDP bound in 50 mM Tris-HCl at pH 7.5, 5 mM $MgCl_2$ and 1 mM DTT in a 5 mm glass NMR tube. The ^{19}F NMR spectrum of RhoA-GDP-MgF_3^--RhoGAP complex shows three resonances which are well-resolved (F_A = –143.4 ppm, F_B = –154.3 ppm, F_C = –173.4 ppm), present in a 1:1:1 ratio that is stoichiometric with RhoA-RhoGAP, corresponding to the three fluorines coordinate to central magnesium (**Figure 6.5a**).

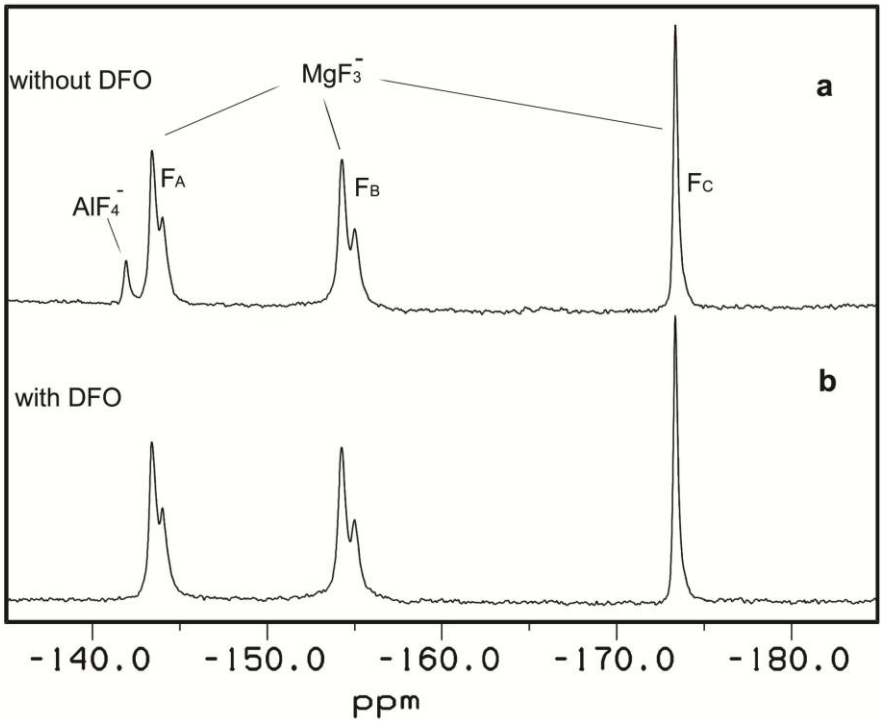


Figure 6.5 Comparison of the ^{19}F NMR of RhoA-GDP-MgF_3^--RhoGAP TSA complex: (**a**) in 50 mM Tris-HCl at pH 7.5, 0.5 mM RhoA (with 1 eq. GDP bound), 5 mM $MgCl_2$, 1 mM DTT, and 10 mM NH_4F; (**b**) the same sample with the supplemented with DFO. The protein-bound MgF_3^- TSA signals are F_A = –143.4 ppm, F_B = –154.3 ppm, and F_C = –173.4 ppm. The protein-bound AlF_4^- TSA signal is at –142.0 ppm.

There is considerable evidence from ^{19}F spectra in **Figure 6.5a** that, in the absence of DFO, trace Al^{3+} can form a RhoA-GDP-AlF_4^--RhoGAP observed at –142 ppm. This addresses the importance of including DFO during any fluoromagnesate survey by NMR or crystallography. To make sure that there is no contamination from Al^{3+} for fluoromagnesate complex, the sample solution has 1 mM DFO present (a powerful aluminium chelator with K_d 10^{-20} M) to chelate the trace Al^{3+} leached out of the NMR tube.(Evers *et al.*,1989) The peak at –142 ppm disappears to give a neat spectrum without aluminium contamination (**Figure 6.5b**).

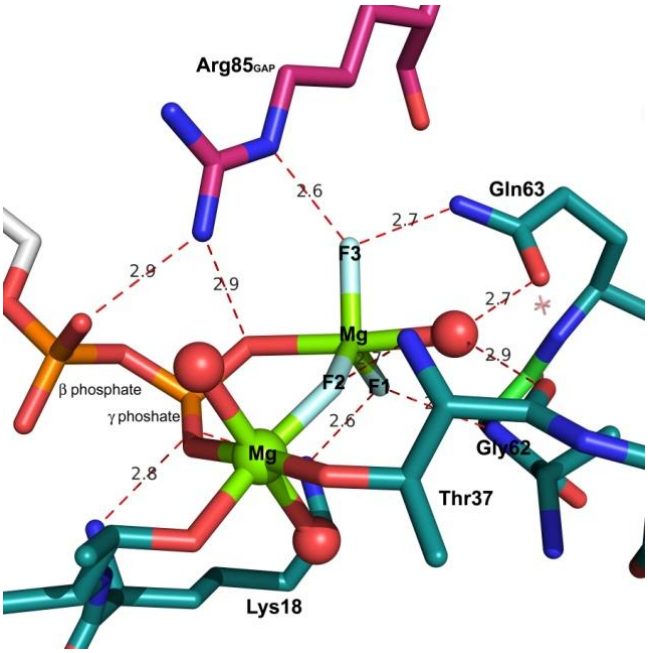


Figure 6.6 The hydrogen bonding network in RhoA-GDP-MgF_3^--RhoGAP TSA complex (PDB:1OW3). RhoA is dark teal and RhoGAP is magenta. Magnesium trifluoride moiety is displayed in kelly and light blue sticks. The catalytic magnesium is a kelly sphere. Waters are red spheres.

SIISs are also measured to verify the consistency of the trifluoromagnesate observed in NMR (**Table 6.2**) and in the existing crystal structure (PDB: 1OW3). The largest SIIS for the RhoA-GDP-MgF_3^--RhoGAP TSA complex is 1.6 ppm for $F_{A,}$ corresponding to F_3 in crystal structure. It correlates with a 2.6 Å hydrogen bond to Arg 85 N–H in RhoGAP and a 2.7 Å hydrogen bond to Gln 63 N–H of RhoA. For F_B, the SIIS is 1.4 ppm, similar to F_A, and it forms a 2.6 Å hydrogen bond to Lys 18 side chain N–H and a 2.7 Å hydrogen bond to backbone amide of Gly 62. The smallest SIIS values F_C (0.7 ppm) result from its primary (1.9 Å) coordination to catalytic magnesium and a 2.7 Å hydrogen bond to backbone amide of Thr 37 (**Figure 6.6**). To conclude, the SIIS of ^{19}F NMR reflects the fact that the three resonances observed in the spectrum are indeed from the tbp MgF_3^- in the RhoA-GDP-MgF_3^--RhoGAP crystal structure.

In addition to the major peaks F_A and F_B, there are two minor resonances at –144.0 ppm and –155.0 ppm which are both 0.6 ~ 0.7 ppm upfield from their major peaks (**Figure 6.7**). The shift is not caused by an isotope effect because a D_2O capillary was used as lock in all the NMR experiments in this work except where the solvent was 100% D_2O. This result may be due to 'conformational breathing' between the two proteins. From the existing crystal structure, it seems that only F_3 (F_A) has direct contact to RhoGAP and right at the interface of the two proteins, but F_1 (F_B) is in contact with Lys 18 and Gln 63, where there is more dynamic flexibility between the P-loop and switch II upon binding than for other parts of the protein.(Ihara *et al.*, 1998) That is the reason why F_C, which is sandwiched between the two metal ions does not have a minor form.

Table 6.2 ^{19}F chemical shifts (ppm) and SIIS (ppm) of the RhoA-GDP-MgF_3^--$RhoGAP_{wt}$ and RhoA-GDP-AlF_4^--$RhoGAP_{wt}$ complexes.

Metal Fluoride Complex		**δF** (D_2O capillary)	**ΔδF** (D_2O capillary-100% D_2O)
RhoA-GDP-MgF_3^--$RhoGAP_{wt}$	F_A (F_3)	–143.4	1.6
	F_B (F_1)	–154.3	1.4
	F_C (F_2)	–173.4	0.7
	average	–157.0	___
RhoA-GDP-AlF_4^--$RhoGAP_{wt}$	F*	–142.0	0.8

* Numbering in the brackets corresponds to the labelling used in PDB crystal structures.

It is worth pointing out that attempts have been made to observe an MgF_x complex in Rho using ^{19}F NMR under the conditions (pH 7.5, 5 mM NaF, 1 mM $MgCl_2$) where a magnesium-dependent Rho-GDP-RhoGAP complex would be expected, but they failed. Graham *et al*. attributed the absence of new resonances to a rapid exchange rate between the free and complex ions being rapid (compared to the chemical shift difference).(Graham *et al.*, 1999) In the light of our success in now forming a fluoromagnesate TSA complex, we believe the explanation is most likely to be a deficient amount of magnesium and fluoride in their sample, considering at 20 mM NaF, the apparent K_d of Mg^{2+} is 1.1 mM in the complex formation with Rho.(Graham *et al.*, 1999)

6.3.2 RhoA-GDP-AlF_4^--RhoGAP TSA complex

RhoA-GDP-AlF_4^--RhoGAP is made by addition of 1 mM $AlCl_3$ to the sample of RhoA-GDP-MgF_3^--RhoGAP TSA complex, which contains 0.5 mM RhoA with 1 eq. GDP

bound, 50 mM Tris-HCl at pH 7.5, 5 mM $MgCl_2$, 10 mM NH_4F and 1 mM DTT in a 5 mm glass NMR tube.

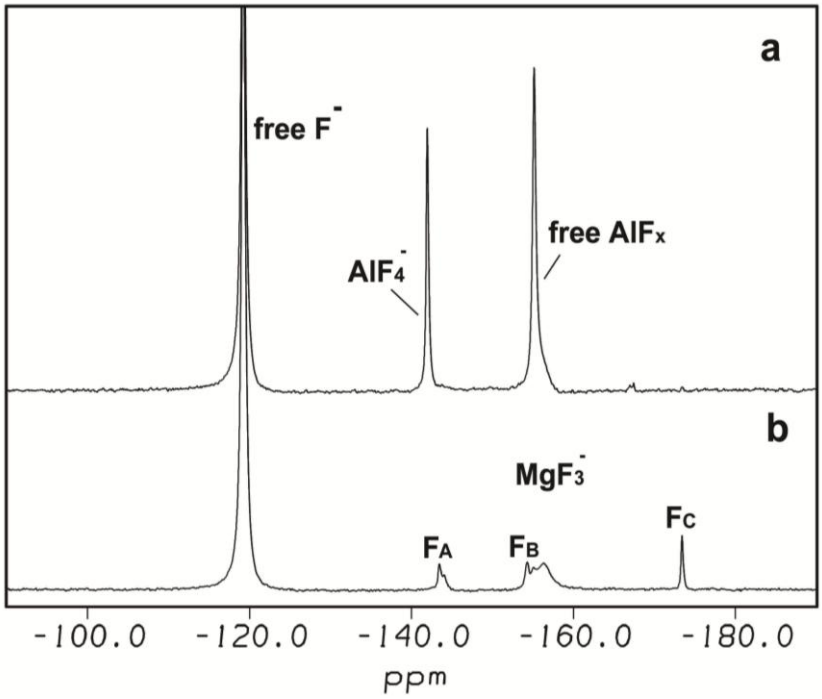


Figure 6.7 ^{19}F NMR spectra of: (**a**) RhoA-GDP-AlF_4^--RhoGAP TSA complex in 50 mM Tris-HCl at pH 7.5, 0.5 mM RhoA (with 1 eq. GDP bound), 5 mM $MgCl_2$, 1 mM DTT, 10 mM NH_4F, 1 mM $AlCl_3$. The signal from protein-bound AlF_4^- TSA as a rotationally averaged peak has a chemical shift at –142.0 ppm and its integration shows there are 4 fluorines in the complex comparing to Fc in (**b**) RhoA-GDP-MgF_3^--RhoGAP TSA complex. The protein-bound MgF_3^- TSA signals are F_A = –143.4 ppm, F_B = –154.3 ppm, and F_C = –173.4 ppm.

RhoA-GDP-AlF_4^--RhoGAP TSA complex shows one rotationally averaged peak in ^{19}F NMR corresponding to the four fluorines coordinate to Al^{3+} (**Figure 6.7a**). A previous study on Gα used ^{19}F NMR to detect AlF_x binding, in which one rotationally averaged peak integrated between 3 to 5 compared to 1 equivalent of protein saturated by stoichiometric Al^{3+}. The calculation of the integration was based on the known value of protein concentration and total fluoride concentration.(Higashijima *et al.*, 1991) In our experiment, since we had already obtained well-resolved resonances in the MgF_3^- complex as an external standard for integration (**Figure 6.7b**), we could establish the accurate value for *x* in AlF_x by comparing the AlF_4^- peak intensity with that of Fc from RhoA-GDP-MgF_3^--RhoGAP. This calibration shows there to be four fluorines in that AlF_x peak corresponding to four equatorial fluorines. The single rotationally averaged NMR peak observed for protein-bound F^- shows the environments of the four protein-

bound F^- ions must be quite similar and/or they must exchange rapidly between the four different loci.

6.3.3 Test of Charge Balance Hypothesis applied to RhoA

As mentioned previously, $RhoGAP_{R85A}$ mutant loses activity by 240-fold relative to the wild type RhoGAP, but it still increases the intrinsic rate of hydrolysis of the GTP by 160-fold. This indicates this small G protein RhoA alone does not have an absolute demand for charge balance (a positive charge from a catalytic residue), at least when compared to other phosphotransfer enzymes we have studied. But charge balance clearly helps the catalysis if it is available. Therefore, it would be very significant and useful to test if Charge Balance Hypothesis (CBH) applies to small GTPases like RhoA when not supported by RhoGAP. The validity of the hypothesis can be easily tested using our TSA setup on RhoA alone and then monitored by ^{19}F NMR.

It is worth pointing out that, by chance, some ^{19}F NMR experiments reported for other purposes used almost the same conditions as ours. But as they used only 5 mM NaF in the setup of other small G proteins in Rho family, such as Cdc42Hs (Hoffman *et al.*, 1998) and Rho (Graham *et al.*, 1999), they found no evidence of formation of aluminium fluoride complexes in the absence of RhoGAP. We suggest that the course was the insufficiency of F^-, based on our experience in the highly concentration-dependent formation of metal fluoride complex in cAPK.

6.3.3.1 A magnesium fluoride complex is not stable for RhoA alone

Although all the studies using ^{19}F NMR to search for metal fluoride complexes in small G proteins have been focusing on aluminium fluoride species, we did not want to miss the possibility of using magnesium fluoride as a probe for charge balance on RhoA. Therefore the ^{19}F NMR spectrum was recorded in 50 mM Tris-HCl at pH 7.5, containing 0.5 mM RhoA (with 1 eq. GDP bound), 5 mM $MgCl_2$, 1 mM DTT, and 10 mM NH_4F. The data showed no formation of RhoA-GDP-MgF_3^- or RhoA-GDP-$MgF_2(H_2O)^0$ in the absence of RhoGAP. This means the binding of MgF_3^- without RhoGAP is much weaker than with RhoGAP, and RhoA alone is not capable of holding MgF_x in the active site when it is exposed on the surface of the protein. Thus attempts to use magnesium fluoride to indicate the net charge in the TSA for RhoA has been unsuccessful.

6.3.3.2 An AlF_3 complex is stable for RhoA alone

Aluminium fluoride complexes have a much greater stability than fluoromagnesate ones not only in free solution but also in the active site of proteins such as hPGK and βPGM and their mutants.

Therefore the ^{19}F NMR spectrum was recorded in 50 mM Tris-HCl at pH 7.5, containing 0.5 mM RhoA (with 1 eq. GDP bound), 5 mM $MgCl_2$, 1 mM $AlCl_3$, 1 mM DTT, and 10 mM NH_4F. Although there is no published data to facilitate identification of this complex lacking RhoGAP, in our experiment, three peaks are present in the ^{19}F spectrum, clearly showing that RhoA-GDP-AlF_3 is formed (**Figure 6.8a**). To make sure that there was no fourth peak obscured by the broad peak of free AlF_x in the spectrum of RhoA-GDP-AlF_3 complex (**Figure 6.8b**), the free F^- signal (at –119 ppm) was suppressed using presaturation, which led to the almost complete suppression of the free fluoroaluminate resonances (at –155 ppm) through exchange. This also shows the complex has a lifetime greater than 1 s, because presaturation of free F^- over this period had no effect on the TSA signals.

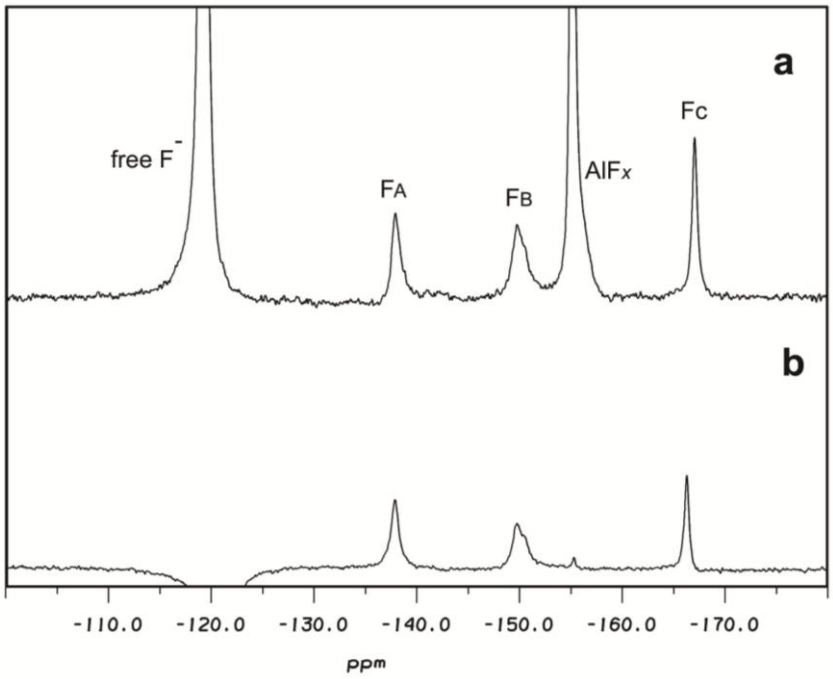


Figure 6.8 ^{19}F NMR spectrum of: (**a**) RhoA-GDP-AlF_3 complex in 50 mM Tris-HCl at pH 7.5, containing 0.5 mM RhoA (with 1 eq. GDP bound), 5 mM $MgCl_2$, 1 mM DTT, 10 mM NH_4F, 1 mM $AlCl_3$ using pulse program *zg*. (**b**) The same sample using presaturation pulse program *zgpr* to saturate the free F^- signal at –119 ppm to remove the problem of non-protein bound AlF_x resonance in fast exchange. The signals from protein bound AlF_3 TSA are F_A = –137.9 ppm, F_B = –149.9 ppm, and F_C = –166.9 ppm.

The direct conclusion is that, even though the conserved arginine from Rho-GTPase-activating protein is essential for efficient catalysis by stabilising the TS, a RhoA-GDP-AlF_3 complex is still stable enough to exist without assistance from GAP, thanks to the intrinsic stability of fluoroaluminate species in water.

6.3.3.3 Choice between tbp AlF_3^0, octahedral $AlF_3(OH)^-$, or $AlF_3(H_2O)^0$ for the RhoA complex

The geometry of the aluminium complex in RhoA-GDP-AlF_3 cannot be assigned using ^{19}F NMR alone. The choice lies between tbp AlF_3^0, octahedral $AlF_3(OH)^-$, or $AlF_3(H_2O)^0$ for the RhoA complex. This choice reduces to the nature of ligand *Z* in the RhoA-AlF_3*Z* complex. Is *Z* OH^-, H_2O, or absent?

We should first consider whether the species could be a tbp AlF_3^0 (*Z* absent). To date there are several cases purporting to be tbp AlF_3^0 complexes that were then proved erroneous by both accurate crystal structure and ^{19}F NMR studies. So far, the only evidence on chemical shift of a tbp AlF_3^0 in βPGM moves 10 ppm upfield on average relative to the octahedral AlF_4^- is only supported by the computational calculation (Webster, unpublished). Even though the calculated value on βPGM is very supportive of our case, but according to the crude nature of the chemical shift calculation, we can not say the geometry is not tbp until we do ^{19}F NMR on that system if we can find one in future or if we can manage to crystallise it.

Table 6.3 ^{19}F chemical shifts (ppm) and SIIS (ppm) of the RhoA-GDP-AlF_3^0 complex.

Metal Fluoride Complex		**δF** (D_2O capillary)	**Δδ F** (D_2O capillary-100% D_2O)
RhoA-GDP-AlF_3^0	F_A	–137.9	0.9
	F_B	–149.9	0.8
	F_C	–166.9	0.3
	average	–151.6	___

It therefore seems likely that the AlF_3 moiety is octahedral as has been shown for other phosphoryl transfer proteins.(Cliff *et al.*, 2010) This raises the question: why did RhoA make an AlF_3*Z* complex and not an AlF_4^-complex? The nature of the fourth ligand (*Z*) needs now to be identified. In principle it could be either H_2O or OH^-.

Let's first consider OH^- as a possibility. At pH 7.5, the OH^- concentration in the buffer is about $10^{-6.5}$ M, $[F^-]$ is at 10 mM (10^{-2} M), some $10^{4.5}$ higher than $[OH^-]$. Lower pH

was also tested at pH 5.5, when the F^- is more than a million times more populated in the buffer that OH^-, but there are still only three resonances arising from AlF_3 in the spectrum (data not shown). This indicates the competition for the fourth position is unlikely to derive from F^- *vs* OH^-.

The competitiveness of H_2O is always high, at about 55 M, and so it has a larger chance to complete a RhoA-GDP-$AlF_3(H_2O)^0$ complex. This is similar to the case of $HsPGK_{K219A}$-3PG-$AlF_3(H_2O)^0$-ADP TSA complex. The removal of the positive charge contributed by Lys 219 in the coordination of the TSA in the wild-type system induces a corresponding unit reduction in negative charge through the loss of one fluoride from the aluminium fluoride moiety in the mutant complex (**Section 1.9**). Apparently, in the crystal structure of the $HsPGK_{K219A}$-3PG-$AlF_3(H_2O)^0$-ADP TSA complex, there is no big perturbation of conformation in the TSA complex. Therefore the protonation state of the oxygen atom in the fourth ligand of our complex can readily be resolved from the ^{19}F chemical shift values as long as they are significantly downfield from those of the $HsPGK_{wt}$-3PG-AlF_4^--ADP TSA complex, consistent with a loss of electron density from the AlF_3^0 moiety. That would result from replacement of F by a less electronegative species such as H_2O.(Cliff *et al.*, 2010)

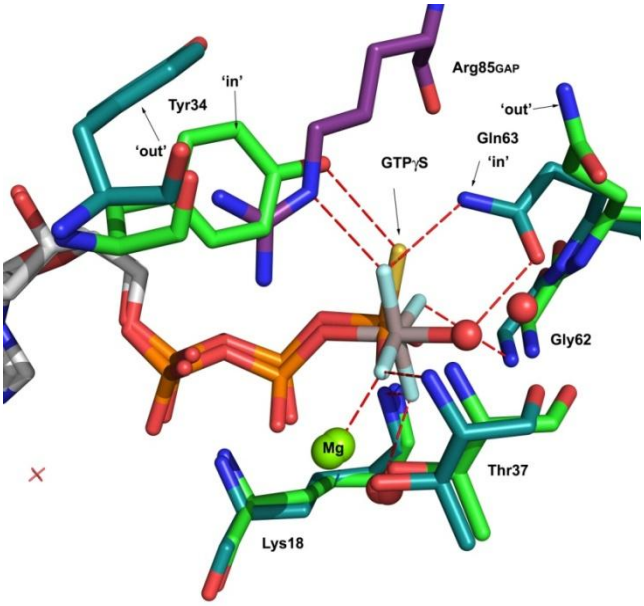


Figure 6.9 Alignment of the ground state structure of $RhoA_{G14V}$-GTPγS complex (PDB: 1A2B, green sticks) with the transition state structure of RhoA-GDP-AlF_4^--RhoGAP complex (PDB: 1TX4, dark teal sticks for RhoA and magenta sticks for RhoGAP). Catalytic magnesium ions are green spheres, fluoride ions are cyan sticks, sulfur is yellow stick and water molecules are red spheres.

By contrast, in the RhoA-GDP-AlF_3Z complex, because of the absence of RhoGAP, not only is the arginine finger missing but the active site of RhoA is very much different for key residues according to the presence or absence of RhoGAP. This is known by comparing the structure of RhoA with the GTP analogue guanosine 5'-3-*O*-

(thio)triphosphate GDPγS (2.4 Å resolution, PDB: 1A2B) and RhoA-GDP-AlF_4^--RhoGAP TSA (1.65 Å, PDB: 1TX4) complex (**Figure 6.9**)

The RhoA-GTPγS complex mimics the active GS in the absence of RhoGAP. In the GS, Tyr 34 is in an 'in' position to form a hydrogen bond with sulfur and occupies the space where the guanine group is when arginine finger (Arg 85) from RhoGAP is present. At the same time, Gln 63 moves out, away from the γ-phosphate and the nucleophilic water. The attacking water is 3.6 Å away from the central phosphorus and only interacts closely with a backbone carbonyl of Thr 37. The backbone amides of Thr 37 and Gly 62 both shift 0.7 Å comparing to those of in RhoA-GDP-AlF_4^--RhoGAP complex, forming a 'looser' active site. However, in both structures the catalytic magnesium and GDP moiety barely move. Based on these data, one can postulate that if the complex under discussion is RhoA-GDP-$AlF_3(H_2O)^0$, its actual structure should be similar to RhoA-GTPγS in the absence of RhoGAP. The distance from the attacking water to aluminium in $AlF_3(H_2O)^0$ moiety should be shorter than 3.6 Å and similar to that in the RhoA-GDP-AlF_4^--RhoGAP structure because of the ionic interaction of Al–O. Although the coordination of this water to Gln 63 side chain might be missing, due to the absence of RhoGAP, it could still be hydrogen bonded to Thr 37 backbone amide and the carbonyl oxgyen of Gly 62. In general, the whole active site of this RhoA-GDP-AlF_3 complex is more exposed to the surface of RhoA and not as tight as in the TS for catalysis with RhoGAP. The relatively smaller SIIS value of all the three resonances relative to the TSA complex supports this (**Table 6.3**). However, the average chemical shift of these three resonances is –151.6 ppm, that is 9.4 ppm upfield from RhoA-GDP-AlF_4^--RhoGAP. We have to conclude that because of the large conformational change in the active site, we cannot simply assign net charge change as the cause of this big change in chemical shift.

Thus far, we still do not know the nature of AlF_3Z. We are only confident on the assignment of F_C, which must be the fluorine between the catalytic magnesium and aluminium because of its upfield chemical shift and smallest SIIS.

In the meantime, a crystal trial has been set up by Erika Pellegrini (ESRF, Grenoble, France) which should identify in due course the coordination of aluminium and other conformational changes in the activity attributed to the binding of RhoGAP (**Appendix III**).

6.3.4 Test of Charge Balance Hypothesis in RhoA-RhoGAP$_{R85A}$

As described previously, CBH has been proved to be valid in HsPGK$_{K219A}$ mutant which forms an HsPGK$_{K219A}$-3PG-$AlF_3(H_2O)^0$-ADP complex and in βPGM with a similar manner (**Chapter 1.9**). ^{19}F NMR of the metal fluoride TSA complex is a very convenient indicator to identify the number of fluorines and from the chemical shift changes of the fluorine signals we can seek to identify the fourth ligand as either H_2O or OH^-.

The role of RhoGAP$_{R85A}$ mutation is to seek to prove the CBH in RhoA-RhoGAP in the same way as for wild type proteins. If the CBH applies, RhoA-RhoGAP$_{R85A}$ should behave similarly to the equivalent mutant of βPGM or HsPGK. However, this does not seem to be the case.

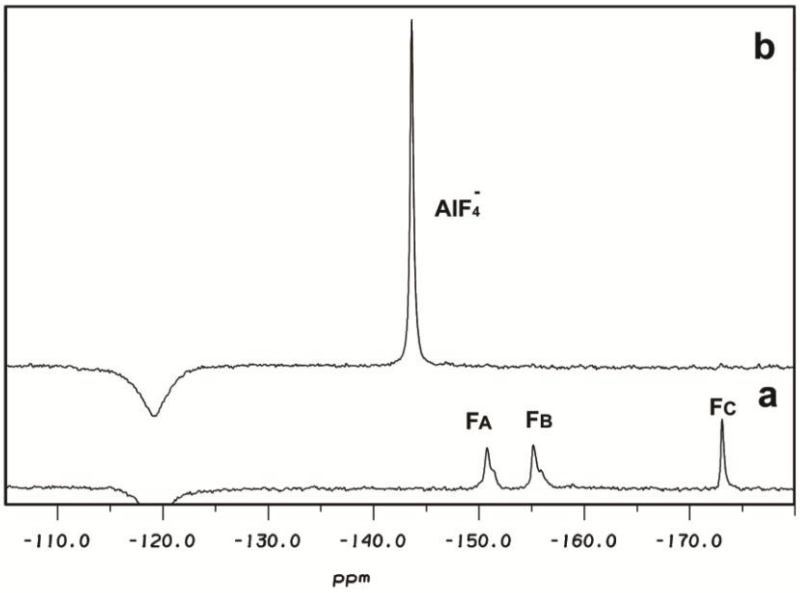


Figure 6.10 ^{19}F NMR spectra of (**a**) RhoA-GDP-MgF_3^--RhoGAP$_{R85A}$ TSA complex in 50 mM Tris-HCl at pH 8.0, 0.5 mM RhoA (with 1 eq. GDP bound), 0.76 mM RhoGAP$_{R85A}$, 150 mM NaCl, 10 mM $MgCl_2$, 5 mM DTT, and 20 mM NH_4F. The signals from protein-bound MgF_3^- TSA are F_A = –150.8 ppm, F_B = –155.1 ppm, and F_C = –173.0 ppm, in 6.5: 7.3: 5.6 ratio. (**b**) RhoA-GDP-AlF_4^--RhoGAP$_{R85A}$ TSA complex in the same conditions as (**a**) supplemented with 2.5 mM $AlCl_3$. The rotationally averaged signal from protein-bound AlF_4^- TSA is at –143.6 ppm, which has an integration of 30.9.

Using the same experiment conditions as for the wild type MgF_3^- complex, ^{19}F NMR (**Figure 6.10**) shows that, RhoA-RhoGAP$_{R85A}$ with GDP formed a TSA complex, giving three ^{19}F resonances in 1:1:1 ratio, characteristic of a trifluoromagnesate, MgF_3^-, complex. The chemical shifts for F_A and F_B in this RhoA-GDP-MgF_3^--RhoGAP$_{R85A}$

TSA complex are typical of fluorides in other MgF_3^- TSA complexes coordinated only by hydrogen bonds. The chemical shift of F_B is –155.1 ppm, 0.8 ppm upfield from that of F_A in the wild type complex. The chemical shift of F_C is –173.0 ppm, 0.4 ppm downfield relative to the corresponding wild type F_C. By contrast, F_A has the largest upfield shift of 7.4 ppm. This indicates the large local shielding effect on F_3 resulting from the deletion of cationic Arg 85 in RhoGAP. It is noteworthy that in this complex of the mutant the two downfield peaks, F_A and F_B, still retain minor forms, but the chemical shift differences are smaller than those in the wild type complex.

Table 6.4 ^{19}F chemical shifts (ppm) and SIIS (ppm) of the RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ and RhoA-GDP-AlF_4^--$RhoGAP_{R85A}$ complexes.

Metal FluorideComplex		**δF** (D_2O capillary)	**ΔδF** (D_2O capillary-100% D_2O)
RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$	F_A (F_3)	–150.8	1.0
	F_B (F_1)	–155.1	1.4
	F_C (F_2)	–173.0	0.6
	average	–159.6	———
RhoA-GDP-AlF_4^--$RhoGAP_{R85A}$	F*	–143.6	0.6

* Numbering in the brackets corresponds to the labelling used in PDB crystal 1OW3 and 1TX4.

The RhoA-GDP-AlF_4^--$RhoGAP_{R85A}$ TSA complex is made by adding 2.5 mM Al^{3+} to the MgF_3^-sample. It gives a rotationally averaged resonance at –143.6 ppm that is 1.6

Table 6.5 Summary of the chemical shifts (δF) and SIIS (ppm) of the fluoride resonances observed for the metal-fluoride TSA complexes of RhoA-RhoGAP.

Metal Fluoride Complex		**δF** (D_2O capillary)	**ΔδF** (D_2O capillary-100% D_2O)
RhoA-GDP-MgF_3^--$RhoGAP_{wt}$	F_A	–143.4	1.6
	F_B	–154.3	1.4
	F_C	–173.4	0.7
	average	–157.0	———
RhoA-GDP-AlF_4^--$RhoGAP_{wt}$	F*	–142.0	0.8
RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$	F_A	–150.8	1.0
	F_B	–155.1	1.4
	F_C	–173.0	0.6
	average	–159.6	———
RhoA-GDP-AlF_4^--$RhoGAP_{R85A}$	F*	–143.6	0.6
RhoA-GDP-AlF_3	F_A	–137.9	0.9
	F_B	–149.9	0.8
	F_C	–166.9	0.3
	average	–151.6	———

* The peak is rotationally averaged.

ppm upfield from the complex formed with wild type RhoGAP. The reason for this shift should be as the same as that of RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ TSA complex, as will be discussed later in this chapter. This result matches a similar experiment published for the same complex in which the chemical shift change was reported as 1.5 ppm, although the authors did not explain why the chemical shift moves upfield so much after the mutation.(Graham *et al.*, 1999)

The formation of a RhoA-GDP-AlF_4^--$RhoGAP_{R85A}$ complex also positively proves that the mutation R85A has only a small effect on the ability of RhoGAP to form the AlF_x complex.(Graham *et al.*, 1999) Our result thus opposes the hypothesis that the deletion of the arginine finger would completely destabilise the TSA.(Hoffman *et al.*, 1998)

Table 6.6 Differences in chemical shifts (ΔδF) of fluoride resonances observed for the metal-fluoride TSA complexes of RhoA-$RhoGAP_{wt}$ and RhoA/ $RhoGAP_{R85A}$.

Metal Fluoride Complex		ΔδF (D_2O capillary)	ΔΔδF (D_2O capillary-100% D_2O)
MgF_3^-	F_3	7.4	0.6
	F_1	0.8	0.0
	F_2	0.4	0.1
AlF_4^-	F^*	1.6	0.2

* The peak is rotationally averaged.

6.3.4.1 ^{19}F NMR identifies this is a different case from PGK_{K219A} and βPGM_{K145A}.

The formation of RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ and RhoA-GDP-AlF_4^--$RhoGAP_{R85A}$ TSA complexes raises the question why a reciprocal chemical mutation in the negative charge of the substrate nucleotide is not observed as a consequence of the protein mutation in this small G protein?

The first expectation was that in the R85A mutant complex it is unlikely for F_1 to not interact with any species occupying the space arising from the change of arginine to Ala 85. There should be a counter cation of fluoride such as Na^+/H_3O^+ or NH_4^+ coordinating to F_A in these two mutant complexes as in the case of βPGM_{K145A} (**Figure 6.11b, section 1.9)** because in both samples, the Na^+ concentration is 150 mM (protein buffer) and NH_4^+ is 20 mM (NH_4F stock). So a control experiment was set up using NaF as fluoride source to avoid interference from ammonium ions, giving a total Na^+ concentration of 170 mM (150 mM from protein buffer and 20 mM from NaF stock) (**Figure 6.11b**). However, the ^{19}F NMR still showed the same profile and all the three

resonances stayed at the same chemical shift. Therefore, the possibility of an ammonium ion (both hydrated and non-hydrated) can confidently be excluded as NH_4^+ had been excluded in the control experiment.

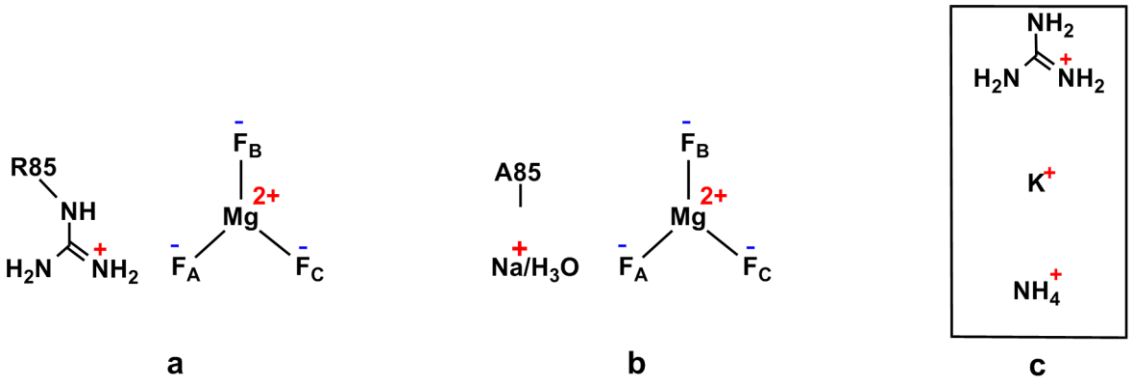


Figure 6.11 Schematic diagram of the fluoromagnesate complexes of RhoA-RhoGAP. (**a**) RhoA-GDP-MgF_3^--$RhoGAP_{wt}$ TSA complex showing the side chain of Arg 85 coordinating the F_A position to balance the charge of the metal fluoride moiety. (**b**) RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ TSA complex showing how sodium/H_3O^+ ions are positioned in the F_A coordination site to balance the charge of the MgF_3^- metal fluoride moiety. **(c)** Other counter cations to F_A in the ^{19}F NMR spectrum of RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ TSA complex which are proved to be not possible.

If we assume that the conformation of the active site does not change much as a result of the R85A mutation, then the size of the cavity is about 5~7 Å in all three dimensions (**Figure 6.12**). Sodium ion itself has an ionic radius of 1.16 Å, which is probably too small to make any remarkable interactions with other lone pair electron donors. The hydrated sodium ion radius is 2.76 Å, which is the right size to be accommodated in the groove and make sufficient ionic and hydrogen bonding interactions to F_3 and surrounding residues (**Figure 6.12**). Furthermore, the possibility of a hydrated sodium cation is particularly important to satisfy the CBH. A further possibility is water or protonated water could be the second ligand to F_3. However, a single water molecule is also too small to fill the cavity, so if F_3 formed a hydrogen bond to a water molecule it would have to be one of a network of waters. Such a hydrogen bond interaction ought to have been signalled by the SIIS measured in D_2O.

Other control experiments were performed by removing all the sodium cations by using KF or guanidinium fluoride as the fluoride source through buffer exchange (**Figure 6.11c**). Because the potassium and guanidinium cations are also potential counter ions of F_A that can bind in the cavity instead of Na^+, the chemical shift of F_A should change for a different counter ion with different electron density. However, ^{19}F NMR of all

these controls gave the same profile which means the chemical shift change either has no relevance to a counter cation, or, if there is a bound cation, it has to be Na^+/H_3O^+ with a very high affinity that cannot be competed out by K^+ or guanidinium or even NH_4^+. However, in the absence of any structural evidence, this argument remains speculative.

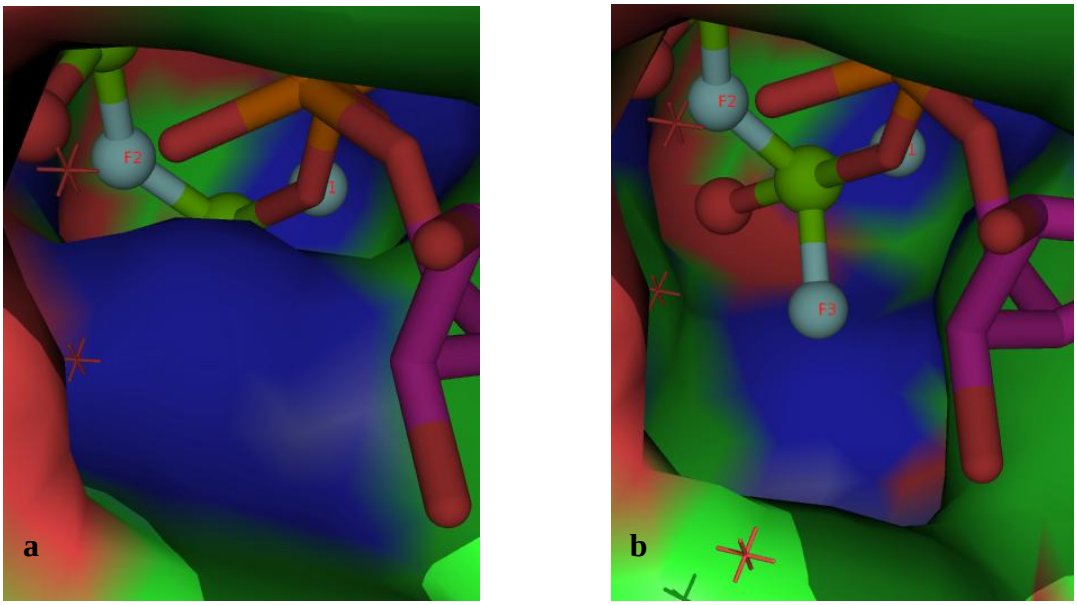


Figure 6.12 Inner molecular surface of the active sites in (**a**) RhoA-GDP-MgF_3^--$RhoGAP_{wt}$ TSA complex (PDB:1OW3): (**b**) RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ complex deduced from structure (**a**). The blue filling in (**a**) is the side chain of Arg 85 of RhoGAP.

6.3.4.2 Crystal structures show Rho proteins behave in the same way, but conformational change determines the charge of metal fluoride moieties.

Simply by using ^{19}F NMR, we cannot explain the above anomalous behaviour of the metal fluoride moieties. We therefore turned to crystal structure evidence and sought for other rationalisations. The first insight was gained by overlaying structures of RhoA-GDP-MgF_3^--$RhoGAP_{wt}$ (PDB: 1OW3) and Cdc42-GDP-AlF_3-$RhoGAP_{R305A}$ (PDB: 2NGR, 1.9 Å), another small G protein. Cdc42 is a close member to RhoA in the Rho family with 51% sequence similarity and its GAP mutant $Cdc42GAP_{R305A}$ (also named as $RhoGAP_{R305A}$) is identical to $RhoGAP_{R85A}$.

In the RhoA-GDP-MgF_3^--$RhoGAP_{wt}$ structure (**Figure 6.13**, cyan sticks), F_A forms a 2.6 Å hydrogen bond to the side chain amide of RhoA Gln 63 and another 2.7 Å hydrogen bond to the δNH in the guanidine group of RhoGAP Arg 85. The Gln 63 side chain carbonyl oxygen accepts a 2.7 Å hydrogen bond from the attacking water. Tyr 34 and Pro 36 sandwich Arg 85 from RhoGAP to force it to stay in the active site through

π-cation interaction.(Ma *et al.*, 1997) In the mutation structure of Cdc42-GDP-AlF_3-$RhoGAP_{R305A}$ (**Figure 6.13**, green sticks), it can be seen that Gln 63 still retains the function of holding the attacking water at an in-line position. But the side chain aromatic ring of Tyr 34 swings back into the active site to fill the space left by removal of Arg 85 and its aromatic ring lies almost in the same plane as the guanidinium group of Arg 85. The phenolic –OH forms a hydrogen bond (2.3 Å) with F_A, as in the GS complex (PDB: 1A2B).. There are other big conformation changes on the side chains of Arg 86 and Glu 81 of RhoGAP in the two structures, but as they have no direct or close interactions with the metal fluoride moiety they are not likely to have caused the big chemical shift change on F_A. It should be mentioned that the assignment of the tbp moiety as neutral AlF_3 in structure 2NGR is unreliable because 2000 fold more Mg^{2+} (0.2 M) than Al^{3+} (0.1 mM) was present in the crystallisation solution at pH 6.0.(Nassar *et al.*, 1998)

The conformational change of Tyr 34 in the mutant was established by solving the crystal structure of the RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ TSA complex at 2.2 Å resolution by Erika Pellegrini (ESRF, Grenoble) (see **Appendix III**). It shows that Tyr 34 is coordinated to F_A and also stacks against the Pro 36 ring at 3.6 Å, suggestive of a stabilising effect π-π interaction.(Imai *et al.*, 2009) The special conformation of Tyr 34 in the mutant may partly explain why it can still stabilise the negatively charged MgF_3^- after deletion of Arg 85in $RhoGAP_{R85A}$. The side chain of tyrosine has a rather high pK_a around 10.5, not quite as high as arginine (pK_a 12.5), but there is substantial electron withdrawn effect from the benzene ring to make it a fairly good hydrogen bond donor without carrying a positive charge. It has been suggested that Tyr 34 acts as a pseudo-arginine finger, and it has been shown that if Tyr 34 is removed, the intrinsic GTP hydrolytic activity of mutant Y34S is reduced at least 20%.(Fidyk *et al.*, 2002)

Does this result mean the CBH does not apply to small G proteins? We need more evidence from structural analysis. Five structures in the PDB and one from Grenoble (Erika Pellegrini) can be compared. These comprise four TSA structures of wild type small G proteins with their GAPs, RhoA-GDP-MgF_3^--$RhoGAP_{wt}$ (PDB: 1OW3, at 1.8 Å), Cdc42-GDP-AlF_3^0-$Cdc42GAP_{wt}$ (PDB: 1GRN, at 2.1 Å), Ras-GDP-AlF_3^0-$RasGAP_{wt}$ (PDB: 1WQ1, at 2.5 Å resolution), and Rac1-GDP-AlF_3^0-$Sptp_{wt}$ TSA complex (PDB: 1G4U, at 2.3 Å resolution) along with two of GAP mutant TSA

structures RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ and Cdc42-GDP-AlF_3^0-$Cdc42GAP_{R305A}$ (PDB: 2NGR, at 1.9 Å resolution). All these four small G proteins are from the Ras homologue superfamily and utilise an arginine finger to stabilise the TS. More importantly, they all have a tyrosine next to arginine finger spatially.

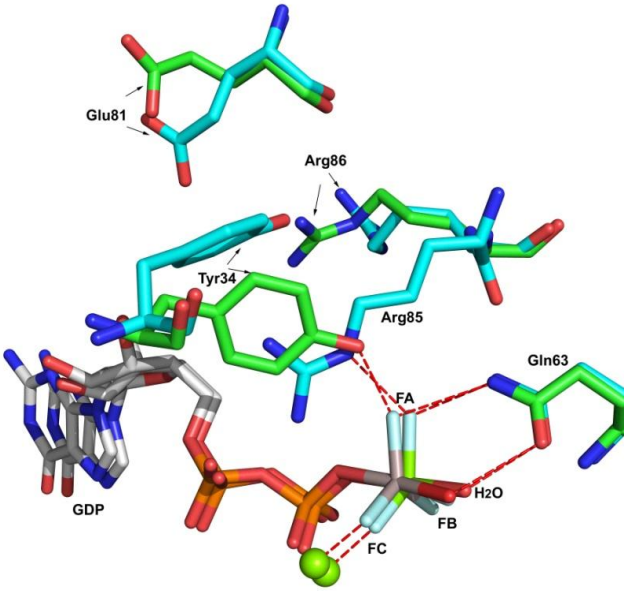


Figure 6.13 Superposition of Cdc42-GDP-AlF_3^0-$RhoGAP_{R305A}$ (PDB: 2NGR, green) and RhoA-GDP-MgF_3^--$RhoGAP_{wt}$ (PDB: 1OW3, cyan) structures. GDP is white sticks, fluorines are sky blue, Mg^{2+} ions are green spheres, and Al^{3+} is grey. The residues with big conformational change are highlighted.

Five of the wild type and GAP mutant TSA structures for these small G proteins (except Ras-RasGAP) have a conserved sp^2-hybridised water molecule Wat1 that always bridges the α-phosphate oxygen, backbone carbonyl oxygen of Val 35 (or Val 33) and the catalytic magnesium (**Figure 6.14 a-d, f**). Because the conformational change in the active site by the mutation makes RhoA and its GAP is still capable of keeping the negative charge on the γ-phosphate in the TSA (so does Cdc42-Cdc42GAP), it means the enzyme complex RhoA-$RhoGAP_{R85A}$ can still stabilise the TS, if less efficiently. This explains why the R85A mutation only reduces activity 160 fold rather than completely abolish it.

However, the deletion of Arg 85 and rotation of Tyr 34 side chain destroys the solvation locally. It is sensible to suggest that a counter cation sits in the gap to interact with anionic α-phosphate, for example in place of Wat2 or Wat3 is in **Figure 6.14c**, and this counter ion could be Na^+ or NH_4^+ or even H_3O^+. This counter ion could also be beneficial to enhance the binding of the α-phosphate when some key hydrogen bonds are lost due to the mutation, such as to backbone carbonyls of Val 35 and Glu 31. By this means, the total charge in the active site could still remain the same, and charge balance still applies. But this 'charge balance' is focused not on the γ-phosphate but on

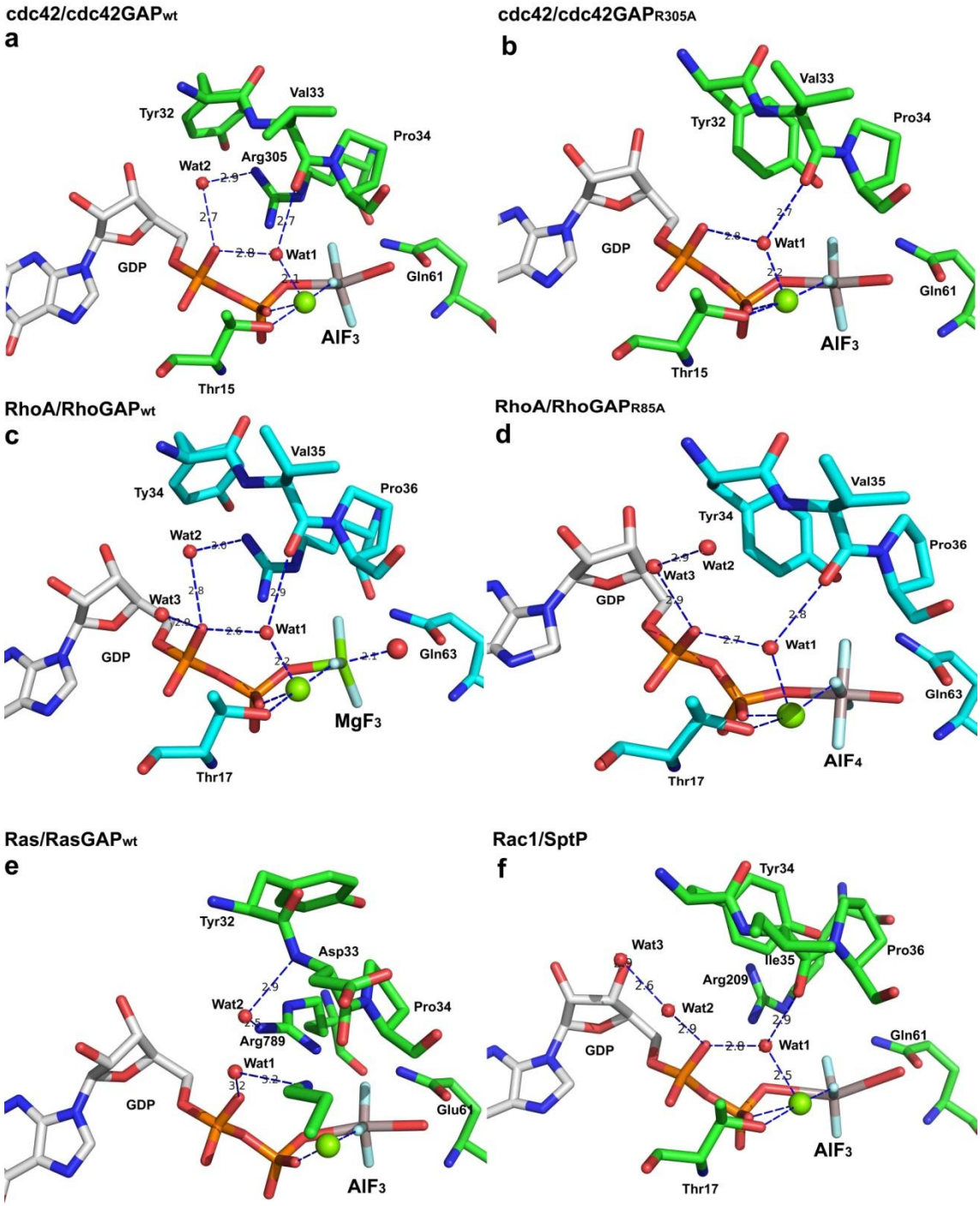


Figure 6.14 Comparison of the structures to show the difference of the position of conserved waters. (**a**) Cdc42-GDP-AlF_3-Cdc42GAP$_{wt}$ TSA and (**b**) Cdc42-GDP-AlF_3-Cdc42GAP$_{R305A}$ TSA complexes. (**c**) RhoA-GDP-MgF_3^--RhoGAP$_{wt}$ TSA and (**d**) RhoA-GDP-AlF_4^--RhoGAP$_{R85A}$ TSA complexes. (**e**) Ras-GDP-AlF_3-RasGAP$_{wt}$ TSA complex. (**f**) Rac1-GDP-AlF_3-Sptp$_{wt}$ TSA complex. Wat1 is conservative in both wt and mutant structures. In all the structures, GDP is white sticks, fluorines are light blue, Mg^{2+} ions are green, and Al^{3+} is grey. Water molecules are red spheres.

the α-phosphate which also interacts with the arginine finger through a water molecule, as Wat2 or Wat3 in the structures as seen in **Figure 6.14**.

6.3.4.3 Is it conceivable that Arg 85 is neutral in wild type structure 1TX4 and 1OW3?

Application of the CBH suggests two scenarios for the wild type and mutant RhoA-RhoGAP complexes: Arg 85 with negatively charged TSA in the wild type, or Tyr 34 with negatively charged TSA in the mutant. Since Tyr 34 cannot be positively charged, it is necessary to ask whether Arg 85 in the wild type structure may be uncharged (deprotonated)? Although there is a general consensus that arginine is always cationic because of its high pK_a of 13, just two cases have been considered as exceptions in the literature.(Harms *et al.*, 2011) In addition, there is a widely held opinion that the TS for phosphoryl transfer of small G proteins is rather dissociative (loose) so that no charge neutralisation from an arginine would be required. The major role for the arginine finger can be to provide the special geometry and multi-hydrogen-bonding capability in the active site. In PDB: 1TX4, all the protons in the guanidine group can be assigned to the corresponding hydrogen bonding acceptors, including H4 which forms a long 3.3 Å hydrogen bond to the benzene ring of Tyr 34. H5 is an exception because it points

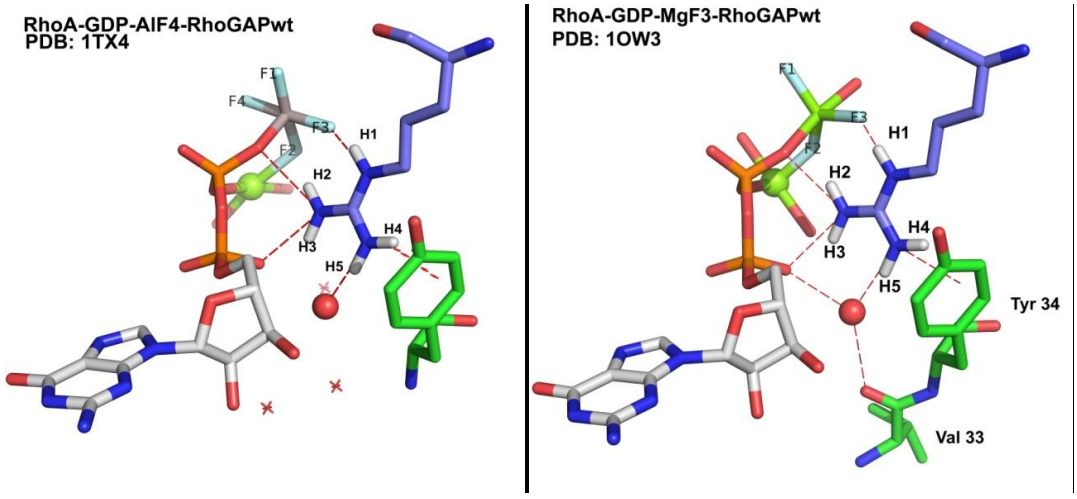


Figure 6.15 Structures of RhoA-GDP-AlF_4^--$RhoGAP_{wt}$ (PDB: 1TX4) and RhoA-GDP-MgF_3^--$RhoGAP_{wt}$ (PDB: 1OW3). Arg 85 is in purple sticks, hydrogens 1 ~ 5 shown in white. GDP is in white sticks, and Tyr 34 is in green sticks. The water molecule coordinating to H5 is a red sphere.

towards a water (red sphere in **Figure 6.15**, left) which can be either a proton donor or an acceptor. After comparing this with the full structure 1OW3, we noticed that there are three residues Phe 31, Glu 32 and Val 33 not present in structure 1TX4. It is clear

that the backbone carbonyl of Val 33 accepts a hydrogen bond from this water (red sphere in **Figure 6.15**, right) as does the non-bridging oxygen on the α-phosphate. Therefore, the hydrogen bond to Arg 85 must be donated from the guanidine H5. This structural analysis rules out the possibility that the Arg 85 guanidine moiety is neutral.

6.3.4.4 Charge balance calculation evidence

As described in **Chapter 1**, a charge balance calculation has been carried out on other proteins to compare the capacity to control charge between phosphoryl transfer and non-phosphoryl transfer enzymes.(Cliff *et al.*, 2010) It shows an apparent common pattern between the net charge and distance from the reaction centre such that within a radius of approximately 15 Å the net charge is close to zero. By contrast, the net charge can either increase or decrease dramatically in non-phosphoryl transfer enzymes, showing there is not efficient charge balance with respect to the substrate.

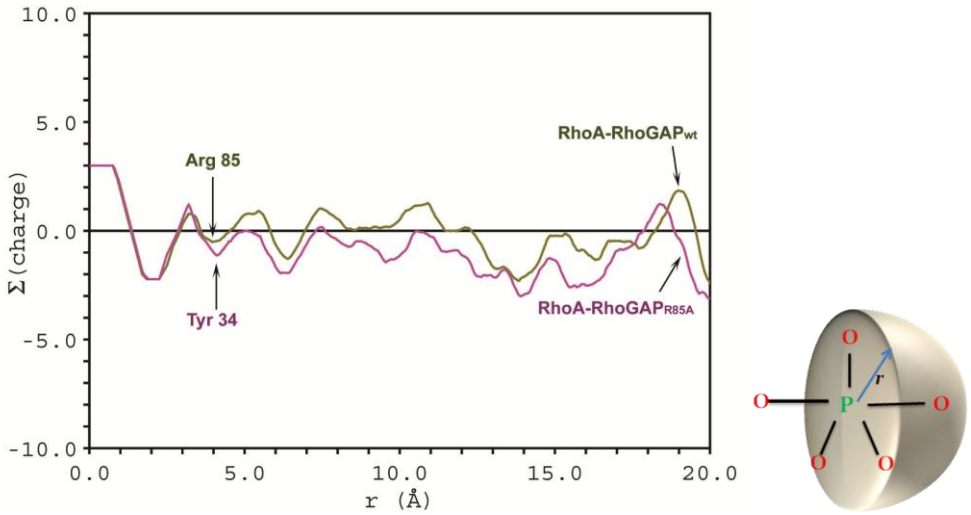


Figure 6.16 Charge balance for RhoA-GDP-AlF_4^--$RhoGAP_{wt}$ (green curve, calculation based on PDB: 1TX4) and RhoA-GDP-AlF_4^--$RhoGAP_{R85A}$ (magenta curve, calculation based on the unpublished 2.2 Å structure). Net charge is calculated for atoms within a sphere of radius *r* centred on the core Al of the AlF_4^- TSA complex.

In the CB calculation, the charge on each atom in all residues is simply assigned without taking into account fractional charges that might affect the status of the residue. So Arg 85 is net +1 on the guanidinium side chain and Tyr 34 is net zero, For example, the partial positive charge provided by the edge of the Tyr rings is ignored in the first-order charge balance calculations. The calculation results for RhoA-GDP-AlF_4^--

$RhoGAP_{wt}$ and RhoA-GDP-AlF_4^--$RhoGAP_{R85A}$ show the total charge in the sphere only begin to deviate out from 6 Å while the key functions of Arg 85 and Tyr 34 are effectively around 4.5 Å. The curve for R85A (**Figure 6.16**, magenta) generally undulates below the zero charge line indicating negative net charge., The curve for the wild type TSA complex (in green) is evenly distributed around zero. Clearly, the contrast between the two curves shows wild type GAP deals with the substrate net negative charge more effectively than does mutant GAP for the same substrate.

6.3.5 Test of Charge Balance Hypothesis in RhoA-RhoGAP using GTPγF

The CBH has been established for some examples of mutant phosphoryl transfer enzymes, such as K145A in βPGM and K219A in PGK (see **Section 1.9**). What is observed in these examples is the loss of one positive charge in the protein causes a complementary switch from MgF_3^- to MgF_2 or of AlF_4^- to AlF_3^0 in response. This shows the sensitivity of charge balance in the whole active site as reflected by the TSA complexes. However, it has not been tested in any of the enzymes we have studied when the *substrate* has one charge diminished. A small G protein, RhoA, was chosen as a model for this purpose because firstly, its hydrolysis reaction is relatively simple, involving only one substrate — GTP and water. Secondly, there is no phosphoenzyme intermediate, so the stability of a chemically-mutated phosphoenzyme does not need to be taken into account. Thirdly, a stable substrate analogue such as GTPγF is synthetically available and can be rigorously purified. Lastly, the hydrolytic mechanism for a small G protein has been well studied structurally but still there remains a strong conflict of opinion on the catalytic roles of some key residues, such as the arginine finger. Our purpose of this exploration is to test CBH in a biochemical manner based on an understanding of the TSAs for both wild type and mutant GAP proteins.

6.3.5.1 Design of the activity assay

The nucleoside triphosphate analogue GTPγF was introduced for this purpose. It has three phosphates accurately replicating the shape of GTP. However, putting a fluorine on the γ-phosphate in the place of an anionic oxygen means that GTPγF has only three negative charges at neutral pH (**Figure 6.17**). Providing no other effects influence catalysis, if the CBH is the dominant feature of this enzyme mechanism, then hydrolysis of GTPγF should occur in the absence of the positive charge from Arg 85 in RhoGAP.

Thus, the experiments are designed to test the hydrolytic activity of RhoA alone, RhoA-$RhoGAP_{R85A}$, and RhoA-$RhoGAP_{wt}$ with GTPγF as a substrate.

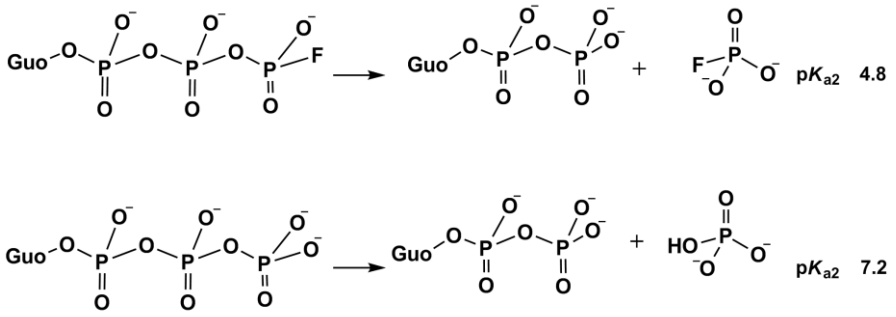


Figure 6.17 The proposed reaction of GTPγF hydrolysis.(Yoza *et al.*, 1994)

Since all the potential products are phosphorus-containing species, the reaction can easily be monitored by ^{31}P NMR. The reaction should use a low enzyme concentration with high substrate concentration so that the chance of observing immediate products increases and the NMR acquisition time can be shortened. To test the hydrolytic activity, several ways of producing nucleotide-free RhoA had to be considered, because GDP intrinsically binds very tightly to RhoA.

(**a**) EDTA could be added to the GST-protein bound on glutathione beads to chelate the catalytic Mg^{2+} and thus weaken GDP binding, followed by buffer exchange or dialysis to remove GDP.(Zhang *et al.*, 2000) This was explored but proved to be unachievable without destabilising the protein at rt and so voiding the NMR experiment. Even if successful, it would still call for a single turnover assay if GDP is the primary product.

(**b**) A much higher concentration of the GTP analogue could be used to out-compete GDP.(Ferguson *et al.*, 1986) This is not realistic as the binding constant for GTPγF can be expected to be lower than that for GTP, which itself is weaker that GDP, so the experiment will fail.

(**c**) Incubation with alkaline phosphatase (1U per mg G protein) at 4°C overnight in the presence of 200 μM ammonium sulphate, 0.1 μM zinc chloride should degrade the GDP followed by removing alkaline phosphatase by size exclusion chromatography.(Stumber *et al.*, 2002) This method can remove GDP but it still cannot avoid the activity assay becoming a single turnover test. Also it would require a very high concentration of protein to improve the signal to noise ratio of the NMR spectra to enable us to see the

product species. (It should be noted that our broadband probe is not very sensitive towards phosphorus signals due to technical reasons). So this method was also seen to be not suitable for our purpose.

(**d**) The best practical way to mount our experiment is therefore to use a guanine nucleotide exchange factor (GEF) specific for RhoA that can catalyse fast exchange between GDP and GTP.

The exchange activity being catalysed by GEFs involves the formation of an initial, low-affinity GEF-GTPase-GDP ternary complex that rapidly converts into a high-affinity GEF-GTPase binary complex concomitant with expulsion of GDP and Mg^{2+} (**Figure 6.18**).(Pan *et al.*, 1998; Cherfils *et al.*, 1999) The binary complex is stable in the absence of exogenous guanine nucleotides. However, the relatively high concentration of GTP *in vivo* (normally ~20-fold higher than GDP) favours the binding of GTP followed by dissociation of the GEF and activated GTPase.(Rossman *et al.*, 2002)

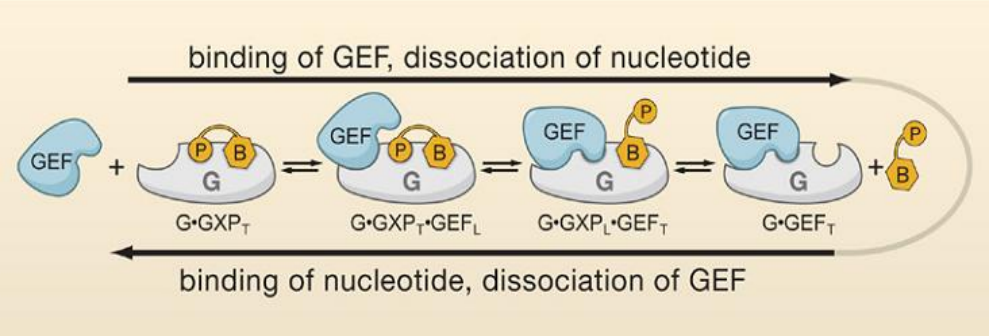


Figure 6.18 GEF catalysed exchange reaction occurs in successive reversible steps. The nucleotide (orange) interacts with the G protein (grey) *via* its base (B) and its phosphate moieties (P). The GEF (blue) competes with the nucleotide for binding with the G protein and thereby promotes nucleotide exchanges. The competition involves the existence of loose (subscript L) and tight (subscript T) interaction of the G protein with the nucleotide and the GEF.(Bos *et al.*, 2007)

6.3.5.2 Controls to test the activity of GEF

Because Dbs is a GEF for RhoA, the biological activity of Dbs can be determined from its ability to catalyse the exchange of GDP on Rho GTPase. Published exchange rate k_{obs} for RhoA is 0.16 ~ 0.31 $\times 10^{-3}$ s^{-1} (no Dbs), and 2.25 ~ 4.11$\times 10^{-3}$ s^{-1} (with wild type Dbs), when 1 μM RhoA is stimulated by 0.2 μM Dbs in the presence of 100 μM *N*-methylanthraniloyl GTP (mant-GTP).(Rossman *et al.*, 2002; Baumeister *et al.*, 2006) To

test the activity of the commercial Dbs, a control experiment was carried out by adding the natural substrate GTP to the RhoA/Dbs in buffer.

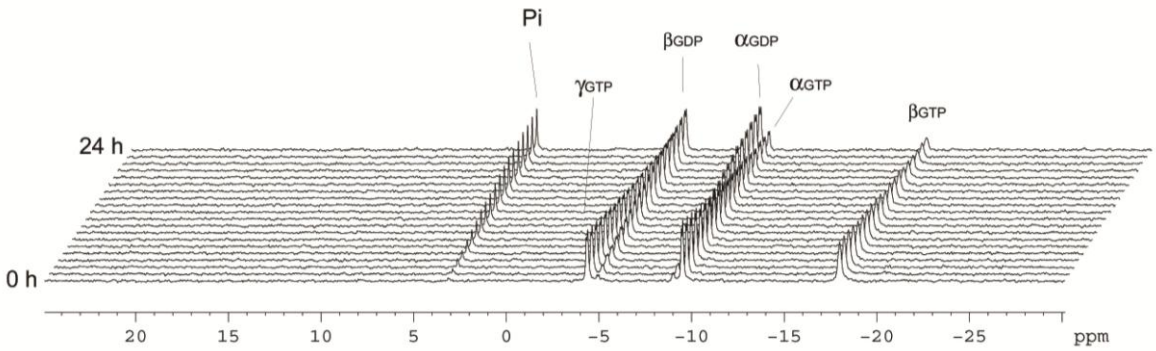


Figure 6.19 GTP hydrolysis assay as a control experiment to test the activity of purchased hDbs monitored by ^{31}P NMR. Conditions: 50 μM RhoA (with 50 μM GDP bound), 2.5 μM Dbs (RhoGEF), 5 mM GTP in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT. Each spectrum includes 3840 scans and takes 1h 21 min.

The ^{31}P NMR time course shows that reconstituted Dbs had adequate activity to exchange GDP with GTP and that the GTP was hydrolysed to GDP and P_i by RhoA at a good intrinsic rate, with a half life around 15 h (**Figure 6.19**). As the catalytic rate constant for the intrinsic RhoA (in the absence of GAP) as a GTPase is 1.4×10^{-4} s^{-1},(Graham *et al.*, 1999) it follows that in our experiment if the catalysed hydrolysis was at the same rate or slower for RhoA behaving as a 'GTPγFase', then the exchange process would not become the rate limiting step which had been the purpose in using Dbs as a GEF.

6.3.5.3 Activity assay using RhoA alone

The CBH was first evaluated by studying the hydrolysis of GTPγF catalysed solely by RhoA. To setup the sample, 50 μM RhoA (with 1 eq. GDP bound) and 2.5 μM Dbs (RhoGEF) in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT were premixed in a 5 mm NMR tube. 5 mM GTPγF was added to the sample and the acquisition of ^{31}P NMR data begun immediately after locking and shimming. Each spectrum had 3840 scans and took 1h 21 min.

The time course of the reaction monitored by ^{31}P NMR (**Figure 6.20**) shows RhoA itself has almost no activity to hydrolyse GTPγF into GDP and PO_3F^{2-} within 35 h. The only reaction evident is the fast hydrolysis of a small amount of GTP and formation of a little Pi. Even after the sample was kept at 4°C for a week, only some 10% of GTPγF was

cleaved to give P_i (data not shown). No apparent increase in GDP or GMP was observed. This means there is almost no activity of hydrolysis activity towards GTPγF. The total 3% GDP and GTP at the starting point were identified as impurities in the GTPγF sample. No ^{31}P NMR resonance for monofluorophosphate was observed. It is evident that GTPγF is stable to RhoA hydrolysis.

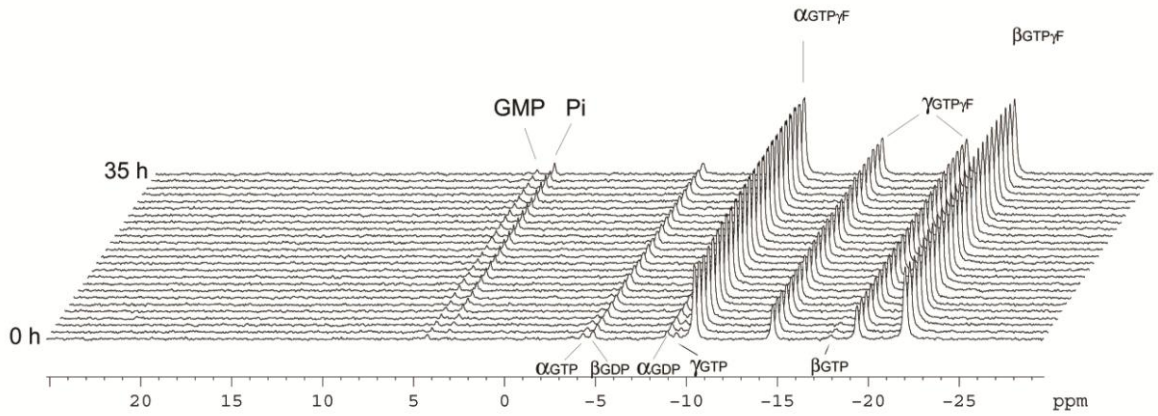


Figure 6.20 RhoA hydrolysis assay to evaluate CBH by GTPγF hydrolysis, monitored by ^{31}P NMR. Conditions: 50 μM RhoA (with 50 μM GDP bound), 2.5 μM Dbs, 5 mM GTPγF in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT. Each spectrum has 3840 scans and takes 1h 21 min.

Possible reasons for the failure to observe GTPγFase activity can be considered. First, in the absence of protein, the affinity of GTPγF to magnesium, that is essential for catalysis, is 0.20 mM^{-1} which is 140-fold smaller than for GTP (28 mM^{-1}).(Stumber *et al.*, 2002) Secondly, given the concentration of the protein in all the assays is very low (50 μM for RhoA), it is impossible to test the binding status of GTPγF to RhoA due to the low sensitivity of our ^{31}P NMR. However, the nonbridging oxygens on the γ-phosphate in GTPγF should be less negatively charged than in GTP both because of the electron withdrawing effect of fluorine and because its one negative charge is shared between the two γ-oxygens. (In GTP, two negative charges are shared between the three γ-oxygens). Therefore the ionic binding interaction to the protein and to the Mg^{2+} cation must be weakened. That will clearly reduce the competitiveness of GTPγF for binding to RhoA as a substrate.

6.3.5.4 Activity assay using RhoA plus $RhoGAP_{R85A}$

We have seen that in the catalysed hydrolysis of GTP, $RhoGAP_{wt}$ contributes not only a positively charged arginine finger interacting with the TS of the reaction to accelerate

the hydrolysis by 240-fold and also increases the intrinsic hydrolysis of the GTP by 160-fold by optimising the conformation of the GTP binding pocket to place water in-line with the breaking P–O bond of GTP. Thus, RhoGAP might facilitate the correct of the GTPγF substrate in the same way. Therefore, the $RhoGAP_{R85A}$ mutant was added to RhoA as a further test of the CBH in the hydrolysis of GTPγF.

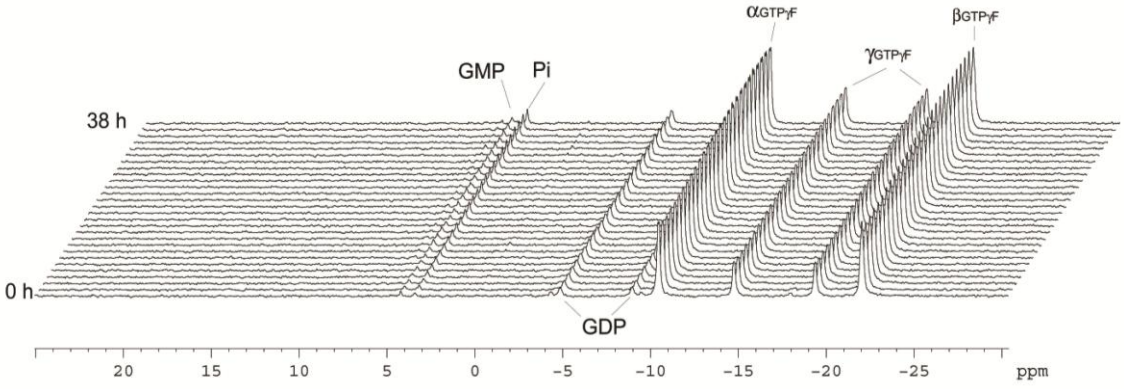


Figure 6.21 Hydrolysis assay to test the activity of RhoA-$RhoGAP_{R85A}$ on GTPγF monitored by ^{31}P NMR. Conditions: 50 μM RhoA (with 50 μM GDP bound), 2.5 μM RhoGEF, 5 mM GTPγF and 50 μM $RhoGAP_{R85A}$ in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT. Each spectrum has 3840 scans and takes 1h 21 min.

As for the setup in the previous assay, 50 μM RhoA (with 1 eq. GDP bound), 50 μM $RhoGAP_{R85A}$ and 2.5 μM Dbs in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT were premixed in a 5 mm NMR tube. 5 mM GTPγF was added to the sample and the acquisition of ^{31}P NMR data begun immediately after locking and shimming. Each spectrum had 3840 scans and took 1h 21 min. The ^{31}P NMR monitored time course shows a result not much different from that with RhoA alone; no activity is detected (**Figure 6.21**). This means $RhoGAP_{R85A}$ has almost no effect on the catalysis of GTPγF or even its binding and only gives a negative result.

6.3.5.5 Activity assay using RhoA-$RhoGAP_{wt}$

Since the first two searches for hydrolysis of GTPγF failed, we considered the possibility that RhoA-RhoGAP might deny charge balance for catalysis, as indirectly proved by TSAs in **Sections 6.3.3** and **6.3.4**. It could be that the conformational difference between RhoA, RhoA-$RhoGAP_{wt}$ and RhoA-$RhoGAP_{R85A}$ is so marked that charge balance might not be important for the RhoA catalysis. So $RhoGAP_{wt}$ was added to the sample, hoping that the binding of GTPγF could be better to wt protein and

thereby achieve the desired hydrolysis reaction, even though the addition of wild type RhoGAP would negate the original plan to explore further the CBH.

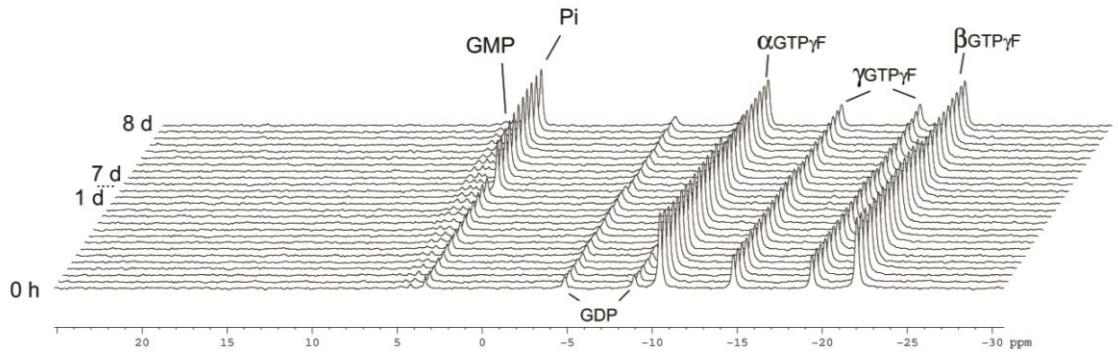


Figure 6.22 Hydrolysis assay to test the activity of RhoA-RhoGAP$_{wt}$ on GTPγF using ^{31}P NMR. Conditions: 50 μM RhoA (with 50 μM GDP bound), 2.5 μM RhoGEF, 5 mM GTPγF and 50 μM RhoGAP in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT. Each spectrum has 3840 scans and takes 1 h 21 min.

Disappointingly, the experimental results in the first 24 h were not different from the two assays in **Section 6.3.5.3** and **Section 6.3.5.4**. After 7 d, the spectrum of the same sample was acquired for a whole day, but the only reaction that was detected by ^{31}P NMR was the consumption of GTPγF and the production of inorganic phosphate (**Figure 6.22**). (NB because of the relaxation properties of Pi, the height and integration of peaks are not reliable enough to compare the concentration of different phosphorus species.)

For the time being, we cannot explain this unexpected side reaction. It does appear that the main result is one of intrinsic chemical inertness of GTPγF. Moreover, it is unlikely that the production of Pi results from the hydrolysis of monofluorophosphate (MFP), because monofluorophosphate is deprotonated at neutral pH to form PO_3F^{2-} and is very stable in the absence of catalysis. Its half-life at pH 6 to 8 at rt is more than one year. That is supported by the observation that, in a control experiment, only 5% of PO_3F^{2-} is hydrolysed within 3 d in the same buffer at the same pH with 50 μM RhoA present (**Figure 6.23**). However, a catalytic amount of alkaline phosphatase (EC 3.1.3.1) is known to accelerate the reaction by 20,000 fold.(Yoza *et al.*, 1994; Stumber *et al.*, 2002)

Yet again, the RhoA-RhoGAP$_{wt}$ catalysed hydrolysis of GTPγF did not lead to observable formation of PO_3F^{2-} by ^{31}P NMR. One possibility is that binding of GTPγF to RhoA has not been improved by RhoGAP because its affinity is intrinsically too

weak. Because wild type RhoGAP has evolved to optimise its active site for a TSA conformation with GTP, its response to the charge changes of GTPγF is unpredictable.

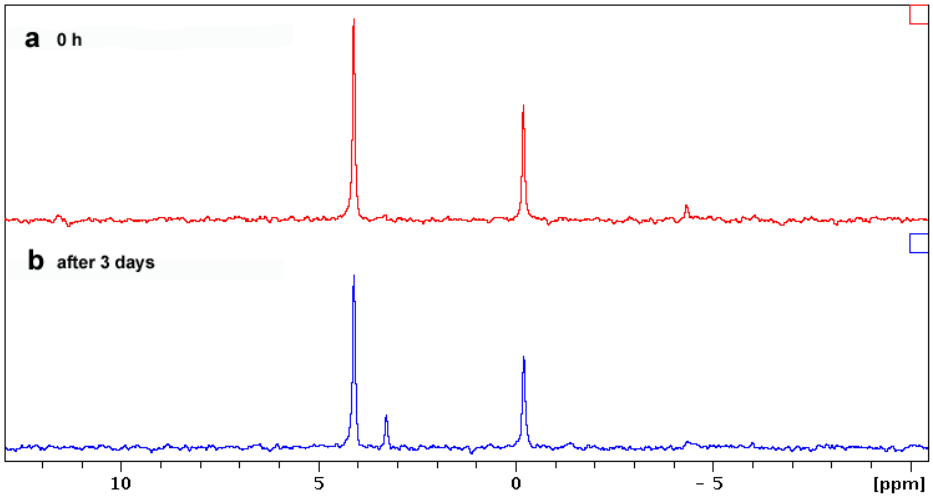


Figure 6.23 ^{31}P NMR spectra of Na_2PO_3F (5 mM) in buffer with 50 μM RhoA, 50 mM Tris-HCl, 150 mM NaCl, 5 mM $MgCl_2$, 1mM DTT, pH 7.5. (**a**) The initial sample showing the chemical shift of phosphorus at 2.1 ppm with $^1J_{PF}$ = 869 Hz (212 MHz). (**b**) The sample after incubation at 25°C for 72 h, the new peak at 3.3 ppm is assigned to P_i. Both spectra have 3840 scans and take 1 h 21 min.

It is also possible that GTPγF may bind to RhoA with such a low affinity that even if hydrolysis is marginally enhanced by RhoGAP, k_1 (**Figure 6.24A**) is still much smaller than exchange rate constant k_{off}, so the production of monofluorophosphate is very slow. If there is very little monofluorophosphate in solution, it can be turned into Pi and F^- at the control rate and it would not be detectable by NMR.

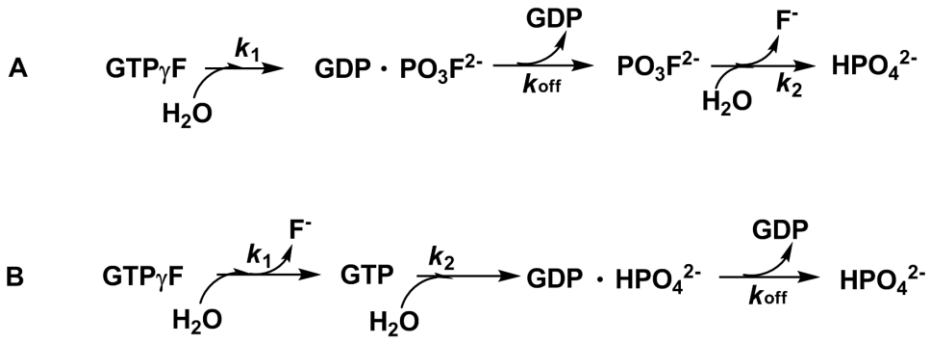


Figure 6.24 Possible reactions catalysed by RhoA-RhoGAP$_{wt}$ for an experiment set up as in Figure 6.22.

A third possibility is that reaction takes place as shown in **Figure 6.24B**. GTPγF is slowly hydrolysed to GTP by water with rate constant k_1, then normal hydrolysis of GTP takes place catalysed by RhoA-RhoGAP. It is inevitable that k_2 (5.4 s^{-1}) must be

much larger than k_1 and k_{off}, so the rate limiting step would be k_1. A previously reported hydrolysis of Ras GTPase using GTPγF as substrate with $k_1 = 4.7 \times 10^{-4}$ s^{-1} (Stumber *et al.*, 2002) is not persuasive because the activity could have been from a contaminating enzyme, such as alkaline phosphatase, and not from GTPase itself.

Finally, it has to be considered that the migrating species for cleavage of GTPγF would be a neutral trigonal PO_2F group rather than the anionic PO_3^- moiety. The effect of loss of negative charge will surely contract the molecular orbitals involved with axial O–P–O bonding in the transition state. While that cannot be quantified here, it could be a challenge for a theoretical chemist!

6.3.5.6 No hydrolytic activity has been found for GTPγF

From these endeavours, we have to draw the conclusion that the CBH cannot be tested by using GTPγF as substrate as all assays failed to show any hydrolysis of GTPγF. This can be attributed to various reasons, including weak binding to the active site, low magnesium affinity because of the consequences of fluorine substitution, and inherent chemical inertness. The large conformational change induced by the RhoGAP binding and mutation also increases the uncertainty of the GTPγF hydrolysis.

6.4 General Conclusions

Our NMR results on RhoA-GDP-MgF_3^--$RhoGAP_{wt}$ TSA complex confirm the PIXE analysis in the solid state (see **Section 6.3.1**). The two wt TSA complexes RhoA-GDP-MgF_3^--RhoGAP and RhoA-GDP-AlF_4^--RhoGAP both obey the CBH (see **Section 6.3.2**) The concept of charge balance in small G proteins has been developed by using RhoA alone to form a stable RhoA-GDP-AlF_3 complex in solution. Our data suggests the nature of AlF_3 in this complex is more likely to be a octahedral $AlF_3(H_2O)^0$ than either a tbp AlF_3^0 or an octahedral $AlF_3(OH)^-$ (see **Section 6.3.3.2**). Special features of the active site of $RhoGAP_{R85A}$ are revealed by symbiotic use of ^{19}F NMR and crystallography. Both the RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ and RhoA-GDP-AlF_4^--$RhoGAP_{R85A}$ TSA complexes maintain negative charge on the metal fluoride moieties using Tyr 34 coodination (see **Section 6.3.4.1** and **Section 6.3.4.2)**. These results show the CB concept in small G proteins differs from that in phosphomutase βPGM and small molecule kinase HsPGK. Finally, a substrate analogue GTPγF proved stable to any hydrolytic acitivity by RhoA alone or with and without a contribution from the arginine

finger in RhoGAP (see **Section 6.3.5**). These results suggest that the CB situation may be more complex for a small G protein than for other systems studied, although the negative result may simply be a manifestation of the intrinsic unreactivity of GTPγF.

Chapter 7. UMP/CMP Kinase (UMPK)

7.1 Introduction

Nucleoside monophosphate kinases (ATP-NMP phosphotransferases) catalyse the reversible transphosphorylation between nucleoside triphosphates and nucleoside monophosphates. Adenylate kinase (ATP-AMP phosphotransferase, EC 2.7.4.3) is the best known member of this class since its discovery.(Kalckar, 1943)

UMP/CMP kinase (EC 2.7.4.14) also operates on nucleoside monophosphate and catalyses the reversible reaction transferring γ-phosphate of ATP to UMP or CMP. Its cDNA from *Dictyostelium. discoideum* was first subcloned into an *E. coli.* expression vector pIMS5–cDUK–1 or pIMS6–cDUK–1 by Wiesmuller *et al.* in 1989 and its molecular and kinetic properties were discussed in comparison to the characteristics of adenylate kinase in the same family.(Wiesmuller *et al.*, 1995) As for adenylate kinases, the UMP/CMP kinase from *D. discoideum* (shortened to $UMPK_{dicty}$ has a broad specificity for nucleoside triphosphates but a more narrow specificity for nucleoside monophosphates. Having 30 ~ 43% sequence identity with prokaryotic and eukaryotic mitochondrial and cytoplasmic adenylate kinase shows the fact that UMP/CMP kinase may have promiscuous phosphorylation activity as an adenylate kinase.

7.1.1 General description of the structure

$UMPK_{dicty}$ has 194 amino acid residues and 2 cysteines in positions 23 and 119. Cys 23 is critical and plays a structural role in favouring an active conformation. We found this in our experiments when we tried to make a TSA complex. Without DTT in the buffer, the TSA complex can't form which indicates that the conformation was not the active one.

All the proteins in the adenylate kinase family have three domains: CORE, NMP, and LID domains (**Figure 7.1**).(Vonrhein *et al.*, 1995) The rigid CORE region includes the central parallel β-sheet and the P-loop that plays an important role in binding the triphosphate moiety of ATP. There are very few contacts with the adenine base of ATP which is also very common in other NMP kinases and this is consistent with the observed low specificity toward ATP. Apart from the exocyclic amino group at C6 that

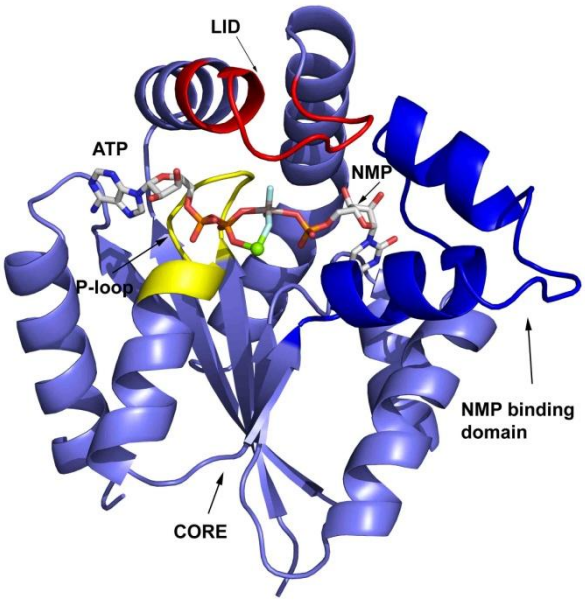


Figure 7.1 Three dimensional cartoon structure of UMP/CMP kinase complex with its TSA(PDB: 3UKD). Catalytic magnesium is a green sphere, AlF_3 is grey and light blue sticks, ADP and NMP are white sticks. The NMP binding domain (residues 36–65) is in blue, LID domain (residues 130-140) is in red, and P-loop (residues 13-21) is in yellow.

forms a 2.9 Å hydrogen bond to backbone carbonyl oxygen of Arg 176, the only other purine interaction is a typical aromatic-amino group stacking between the Arg 127 side chain and the adenosine ring.(Burley *et al.*, 1986) Lys 19 in the highly conserved P-loop sequence(Saraste *et al.*, 1990) is also thought to be an important positive charge donor in the transition state.(Reinstein *et al.*, 1990; Tian *et al.*, 1990; Byeon *et al.*, 1995) Also in structure PDB:1UKE in which Up_5A is the bisubstrate analogue, the amino group of this Lys 19 interacts with P2 and P4 strongly.(Scheffzek *et al.*, 1996)

The nucleoside monophosphate is bound in the NMP region. After the substrates are bound, NMP domain closes upon the NMP moiety of the bound dinucleotide and the LID domain closes over the negatively charged phosphate groups of the bound nucleotides with only highly conserved, positively charged arginines (Arg 137 and Arg 131 in $UMPK_{dicty}$). The high flexibility of these two functional domains is indicated by the *B*-factor.(Scheffzek *et al.*, 1996) Asp 139 and Asp 140 are believed to be important for stabilising the TS but not for the substrate-induced conformational changes.(Dahnke *et al.*, 1994) Arg 42 and Arg 93 interact with the phosphate of UMP/CMP to enhance the binding of nucleoside monophosphate. Especially the highly conserved Arg 93 also interacts with γ-phosphate of ATP in the TS.

Experimental evidence shows that the presence of Mg^{2+} is critical for the phosphoryl transfer process but not for the binding of ATP or ADP at the donor nucleotide binding site (Reinstein *et al.*, 1990) because Mg^{2+} is more strongly bound in the TS for the enzyme reaction than in the TS for the uncatalysed reaction.(Stockbridge *et al.*, 2009) $UMPK_{dicty}$ seems to be an extreme case for kinases since crystal structures show that the active site residues of $UMPK_{dicty}$ make no contact with the metal ion **(Figure 7.2)**. The requirement for Mg^{2+} can be considered to arise entirely from interactions between the metal ion, the substrate and the water ligands.(Scheffzek *et al.*, 1996; Stockbridge *et al.*, 2009)

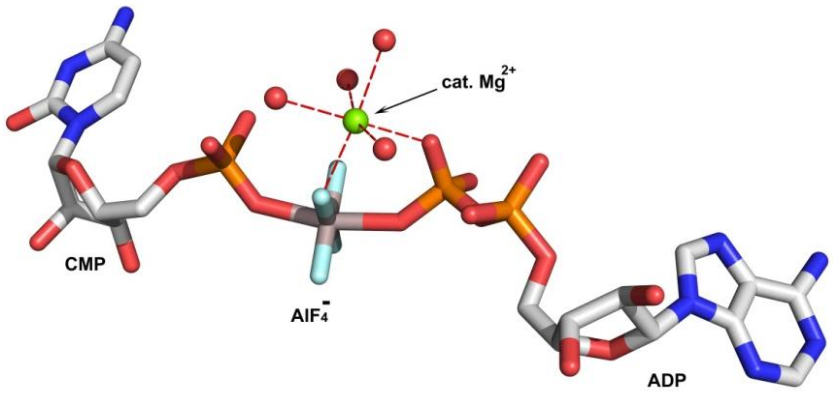


Figure 7.2 Structure of UMPK-ADP-AlF_4^--CMP TSA complex (PDB: 1QF9) showing catalytic magnesium (green sphere) is required for substrate interactions. AlF_4^- is grey and light blue sticks, and four water molecules are in red spheres.

7.1.2 Background

Over the last two decades, the $UMPK_{dicty}$ has been studied intensively to investigate phosphoryl transfer mechanisms by X-ray crystallography and computation, because of its stability over a wide range of pH and its structural similarity to adenylate kinase. (Wiesmuller *et al.*, 1995; Scheffzek *et al.*, 1996; Schlichting *et al.*, 1997; Schlichting *et al.*, 1999; Hutter *et al.*, 2000) Even though there have been several crystal structures of $UMPK_{dicty}$, and even more for adenylate kinase, the actual reaction mechanism of phosphoryl transfer is still unclear. The main discussion has focused on whether the reaction is associative or dissociative. If the reaction is fully dissociative, the γ-phosphate breaks away from ATP to form a monoanionic metaphosphate (PO_3^-) in the TS which is then captured by the α-phosphate of the mononucleoside. This mechanism is valid if the distances between the donor oxygen and acceptor oxygen in the TS are larger than 6.6 Å.(Mildvan, 1997) In contrast, for a fully associative mechanism the pentaoxyphosphorane group in the TS is highly negatively charged and the distances between its two apical oxygens should be around 3.4 Å corresponding to the two covalent single axial P–O bond lengths for a tbp.(Jones, *et al.*, 2002)

There are two structures in the PDB: 3UKD and 5UKD for $UMPK_{dicty}$-ADP-AlF_3^0-CMP complexes both crystallised at pH 8.5 and with the tbp TSA complex identified as AlF_3^0. In 5UKD, the distances between the central metal and the axial phosphate oxygens are 2.2 and 2.3 Å, and the three equatorial metal fluoride bond lengths are 1.7, 1.7, and 1.8 Å. But in 3UKD, the distances between metal and the axial phosphate oxygens are 2.0 and 2.2 Å, and the three equatorial metal-fluoride bond lengths are 1.6, 1.7, and 1.7 Å (**Figure 7.3**).(Schlichting *et al.*, 1997; Schlichting *et al.*, 1999) Both of these suggest an associative mechanism, although there may be an underlying reason to cause the difference between the axial O–Al–O and equatorial Al–F bond lengths in the two structures which will be discussed later (**Section 7.3.1**). In another structure PDB: 1UKE, a bisubstrate analogue P^1-(5'-adenosyl)-P^5-(5'-uridyl)-pentaphosphate (Up_5A) binds at the active site to form a $UMPK_{dicty}$-Up_5A complex, in which the distance between the β-bridging oxygen on adenosine end and α-oxygen on uridine end is 4.5 Å.

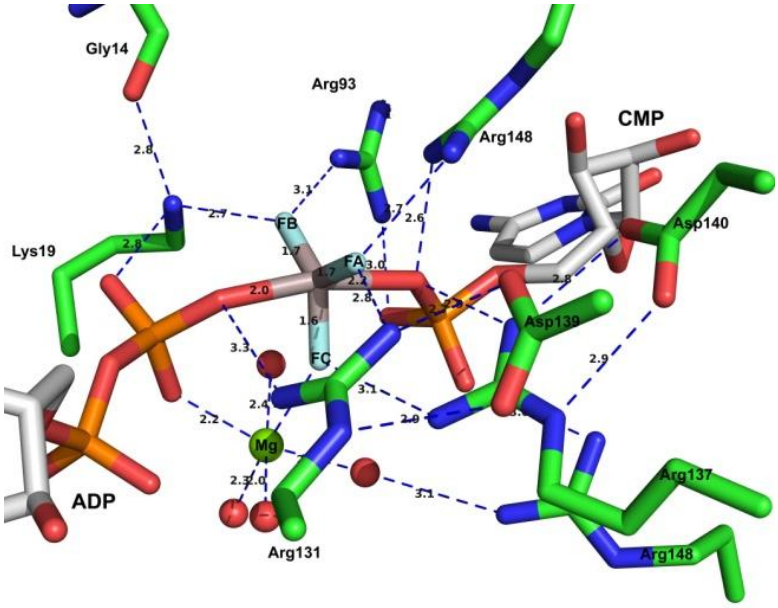


Figure 7.3 Crystal structure of UMPK-ADP-AlF_3^0-CMP TSA complexes (PDB: 3UKD). Complex crystallised from a solution of 0.64 mM $UmpK_{dicty}$, 0.8 mM ADP, 0.8 mM CMP, 100 mM Tris-HCl (pH 8.5), 30 mM $MgCl_2$, and 40 mM DTE in the presence of 0.8 mM $AlCl_3$ and 4 mM NaF using PEG 4000 as precipitant.(Schlichting *et al.*, 1999).

This structure also indicates an associative mechanism.(Scheffzek *et al.*, 1996) But the computational result based on the 'AlF_3^0' shows it is rather more dissociative in the phosphoryl transfer, being accompanied by the synchronous shift of a proton from CMP to the transferred PO_3^- group.(Hutter *et al.*, 2000) Interestingly, the same calculated

result based on the original crystal structure of 'cAPK-ADP-AlF_3^0-SP20' complex by the same group also suggested the mechanism is dissociative even though the distances between the two apical oxygens (PDB: 1L3R) is 4.5 Å in the TSA structure.(Hutter *et al.*, 1999) (NB after refinement by Dr. Matthew Bowler, ESRF, Grenoble, this distance is even shorter at 4.3 Å.)

In addition to the controversial mechanism discussed for this enzyme, the nature of the TSA complex geometry for crystals formed at different pHs is also a puzzle. It has been claimed that for most cases, the bound metal fluoride occurs as a tbp AlF_3^0 when the pH of crystallisation is above 7.5 but as an octahedral AlF_4^- when the crystallisation pH is lower than 7.5 (**Figure 7.4**). Only in two cases, the tbp complexes were formed at low pH, but these had a fairly low Al^{3+} concentration in the sample.(Schlichting *et al.*, 1999)

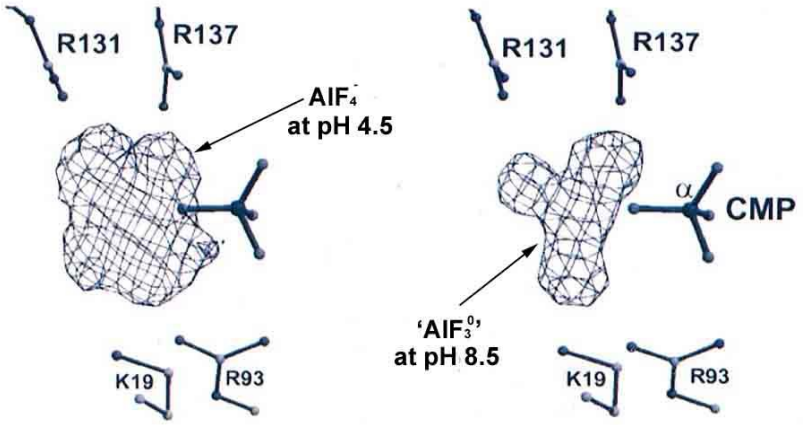


Figure 7.4 Residual difference electron density ($F_{obs} - F_{calc}$, 3σ cut off) as observed before placing the models of (left) AlF_4^- in the square planar electron density for the crystal grown at pH 4.5 or of (right) 'AlF_3^0' in the trigonal planar electron density for the crystal grown at pH 8.5.(Schlichting *et al.*, 1999)

7.1.3 Purpose of this study on UMP/CMP kinase

According to our experimental results, the tbp TSA metal fluoride moieties seen in PSP, βPGM, human PGK, as well as for the structures of the small G protein RhoA and cAPK have all been proved to be MgF_3^- by ^{19}F NMR in work in this thesis (**Section 1.8**). While it is not realistic for us to test all the phosphoryl transfer enzymes in the PDB list (**Table 1.3**), at least we can analyse a key kinase, $UMPK_{dicty}$, itself the paradigm for AlF_3^0 TSA complexes, in order to verify our general analysis that many, perhaps most, if not all of the tbp TSA metal complexes assigned in the PDB are in fact MgF_3^- and not AlF_3^0. It will be useful and timely to correct errors in the literature that appear to have

misinterpreted most of the complex structures of this same nature.(Sudom *et al.*, 2001; Miles *et al.*, 2002; Li, 2003) The charge on the phosphoryl group in transfer is a key feature to identify whether the reaction is more associative or dissociative. It is also important for the computational calculations to have the right model based on the correct interpretation of the X-ray data. We note that modifications to the protonation status of substrates and ionised catalytic residues cannot be ruled out as the local enzymatic environment can influence their effective pK_a a great deal.(Warshel *et al.*, 1991) Therefore, we believe the charge of the metal fluoride mimicking the phosphate in flight is not simply a consequence of the pH of crystallisation of the sample, which was attributed to the state of ionisation of the phosphoryl group(s) in the substrate(s).(Schlichting *et* al.,1999)

In this chapter, both MgF_3^- and AlF_4^- TSA complexes have been tested over a wide range of pH to show that the MgF_3^- is stable at all pHs from 6 to >9 while AlF_4^- is stable only at pH below 8.5. A new intermediate 'AlF_3' species assigned as $AlF_3(OH)^-$ has been discovered both in UMPK-ADP-$AlF_3(OH)^-$-UMP and UMPK-AMPCP-$AlF_3(OH)^-$-UMP complexes in the course of the pH titration and its nature analysed. The formation of UMPK-ADP-BeF_2^0-UDP TSA complex explains residual adenylate kinase activity in $UMPK_{dicty}$. The CBH has also been validated because a neutral AlF_3^0 does not form in UMPK-ADP-AlF_3^0-UDP complex. Attempts to make TSA complexes using fluorine-containing analogues of UMP and ADP nucleotides as ^{19}F NMR standards for internal integration reference all failed.

7.2 Materials and Experiments

7.2.1 Cell growth and expression

A plasmid of $UMPK_{dicty}$ was generously given by Prof. Lisa Miesmüller in Department of Obstetrics and Gynaecology of the University of Ulm. The cDNA-derived sequence was confirmed by DNA sequencing as follows:

```
        10         20         30         40         50         60
MMEKSKPNVV FVLGGPGSGK GTQCANIVRD FGWVHLSAGD LLRQEQQSGS KDGEMIATMI

        70         80         90        100        110        120
KNGEIVPSIV TVKLLKNAID ANQGKNFLVD GFPRNEENNN SWEENMKDFV DTKFVLFFDC

       130        140        150        160        170        180
PEEVMTQRLL KRGESSGRSD DNIESIKKRF NTFNVQTKLV IDHYNKFDKV KIIPANRDVN
```

```
      190
EVYNDVENLF KSMGF
```

Number of amino acids is 195 and the molecular weight is 22074.0. The theoretic pI is 6.5. Extinction coefficient is 13,980 M^{-1} cm^{-1} at 280 nm in water.

The protein was proposed to be expressed in *E. coli.* strain XL1-Blue. However, *E. coli.* strain BL21-DE3-Rosetta-pLys was also tested as an expression strain to improve the yield but was not successful. Competent cells were transformed with plasmid p5–WT4 and then plated out on LB agar with 100 μg/ml ampicillin as the antibiotic. 3 × 50 mL starter LB media and 700 mL LB media in 8 × 2L flasks were made up for autoclaving. To each of the three 50 mL flasks was added the proper amount of ampicillin and a selected colony to make starter culture and incubated overnight (12~16 h) at 37°C, 250 rpm. Ampicillin was added to 8 × 2 L flasks which were then inoculated with 10 mL overnight culture each. The cells were grown at 37°C, 250 rpm until OD_{600} reached ~0.6. The expression was induced with 1 mM IPTG and the culture was incubate for a further 5~6 h at 25°C, 250 rpm. The cells were harvested at 4°C by centrifugation at 4200 rpm for 20 min to yield a cell pellet with wet weight of 20~30 g and the pellet was stored at –80°C.

7.2.2 Protein purification

Table 7.1 Protocol of buffer for $UMPK_{dicty}$ purification.

Binding buffer for Blue-Sepharose column:	50 mM	Tris-HCl, pH 7.5
	1 mM	EDTA
	5 mM	DTT (freshly added)
	1 mM	NaN_3
Elution buffer for Blue-Sepharose column:	50 mM	Tris-HCl, pH 7.5
	1 mM	EDTA
	5 mM	DTT (freshly added)
	1 M	NaCl
	1 mM	NaN_3
S buffer for Gel Filtration column:	50 mM	Tris-HCl, pH 7.4
	5 mM	$MgCl_2$
	200 mM	NaCl
	5 mM	DTT (freshly added)
	1 mM	NaN_3

The pellet was resuspended in 40 mL binding buffer with supplement of a tablet of Complete Cocktail Protease Inhibitor. The thawed cells suspension was then lysed by sonication on ice for 4 × 20 s with 1 min gap and then centrifuged at 4°C for 40 min at 24,500 rpm. The supernatant was filtered through 0.22 μm filter syringe before being loaded onto a Blue-Sepharose column (1.5 × 30 cm) at a flow rate of 2 mL/min. The

column was further washed with binding buffer until the baseline was flat. Then a linear gradient of NaCl was applied to the column, from 0 M to 1 M over the volume of 300 mL at a flow rate of 3 mL/min. The $UMPK_{dicty}$ should be eluted after NaCl concentration reached 0.6 M. A SDS-PAGE was run to check the existence of desired protein in the fractions and the fractions with $UMPK_{dicty}$ (MW 22,000) were poured

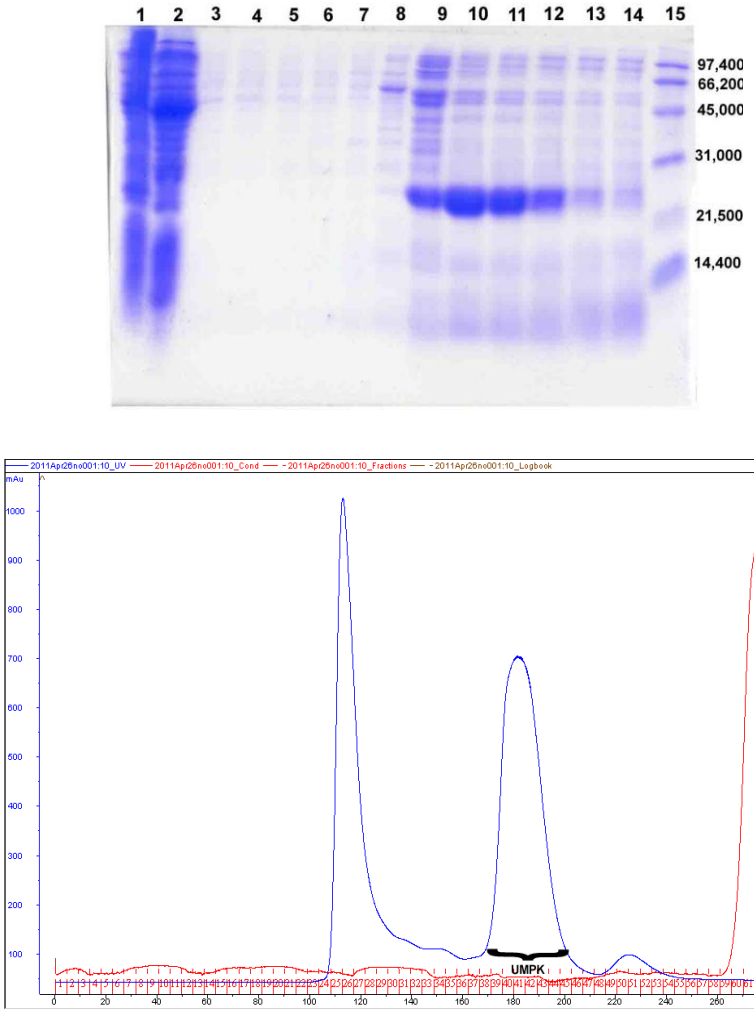


Figure 7.5 Elution of $UMPK_{dicty}$ during affinity chromatography shown on SDS-PAGE. Lane1 – lysate; lane2 – flow through after Blue-Sepharose column; lanes 3-14– fractions of the elution; Lane15 – low range marker. The UV trace is for gel filtration chromatography in the second step.

together. The preliminarily purified protein solution was concentrated to 4 mL and loaded on a pre-equilibrated (with S buffer) Superdex^{TM} G75 Gel Filtration column (GE Healthcare) being followed by the elution at a flow rate 1.5 mL/min with S buffer. A SDS-PAGE was run to check the purity and then the purified protein was concentrated to 0.5 mM using Sartorius Vivaspin^{TM}. The Sartorius Vivaspin^{TM} were rinsed thoroughly with water or buffer prior to addition of the protein sample because the trace

amounts of glycerol on the Sartorius Vivaspin^{TM} membrane may cause protein aggregation. The $UMPK_{dicty}$ exists in the solution as a monomer.

Remarks. The presence of DTT in the buffer is essential for $UMPK_{dicty}$ because it is only active when its single pair of cystines are reduced. The purified protein aggregates after being frozen, therefore it was kept at 4°C with 200 mM NaCl included in the final solution, which improves the $UMPK_{dicty}$ solubility dramatically, but still it can't stop aggregation. So the protein had to be used within 1 month after it was prepared.

7.2.3 NMR sample preparation

pH measurement was carried out in an indirect way, due to unsuitability of the semi-micro, combination, single junction pH electrode (Cat. No.CMAW711/3.7/180, Russell, Thermo) for this sample because it has a high concentration of protein in a Tris buffer.

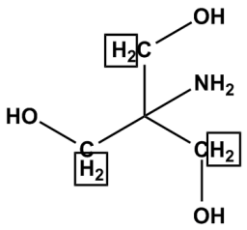


Figure 7.6 Chemical structure of Tris. The ^{1}H chemical shifts of protons in the square were measured.

Tris can be used as a pH indicator because of the ^{1}H NMR chemical shifts of the protons on the three –C**H**$_2$OH give a singlet resonance around 3.5 ppm that is sensitive to pH within the region of 6.7~9.8 (corresponding to the δ 3.725~3.516 ppm, respectively).(Baryshnikova *et al.*, 2008) Therefore, the pH of the sample can be calculated by the equation:

$$pH = pK_a - \log(\delta_{obs} - \delta_{HA})/(\delta_A - \delta_{obs})$$

where δ_{HA} and δ_A are the chemical shifts for protonated and deprotonated forms correspondingly, and δ_{obs} is the ^{1}H chemical shift observed for the solution of unknown pH. Here, the pK_a of Tris is 8.23 ± 0.08, δ_{HA} = 3.732, δ_A = 3.509.

Table 7.2 Comparison between the calculated pH according to the ^{1}H chemical shift and the measured pH. The measured pH value was averaged value from three measurements each time with an error of about ±0.3 pH unit

Experiment name	δ_{obs}	Calc. pH	Measured pH
uckinase_r 2000	3.721	6.95	6.9
uckinase_r 2002	3.730	6.23	5.8
uckinase_r 2004	3.729	6.36	5.7
uckinase_r 2006	3.723	6.85	6.8
uckinase_r 2008	3.715	7.15	7.2
uckinase_r 2010	3.671	7.81	7.6
uckinase_r 2012	3.652	7.98	8.0
uckinase_r 2014	3.605	8.35	8.4
uckinase_r 2016	3.569	8.66	8.8
uckinase_r 2018	3.551	8.87	8.8
uckinase_r 2020	3.541	9.01	Not measured
uckinase_r 2022	3.532	9.18	Not measured
uckinase_r 2024	3.526	9.32	Not measured
uckinase_r 2026	3.518	9.61	Not measured

7.3 Results and Discussion

7.3.1 UMPK-UMP-MgF_3^--ADP TSA and UMPK-UMP-AlF_4^--ADP TSA complexes

The analysis of the MgF_3^- TSA complex in ^{19}F NMR was setup in 50mM Tris-HCl (pH 6.9), 200 mM NaCl, 50 mM DTT, 15 mM $MgCl_2$, with 0.5 mM $UMPK_{dicty}$, 2 mM UMP, 2 mM ADP, 10 mM NH_4F, and 0.5 mM DFO. The top NMR spectrum (**Figure 7.7a**) shows that at pH 6.9 when there is no Al^{3+} present, there are three peaks (–142.1, –147.1, and –168.4 ppm) that correspond to the three fluorides in UMPK-UMP-MgF_3^--ADP. At the same pH, after Al^{3+} was added, the ^{19}F NMR spectrum shows only one rotationally averaged signal at –139.6 ppm integrating for four fluorines in total (**Figure 7.7b**).

We made this UMPK-UMP-MgF_3^--ADP TSA complex to establish that the tbp species in the so-called UMPK-UMP-AlF_3^0-ADP TSA complex structure is indeed MgF_3^-. As mentioned in **Section 7.1.2**, there are two UMPK-UMP-AlF_3^0-ADP complex structures in the PDB obtained under the same conditions. In 5UKD, the distances between aluminium and the axial phosphate oxygens are 2.2 and 2.3 Å, and the three equatorial Al–F bond lengths are 1.7, 1.7, and 1.8 Å (**Figure 7.8**). But in 3UKD, the axial O–Al–O distances and equatorial Al–F bond lengths are 0.1~0.2 Å shorter (**Figure 7.3**). When these two crystal structures are compared side by side, our analysis is that the TSA

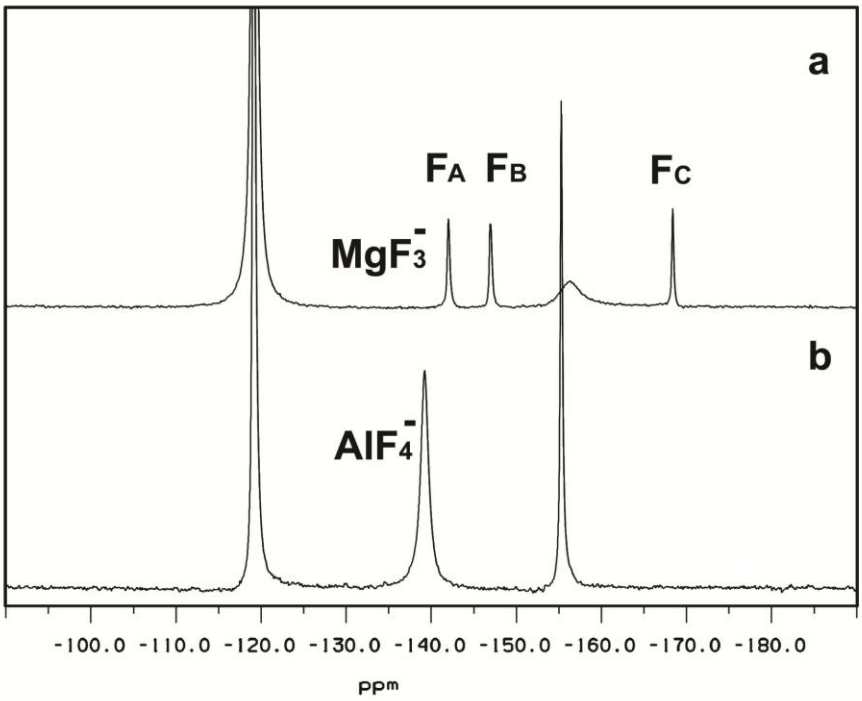


Figure 7.7 ^{19}F NMR spectra of (**a**) UMPK-UMP-MgF_3^--ADP TSA complex in 50mM Tris-HCl, pH 6.9, 200 mM NaCl, 50 mM DTT, 15 mM $MgCl_2$, 2 mM UMP, 2 mM ADP, 10 mM NH_4F, 0.5 mM DFO; The protein-bound MgF_3^- TSA signals are F_A = –142.0 ppm, F_B = –146.9 ppm, and F_C = –168.3 ppm; and **(b)** UMPK-UMP-AlF_4^--ADP TSA complex, recorded in the same conditions as (**a**) with addition of 2 mM $AlCl_3$. The protein-bound AlF_4^- TSA signal is rotationally averaged at –139.0 ppm.

structure 5UKD with longer bond lengths is more similar to the real MgF_3^- TSA structure. In other existing MgF_3^- TSA structures the axial O–Mg distances are around 2.1 ~ 2.2 Å and the Mg–F bond length is 1.8 ~ 1.9 Å, as in PSP and RhoA-RhoGAP. The compromised equatorial bond length observed in 5UKD may have been deliberately preset in the data analysis at 1.7 Å, as we found in the case of βPGM.(Baxter *et al.*, 2010) Alternatively, it may be another case of double occupancy which makes the equatorial bond length more ambiguous as in cAPK and PSP. In 5UKD, F_A forms two short H-bonds with Arg 148 and Arg 131 at 2.5 and 2.9 Å, respectively. F_B forms two H-bonds with Lys 19 and Arg 93 at 2.8 and 3.0 Å, respectively. These good H-bonds sensibly induce the big SIISs for F_A and F_B at 1.5 and 1.4 ppm respectively (**Table 7.3**). By comparison, F_C has a small SIIS at 0.5 Å as it only coordinates to the catalytic magnesium and Arg 137 2.8 Å away. The chemical shifts of all three resonances are similar to those of other MgF_3^- TSA complexes. Therefore, our ^{19}F NMR data of the UMPK-ADP-MgF_3^--UMP TSA complex fully supports the reinterpretation of the tbp AlF_3^0 in PDB: 5UKD structure as MgF_3^-.

Table 7.3 ^{19}F chemical shift (ppm) and SIIS (ppm) of the UMPK-ADP-MgF_3^--UMP, UMPK-ADP-AlF_4^--UMP TSA complexes.

Metal Fluoride Complex		δF (D_2O capillary)	ΔδF (D_2O capillary-100% D_2O)
UMPK-ADP-MgF_3^--UMP	F_A (F_1)	–142.0	1.5
	F_B (F_2)	–146.9	1.4
	F_C (F_3)	–168.4	0.5
	average	–152.4	—
UMPK-ADP-AlF_4^--UMP	F^*	–139.0	0.8

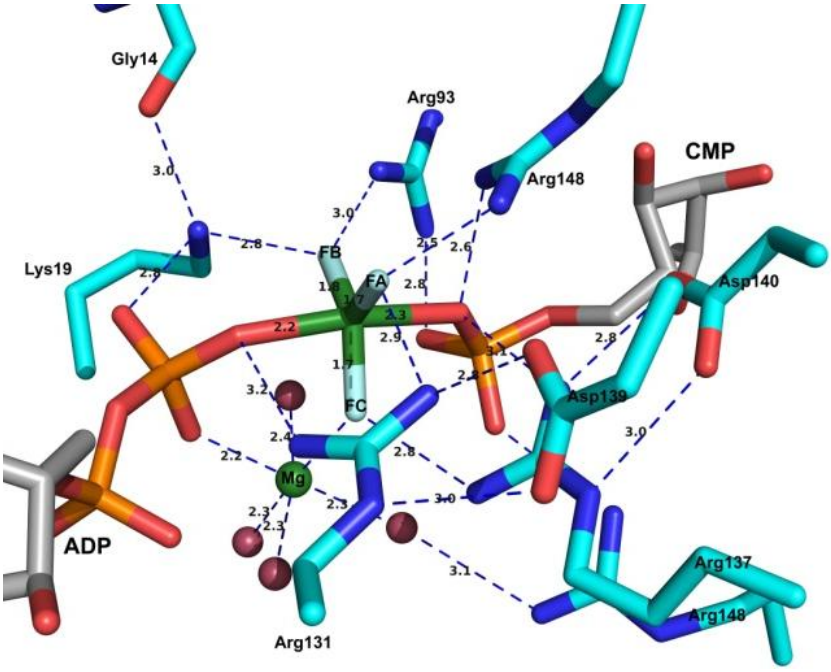


Figure 7.8 Crystal structure of UMPK-ADP-AlF_3^0-CMP TSA complex (PDB: 5UKD). Complex crystallised from 0.64 mM $UmpK_{dicty}$, 0.8 mM ADP, 0.8 mM CMP, 100 mM Tris-HCl (pH 8.5), 30 mM $MgCl_2$, and 40 mM DTE in the presence of 0.8 mM $AlCl_3$ and 4 mM NaF using PEG 4000 as precipitant.(Schlichting *et al.*, 1997)

7.3.2 pH Titration of UMPK-UMP-AlF_4^--ADP TSA complex

Our original purpose of the pH NMR titrations of UMPK-UMP-AlF_4^--ADP TSA complex was to test the validity of the 'priority of anionic charge over geometry' within the full range of pH values for published crystallisation conditions.(Schlichting *et al.*, 1997; Schlichting *et al.*, 1999) The transition between the AlF_4^- and MgF_3^- TSA complexes but maintaining the negative charge can be expected (**Figure 7.9**).(Baxter *et al.*, 2008)

In expectation of the instability of aluminium complexes at high pH, we chose to titrate from low to high pH. The results are given for 12 pH values from 6.23 to 9.61 and for both ^{19}F NMR (**Figure 7.10**, left) and downfield ^{1}H NMR spectra (**Figure 7.10**, right). From pH 6.23 to 7.15, the fluorine spectra show a single resonance, characterised above

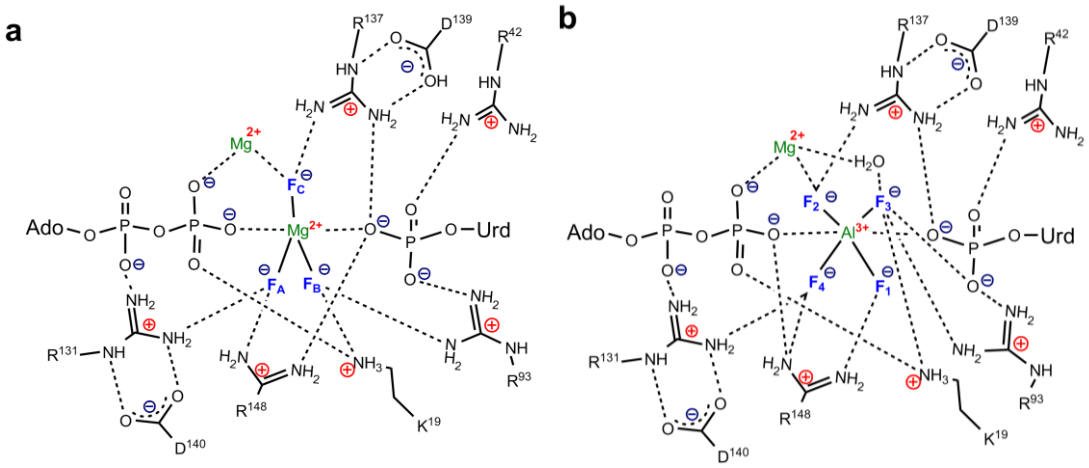


Figure 7.9 Interactions in the active centre of $UMPK_{dicty}$ with (**a**) ADP, UMP and MgF_3^- derived from PDB: 3UKD; and (**b**), ADP, UMP and AlF_4^- derived from PDB: 1QF9.

as that of a tetrafluoroaluminate complex (**Figure 7.10 b-e**). This signal has completely disappeared by pH 9.61 while, starting at pH 7.81, a group of three downfield resonances grows in and then also disappears at pH 9.61 (**Figure 7.10 f-l**). Finally, from pH 8.87 upwards, three new major resonances appear at chemical shifts characteristic of a trifluoromagnesate complex (**Figure 7.10 i-m**). That assignment was confirmed by comparison with the aluminium-free spectrum of the trifluoromagnesate complex at pH 6.9 (**Figure 7.10 a**). It was not possible to obtain spectra above pH 9.6 because the protein precipitated (data not shown).

The corresponding ^{1}H NMR spectra are more difficult to interpret because they only report on the metal complex in an indirect way. It can be said simply that they also show a three-component behaviour in going from low to high pH and that the high pH signals are identical to those of an aluminium-free spectrum (**Figure 7.10, right**). Some characteristic signals at further downfield can to some extent indicate the nature of the fluorine it's hydrogen bonded to and this will be discussed later in this chapter.

The overall result is clear and rather unexpected. The changes in ^{19}F NMR signals show a three-component transitional series. The simplest deduction is that at the highest pH, the tetrafluoroaluminate complex has completely given way to a

trifluoromagnesate complex. It is therefore proper to reassign the tbp in the pH 8.5 X-ray structure PDB: 5UKD and 3UKD as a trifluoromagnesate complex (**Figure 7.9**).

The unexpected feature is the rise and fall of signals for an intermediate complex. This is the first time we have seen a new metal fluoride species that appears only at intermediate pH and has three well resolved peaks that are much more downfield close to the rotationally averaged AlF_4^- (**Figure 7.10 e~l**, left) From its bell-shaped pattern of growth and decline, especially the fact that the signals disappear at highest pH, we can deduce they are most likely to be from a second aluminium fluoride species. To prove this, deferoxamine (DFO) was added to the sample at pH 8.6 when all three TSA species are present. The result is the three downfield peaks (–130.0, –134.6, and –137.4 ppm) as well as the rotationally averaged AlF_4^- peak at –139.2 ppm disappeared (**Figure 7.11**) and this shows the intermediate complex is indeed aluminium fluoride species.

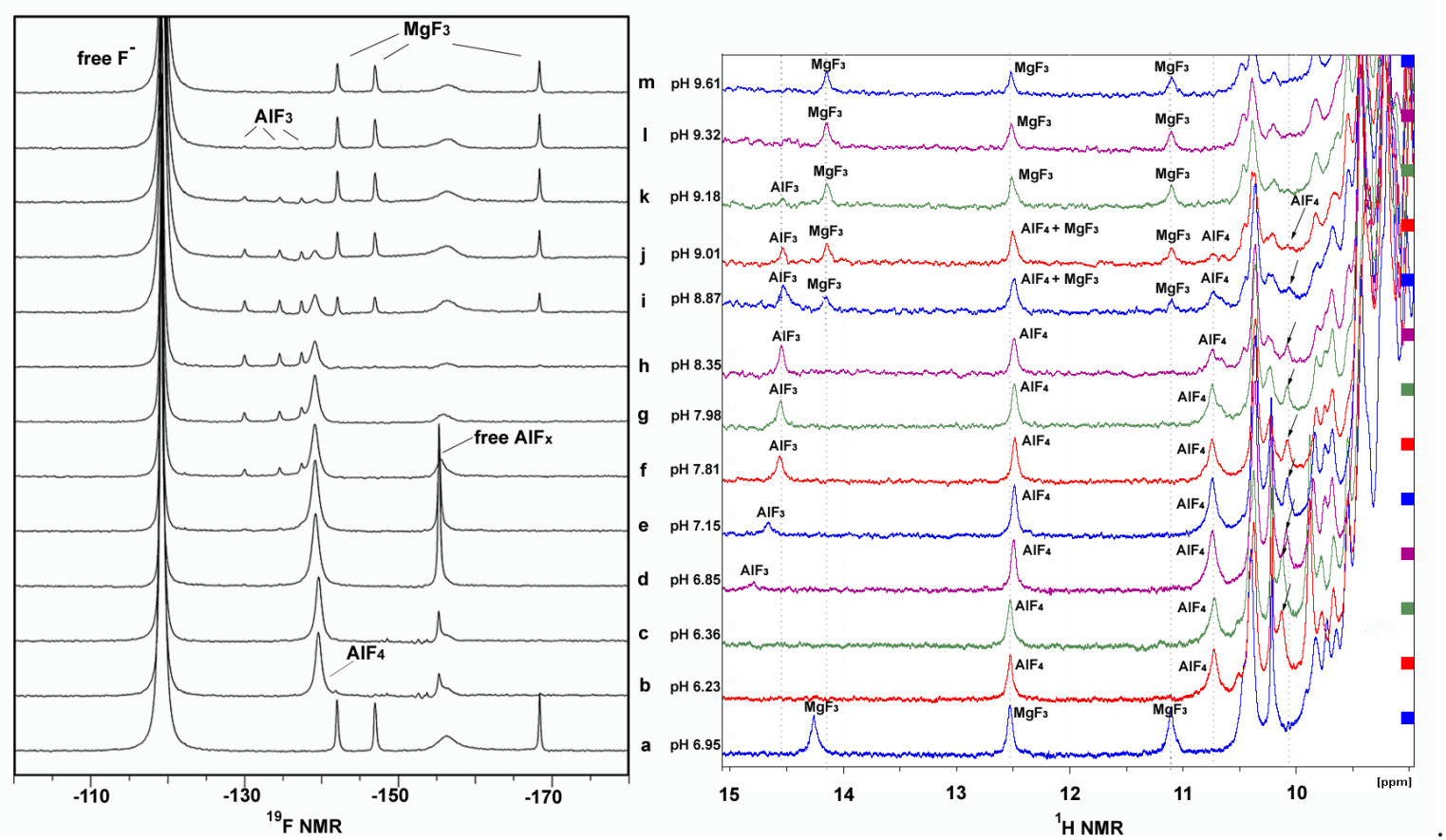


Figure 7.10 ^{19}F and ^{1}H NMR spectra for pH titration of the $UMPK_{dicty}$ metal fluoride TSA complexes over the pH range 6.23~9.61. (**a**) UMPK-UMP-MgF_3^--ADP TSA complex recorded in 0.5 mM $UMPK_{dicty}$, 50 mM Tris-HCl, 15 mM $MgCl_2$, 50 mM DTT, 10 mM NH_4F, 2 mM UMP, 2 mM ADP at pH 6.95. (**b**-**m**) UMPK-UMP-MF_x-ADP TSA complexes from pH 6.23 to 9.61 in 0.5 mM UMPK, 50 mM Tris-HCl, 15 mM $MgCl_2$, 2 mM $AlCl_3$, 50 mM DTT, 10 mM NH_4F, 2 mM UMP, 2 mM ADP. The identity of the signals and the pH in each spectrum are indicated in the middle of titration panel.

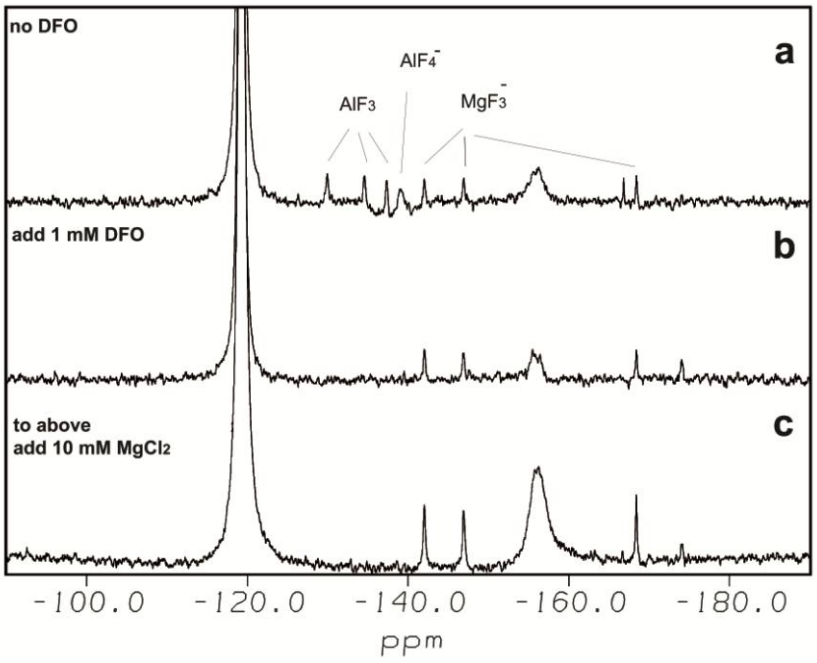


Figure 7.11 ^{19}F NMR spectra showing the four downfield signals are Al^{3+} associated TSA species. (**a**) The spectrum of a mixture of three metal fluoride TSA species in the sample, recorded in the buffer with 50 mM Tris-HCl, 0.5 mM $UMPK_{dicty}$, 5 mM $MgCl_2$, 1 mM $AlCl_3$, 50 mM DTT, 10 mM NH_4F, 2 mM UMP, 2 mM ADP at pH 8.6; (**b**) 1 mM DFO was added to (**a**) to quench Al^{3+} at pH 8.6, leaving only MgF_3^- TSA species; (**c**) a further 10 mM $MgCl_2$ was added to sample (**b**) to recover the full occupancy of the MgF_3^- TSA species at the same pH.

To make the trend of the three complex species more clear, the signal integrals in ^{19}F NMR spectra are plotted as a function of pH (**Figure 7.12**). The integrals were normalised to account for three fluorines in the magnesium and intermediate complexes and four in the tetrafluoroaluminate species. The failure of the blue magnesium curve to reach full height is because protein precipitation has already begun at high pH.

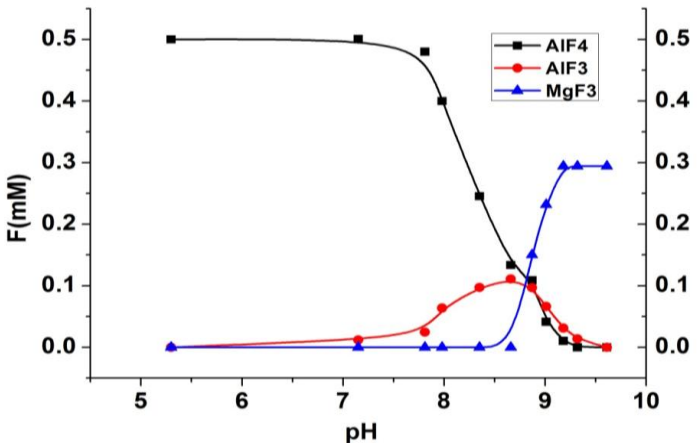


Figure 7.12 A preliminary titration curve to show the transition from AlF_4^- to MgF_3^- *via* $AlF_3(OH)^-$ complex.

It could be that there is a fundamental structure change in the metal fluoride moiety to stop free rotation of AlF_4^- and also shift these three fluorine peaks downfield. So far, we cannot determine the geometry of this aluminium-associated intermediate species only based on the ^{19}F NMR SIIS and chemical shift data (**Table 7.4**). But these data can be interpreted in the light of the published structures to deduce the possible identity of this aluminium trifluoride complex.

Table 7.4 ^{19}F chemical shift (ppm) and SIIS (ppm) of the UMPK-ADP-AlF_3-UMP TSA complex.

Metal Fluoride Complex		**δF** (D_2O capillary)	**ΔδF** (D_2O capillary-100% D_2O)
UMPK-ADP-AlF_3-UMP	F_A (F_4)	–129.9	1.0
	F_B (F_1)	–134.6	0.3
	F_C (F_3)	–137.3	1.0
	average	–133.3	—

First, the three fluorine signals could be from a UMPK-ADP-AlF_3^0-UMP TSA complex in which AlF_3^0 moiety is tbp as in the crystal structures PDB: 3UKD. Because in solution, pH titration shows this intermediate complex with constant chemical shifts exists from pH 7.15 to 9.32 including the pH for the crystallisation for PDB: 3UKD and 5UKD structure. But there is no obvious reason why the intensity of this hypothetical AlF_3^0 complex should have a bell-shaped pattern with respect of the pH. Moreover, if the tbp AlF_3^0 moiety binds in the same way as UMPK-ADP-MgF_3^--UMP, its SIIS should follow the same pattern as MgF_3^- as in **Table 7.3**. However, F_B in AlF_3 at –134.6 ppm has the smallest SIIS at 0.3 ppm, much smaller than F_A and F_C (**Table 7.5**). The most upfield F_C which should have shown the smallest SIIS as it is coordinated to the catalytic magnesium in structure PDB: 3UKD by contrast has a fairly big SIIS (1.0 ppm). Thus, a UMPK-ADP-AlF_3^0-UMP TSA complex as is assigned in crystal structures 3UKD and 5UKD appears to be incorrect.

Second, from SIIS point of view, among the four ligands of the central Al^{3+} in structure 1QF9 (**Figure 7.13**), F_4 has very good hydrogen bonding connections to Arg 148 and Arg 131 at 2.7 Å. F_1 coordinates to the side chain of Arg 148 with two hydrogen bonds at 2.9 and 3.1 Å, but it is 3.3 Å from Arg 93 side chain and 3.2 Å from Gly 15 backbone amide with a poor hydrogen bond angle. These interactions match the SIIS of F_4 and F_1 very well. This means one of the F_2 or F_3 fluorines in the UMPK-ADP-AlF_4^--UMP TSA complex could be replaced by water or hydroxide and the aluminium remains octahedral

to give a UMPK-ADP-$AlF_3(H_xO)$-UMP complex. If the replacing ligand is water (x = 2), it has to donate two hydrogen bonds to receptors in the complex. That is only possible for F_3 which can donate one hydrogen bond to α-phosphate oxygen of UMP (**Figure 7.9**) and one to Wat1 (**Figure 7.13**). However, that means that F_C has to be located in the position F_2 and that is incompatible with its downfield chemical shift and large SIIS (**Table 7.4**). Moreover, the substitution of F^- by a H_2O molecule in tetrafluoroaluminate complex only occurs after a positive charged residue coordinating to fluoride has been mutated to a neutral residue (alanine in most cases), as in PGK and βPGM.(Cliff *et al.*, 2010) Clearly, in our pH titration the presence of negative charged metal fluoride complexes throughout the whole pH range indicates the positive charges in the active site are constant, the F^- to H_2O substitution does not response to changes in the charge-charge interactions. Further, the concentration of water in the sample is 55 M and there is no reason why its competitiveness dramatically increases only inside this narrow pH 8 ~ 9 window. Therefore, octahedral UMPK-ADP-$AlF_3(H_2O)^0$-UMP TSA complex is unlikely to be the answer.

Therefore, the replacement of one of the four fluorides by hydroxide appears to be more appropriate, as an octahedral UMPK-ADP-$AlF_3(OH)^-$-UMP TSA complex. This substitution only requires one hydrogen bond to be donated from the OH group and that is easily accommodated by occupancy of site F_2 that can hydrogen bond to αP oxygen. It means that Fc corresponds to site F_3 and should have a downfield chemical shift and large SIIS, as is observed in the structure that, F_3 coordinates to a water molecule at 2.6 Å, Lys 19 side chain at 3.0 Å and it is 3.3 Å from side chain of Arg 93. The bell-shaped pattern of AlF_3OH^- in the replacement of F^- by OH^- can be explained simply because the population of OH^- increases relative to that of fluoride as the pH rises. The signals then decline with the high pH precipitation of aluminium. This proposal should be easily tested by increasing the fluoride concentration to a significant level in order to change the equilibrium between protein-bound AlF_4^- and AlF_3OH^- complexes while the pH retains the same. This experiment failed because when the concentration of F^- was elevated by 20 fold (from 10 mM to 200 mM) the protein precipitated by the action of NH_4F, which destroys the solvation of $UMPK_{dicty}$.

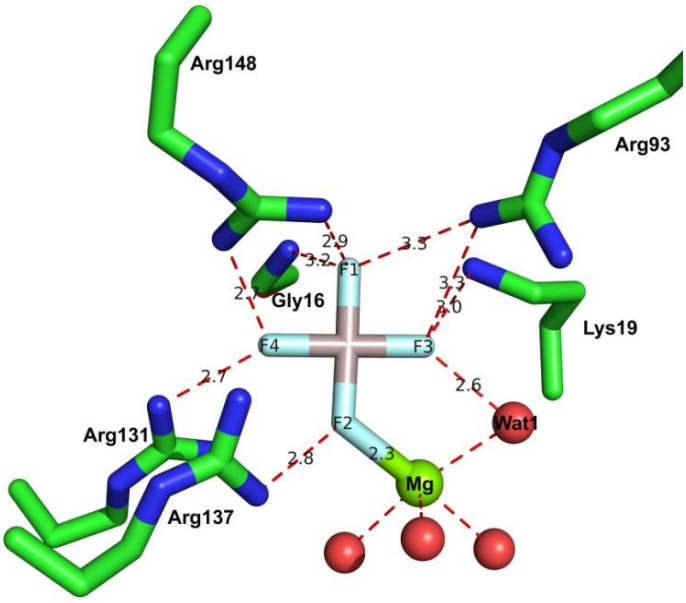


Figure 7.13 The hydrogen bonds length of four fluorines in structure PDB:1QF9 (UMPK-ADP-AlF_4^--UMP TSA complex) for the reference of SIIS in ^{19}F NMR.

^1H NMR is supportive to this interpretation. The most downfield peak at around 14.5 ppm is a signature for the protons which we identify as hydrogen bonded to AlF_3 moiety. It shifts 0.3 ppm upfield as the pH is increased from 6.85 to 9.18, showing this hydrogen is strongly affected by pH change. Since we think in 1QF9, where F_2 is the most likely place for OH^- and this OH^- has the proton acceptor α-phosphate oxygen and only accepts an H-bond from Arg 137 which is strongly hydrogen bonded to Asp 139 (**Figure 7.9**). This hydrogen on Arg 137 directly reacts to the pH change at low pH because of the protonation status of Asp 139, which is very likely to be the resonance at 14.5 ppm and shifts along the pH.

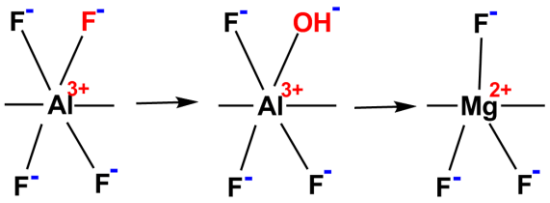


Figure 7.14 Diagram to show how metal fluoride moiety may switch from AlF_4^- to $AlF_3(OH)^-$ to MgF_3^- with rising pH.

In pH titration experiments of all the other four proteins (PSP, βPGM, PGK, cAPK) we have studied, the four resonances from AlF_4^- always appear together at the same time and gradually recede together along with rising pH due to aluminium precipitation. The loss of just one fluoride from their tetrafluoroaluminium species has not been observed.(Baxter *et al.*, 2008; Jin *et al.*, submitted) Why is $UMPK_{dicty}$ so special? The

above proposal satisfies both the analysis of the peak intensities and also the CBH. As no similar case has been described for other enzymes, further experiments need be done to try to determine the identity of the fourth aluminium ligand.

7.3.3 UMPK-ADP-BeF_2^0-UDP TSA complex

Since the nature of the major TSA complexes in $UMPK_{dicty}$ have been established, a ground state complex becomes desirable to tell us more about the conformational change between ground state and transition state. The obvious way to make a GSA in this case is to add ADP, Be^{2+}, F^-, and UMP to form an 'UMPK-ADP-BeF_3^--UMP' ground state analogue (GSA) complex in which ADP-BeF_3^- mimics ATP in the ground state. BeF_3^- complexes generally show two different profiles depending on the binding environment for other phosphoryl transfer proteins. In PSP, hPGK, and cAPK, BeF_3^- gives three well resolved resonances in 1:1:1 ratio corresponding to three non-equivalent fluorines. In βPGM, depending on the hexose phosphate used in the complex, the complex can be seen either as a rotationally averaged broad peak with integration for 3 fluorines or as three resolved peaks. However, when ADP, Be^{2+}, F^-, and UMP were added to the UMP/CMP kinase sample, the ^{19}F NMR unexpectedly shows only two peaks of small line width and this does not match any of the NMR situations previously observed.

The only published beryllium fluoride complex X-ray structure is that of a UMPK-UDP-BeF_2^0-ADP inhibitor complex (PDB: 4UKD). It was produced by a serendipitous discovery when UMP and ADP were added to the protein sample. The authors' objective had been to form a GSA complex – just as we had intended.(Schlichting *et al.*, 1997) The authors concluded that the production of UDP from UMP, given all the rigorous purification procedures by carried out, was due to a trace amount of contaminative adenylate kinase able to convert two equivalents of ADP into one AMP and ATP. The $UMPK_{dicty}$ then used the ATP to phosphorylate UMP to UDP. The published K_m for ATP is 25 μM, and K_m for UMP is 400 μM. The binding constant and association rate of UDP-BeF_2^0-ADP complex were measured to be k_{on} = 103 M^{-1} s^{-1}, k_{off} = 0.005 s^{-1}, K_d = 5 μM (measured at 250 μM ADP and 500 μM CDP), and that may not be low enough to inhibit the phosphorylation.(Schlichting *et al.*, 1997)

We believe, on the other hand, that the generation of ATP (2ADP → AMP + ATP) is one of the promiscuities of $UMPK_{dicty}$ that takes place when the concentration of

diphosphates is high enough. We used the strong inhibitor Ap_5A (50 μM) to inhibit any contaminative adenylate kinase in the sample at a concentration higher than K_d (50 nM to AK), but its inhibition effect did not stop the production of UDP. So this proves the purity of the $UMPK_{dicty}$ prep is very good and the observed adenylate kinase activity must come from $UMPK_{dicty}$ itself.

We therefore designed a positive control experiment and added only UDP and ADP to the sample to see if the NMR signals we observed were due to the same reaction as in the crystallography experiment. The ^{19}F NMR spectra showed that the two major peaks in the ADP-BeF_2^0-'UMP' complex appear at exactly the same chemical shifts as that of ADP-BeF_2^0-UDP (**Figure 7.15**). This thus proves the assignment of the nucleoside diphosphate as UDP is correct.

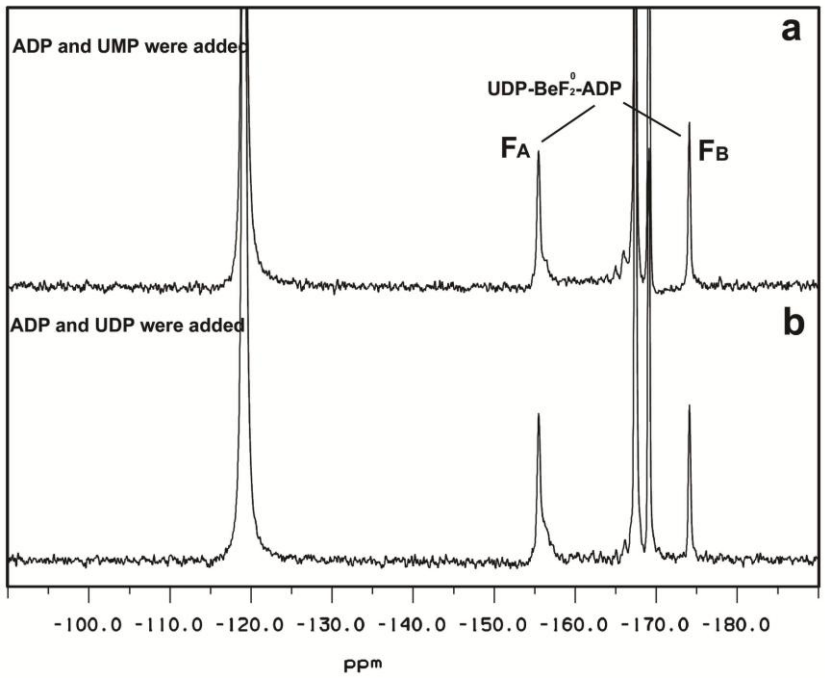


Figure 7.15 ^{19}F NMR spectra of UMPK-UDP-BeF_2^0-ADP complex: (**a**) recorded in buffer with 0.5 mM $UMPK_{dicty}$, 2 mM UMP, 2 mM ADP, 5 mM $BeCl_2$, 10 mM NH_4F at pH 7.3; (**b**) recorded in buffer with 0.5 mM $UMPK_{dicty}$, 2 mM UDP, 2 mM ADP, 5 mM $BeCl_2$, 10 mM NH_4F at pH7.3. The protein-bound BeF_2^0 bisubstrate complex signals are F_A = –155.4 ppm, and F_B = –174.1 ppm. Free F^- = –191.3 ppm and the signals between –165.0 and –170.0 ppm are free BeF_x speices.

The two fluorides in UDP-BeF_2^0-ADP inhibitor complex have chemical shifts at –155.4 ppm and –174.1 ppm. The ~20 ppm chemical shift difference between the two fluorines can now be matched to the chemical environment around those two Fs in the light of the published crystal structure (**Figure 7.16**, PDB: 4UKD). The high field

fluorine F_B at –174.1 ppm is bridging between the beryllium and catalytic magnesium, being distantly hydrogen bonded to the side chain of Arg 137 (3.3 Å). F_A is more deshielded by three short hydrogen bonds with Arg 131 (2.7 Å), Arg 137 (2.8 Å) and Arg 148 (3.1 Å) around it to shift downfield to –155.4 ppm.

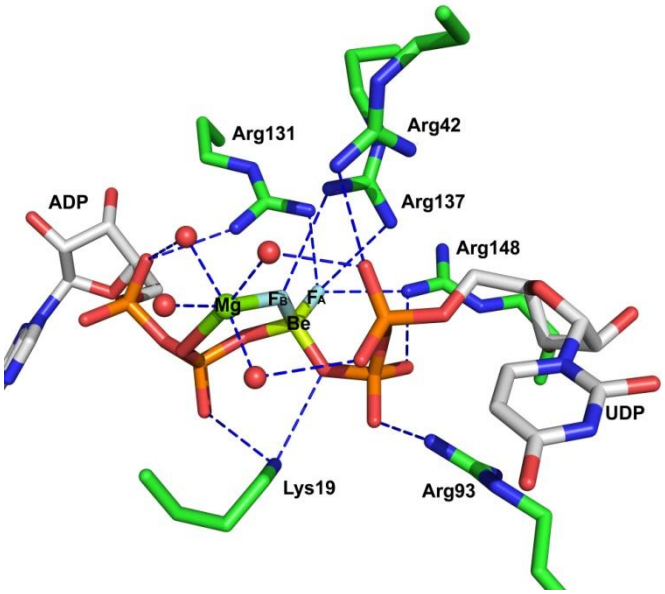


Figure 7.16 Structure of UMPK-ADP-BeF_2^0-UDP complex (PDB: 4UKD). Catalytic magnesium is a green sphere, BeF_2 is in yellow and light blue sticks, ADP and UDP are in white sticks, and water molecules are in red spheres. All the H-bonds are in red dashed lines.

It has been established that Ap_5A at 20 μM inhibits $UMPK_{dicty}$ by 45%, whereas the same concentration of Up_5A completely inhibits the enzyme. For Up_5A, K_i = 50 nM measured with UMP as the constant substrate (10 nM if ATP is constant). $UMPK_{dicty}$ can also bind Ap_5A with K_d = 160 nM, which is 53-fold weaker than Up_5A binding (K_d = 3 nM).(Wiesmuller *et al.*, 1995) So an ADP-BeF_2^0-ADP assembly may also be a possible inhibitor complex. We used ^{19}F NMR to prove that binding of such an ADP-BeF_2^0-ADP complex is much weaker (20 times at least) than for a UDP-BeF_2^0-ADP by comparing the intensity of signals when UDP was replaced by the same concentration of ADP (the ADP total concentration was thus 4 mM) (**Figure 7.17**). The two peaks corresponding to the two fluorines in this complex are not affected by the nucleotide species change from UDP to ADP.

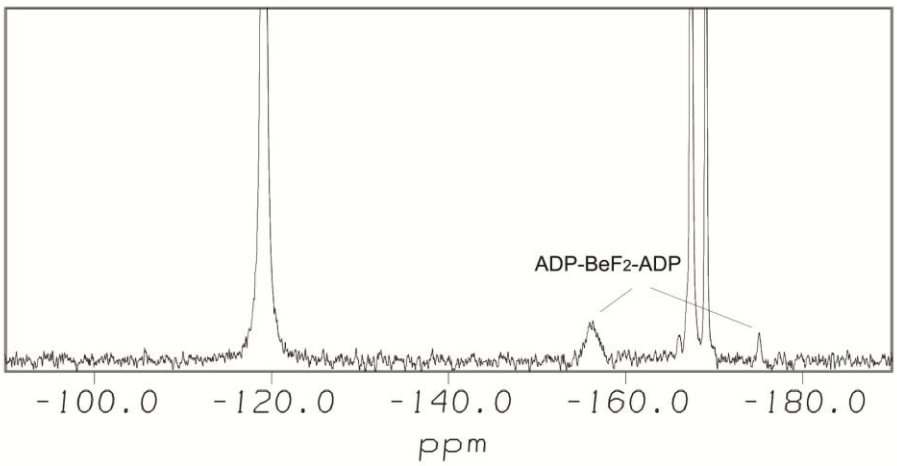


Figure 7.17 ^{19}F NMR of the weak ADP-BeF_2^0-ADP inhibitor complex recorded in buffer with 0.5 mM $UMPK_{dicty}$, 4 mM ADP, 5 mM $BeCl_2$, 10 mM NH_4F at pH 7.3. The protein-bound ADP-BeF_2^0-ADP bisubstrate complex signals F_A = –155.8 ppm, and F_B = –175.0 ppm. Free F^- = –191.3 ppm and the signals between –165.0 and –170.0 ppm are free BeF_x speices.

7.3.4 Enzyme activity inhibition test of UMP-AlF_4^--ADP TSA complex

To solve the enigma of the formation of UDP in the sample even with AlF_4^- present (which has a binding constant K_d = 25 μM), ^{31}P NMR was used to monitor the turnover of the nucleotides and a time course was recorded. To setup the sample, 0.4 mM UMPK in 50 mM Tris-HCl (pH 7.0), 1 mM $AlCl_3$, 5 mM $MgCl_2$, 10 mM NH_4F and 50 mM DTT were premixed in a 5 mm NMR tube. 5 mM UMP and 5 mM ADP were orderly added to the sample and the acquisition of ^{31}P NMR data begun immediately after locking and shimming. Each spectrum had 2560 scans and took 54 min 10 s.

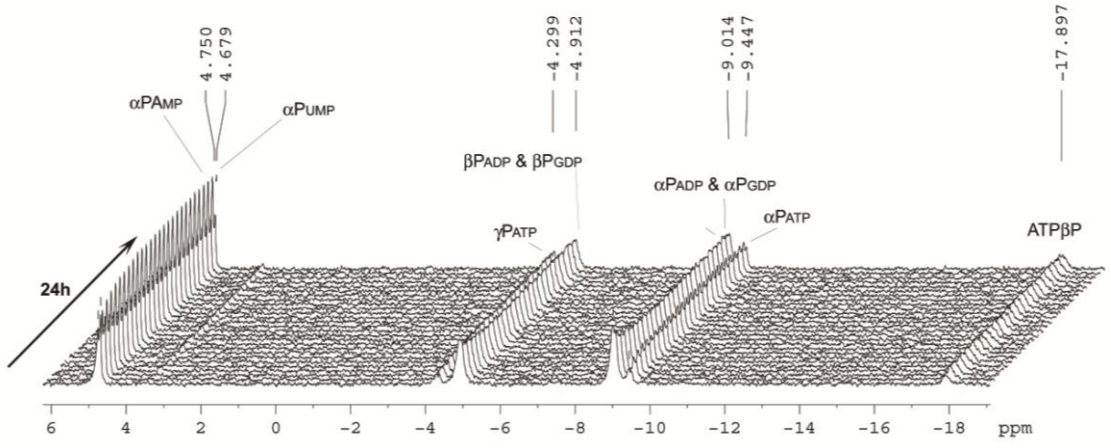


Figure 7.18 ^{31}P NMR spectra stack monitoring the turnover of UMP and ADP into other nucleotides. The time course is recorded using 0.4 mM $UMPK_{dicty}$ in 50 mM Tris-HCl (pH 7.0), 1 mM $AlCl_3$, 5 mM $MgCl_2$, 10 mM NH_4F and 50 mM DTT. Each spectrum has 2560 scans and takes 54 min 10 s.

What should be pointed out is, because the sensitivity of distinguishing two close resonances in ^{31}P NMR depends on a lot of factors, such as chemical shift difference, *J* coupling constant, and line width. Moreover the ability to identify different diphosphate or triphosphate species is fairly limited because the phosphorus signals for ADP and UDP overlap with each other intrinsically so cannot be discriminated. That is also the case for ATP and UTP.

The whole observed activity assay is, to some extent, a black box. Even so, we can still identify the major reactions happening during the time course. We know the nucleotides added into the sample at the start are UMP and ADP. Along the time course, consumption of UMP (main reaction) and the production of AMP (side reaction) are observed, also from the result of UDP-BeF_2^0-ADP we are sure that UDP is also produced (main reaction). ATP is produced and accumulated too, although it has to be the phosphate donor for UMP to make UDP. From the phenomenon that the peaks at –4.9 and –9.0 ppm look diminished shows the reduced amount of ADP is more than the increased amount of UDP. So the loss of total ADP in the sample equals the production of AMP and ATP. In the same way, the loss of UMP equals the production of UDP. At the end of the time course, there would be AMP, ADP, ATP, UMP and UDP in the sample.

$$ATP + UMP \rightarrow ADP + UDP \text{ (main reaction)}$$
$$UDP + ADP \rightarrow ATP + UMP \text{ (reverse reaction)}$$
$$2ADP \rightarrow ATP + AMP \text{ (side reaction)}$$

Because the inhibition of ADP-AlF_4^--AMP complex to UMPK is not as strong, (supposing it should inhibit AK), when the ADP's concentration is fairly high, it is certain that the only equilibrium of the side reaction is towards making ATP and AMP. ATP is utilised as the phosphorylation reagent for UMP as soon as it is produced (main reaction). At low concentration, this main reaction is not inhibited by the aluminium fluoride concentration used in the assay.

7.3.5 Attempts to make UMPK-UDP-AlF_3^0-ADP and UMPK-UDP-AlF_3^0-AMPCP complexes to test charge balance in UMP/CMP kinase

The CBH has evolved from binding of TSA. Transition state theory predicts that TSAs bind more tightly to the enzyme than substrates by the factor of k_{cat}/k_{uncat} in which charge plays a big role.(Schramm, 2007)

So far, we have shown that metal fluoride TSA complexes for UMPK obey the CBH, such as UMPK-ADP-AlF_4^--UMP TSA complex, UMPK-ADP-MgF_3^--UMP TSA complex, and UMPK-UDP-BeF_2^0-ADP complex, and they all have 6 net negative charges on the combined substrates (**Section 7.3.1** and **7.3.3**). The UMPK-Up_5A complex has five net negative charges on Up_5A, and the spacing of its five-phosphates is very similar to that of the UDP-BeF_2^0-ADP complex, so it also binds very well as a bisubstrate analogue complex.

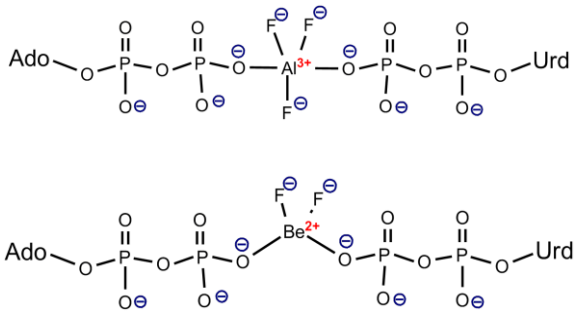


Figure 7.19 The proposed structures of a UMPK- UDP-AlF_3^0-ADP complex compared to UMPK-UDP-BeF_2-ADP.

It is especially valuable to have metal fluoride complexes as a tool to look at the TS of phosphoryl transfer enzymes without complicated synthetic processes to make stable bisubstrate phosphate analogues. If indeed a tbp AlF_3^0 TSA complex can exist, it is very straightforward to synthesise it using the well established method of making metal fluoride anion complexes to be observed by ^{19}F NMR.

In **Figure 7.19**, it is possible that an AlF_3^0 complex might be formed in similar fashion to the BeF_2 complex. It would have tbp geometry and two axial ligands as the two βOs from ADP and UDP. To set up the formation of such a complex, 2 mM ADP and 2 mM UDP were added to 0.5 mM $UMPK_{dicty}$ solution together with 1 mM $AlCl_3$ and 10 mM NH_4F. However, the ^{19}F NMR showed only one resonance present and it corresponds to that of the known UMPK-ADP-AlF_4^--UMP complex (**Figure 7.20**), even though no UMP had been added. ^{31}P NMR shows not only ADP and UDP present in the sample, but also significant amounts of ATP, AMP, UMP and Pi produced by side reactions (**Figure 7.21**).

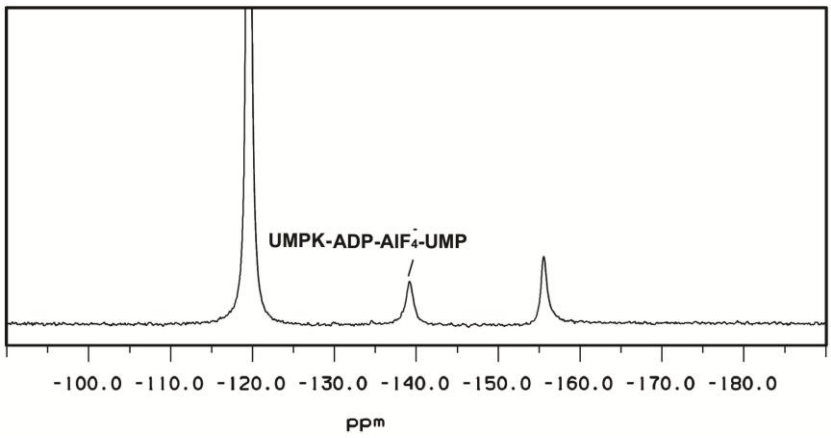


Figure 7.20 ^{19}F NMR spectrum of a sample intended to make UMPK-UDP-AlF_3^0-ADP complex. Instead, a UMPK-ADP-AlF_4^--UMP complex is observed (−139.0 ppm). The spectrum is recorded in 0.5 mM $UMPK_{dicty}$, 50 mM Tris-HCl, 5 mM $MgCl_2$, 1 mM $AlCl_3$, 50 mM DTT, 10 mM NH_4F, 2 mM UDP, and 2 mM ADP at pH 7.5.

As a result of the production of additional nucleotides by residual phosphoryl transfer activity, it proved impossible to form a complex with only the desired combination of nucleotides (UDP and ADP). With this in mind, and in light of the failure to detect the UDP-AlF_3^0-ADP complex, it appeared useful to recruit a suitable non-hydrolysable ADP analogue to avoid these undesirable side reactions. A stable ADP analogue adenosine 5'-(α,β-methylene)diphosphate (AMPCP) was employed instead of ADP and see if a UMPK-AMPCP-AlF_3^0-UDP complex can be formed. AMPCP as its sodium salt was generously given by Prof. Mike Blackburn. For convenience, it was made into a 50 mM stock and ^{31}P NMR showed it is stable in water at neutral pH, rt for at least one month.

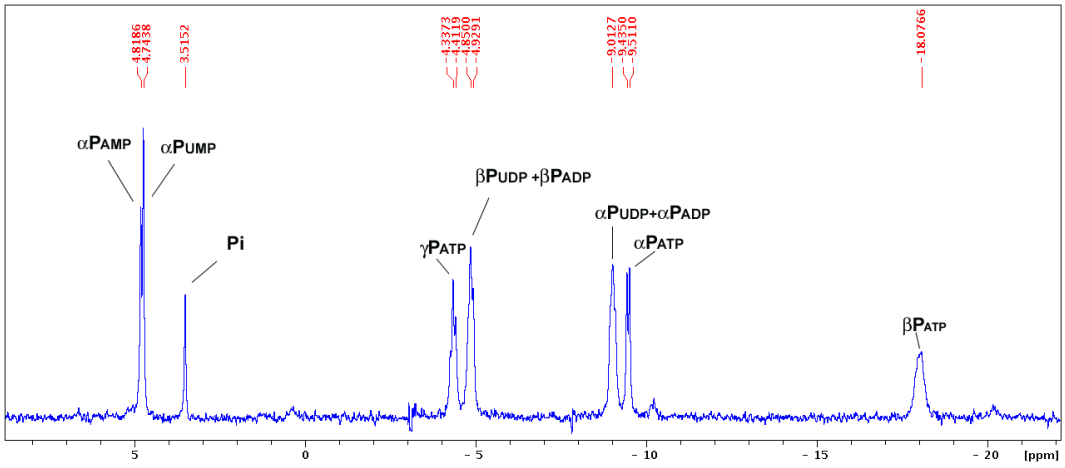


Figure 7.21 ^{31}P NMR spectrum of the sample intended to make UMPK-UDP-AlF_3^0-ADP complex, recorded in 0.5 mM $UMPK_{dicty}$, 50 mM Tris-HCl, 5 mM $MgCl_2$, 1 mM $AlCl_3$, 50 mM DTT, 10 mM NH_4F, 2 mM UDP, and 2 mM ADP at pH 7.5. All phosphorus species are labelled with abbreviations: *e.g.*, αP_{ATP}.

To 0.5 mM $UMPK_{dicty}$ solution, 1 mM $AlCl_3$ and 10 mM NH_4F were added, and only AMPCP and UDP were used as substrates at pH 7.5, and ^{31}P NMR data was also collected after the ^{19}F NMR experiment. In the ^{19}F NMR there was almost no detectable complex at the beginning and only after 12 h was there a weak, rotationally averaged peak which was in fact UMPK-AMPCP-AlF_4^--UMP complex (**Figure 7.22**), to be discussed later in this chapter.

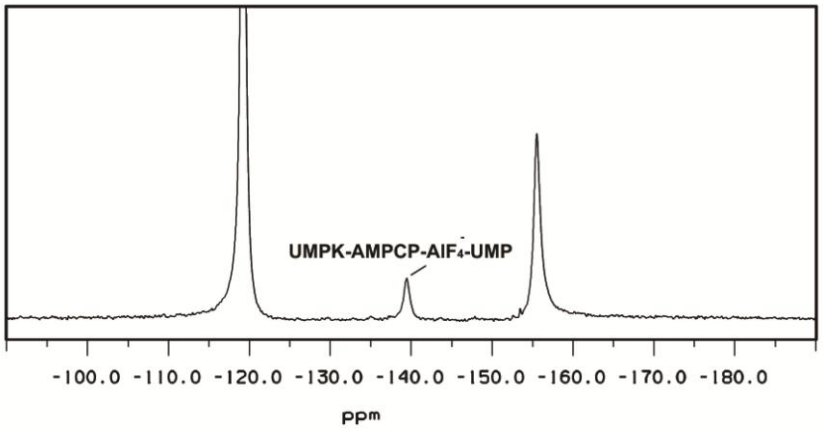
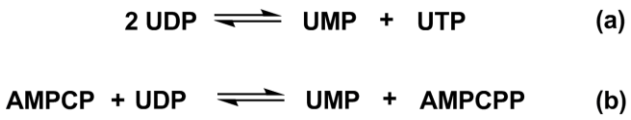


Figure 7.22 ^{19}F NMR spectrum of sample intended to make a UMPK-AMPCP-AlF_3^0-UDP complex. Instead, a UMPK-AMPCP-AlF_4^--UMP complex is seen (−139.6 ppm), recorded in buffer with 0.5 mM $UMPK_{dicty}$, 50 mM Tris-HCl, 5 mM $MgCl_2$, 1 mM $AlCl_3$, 50 mM DTT, 10 mM NH_4F, 2 mM UDP, and 1 mM AMPCP at pH 8.0.

From the ^{31}P NMR spectrum, UDP, UTP, UMP, AMPCPP and small amount of Pi were seen in the sample at the end of 2 d. The two obvious reactions taking place in the sample are:



Reaction (**a**) could be another promiscuity which $UMPK_{dicty}$ can do even though most of the studies so far have only focused on a residual activity of $UMPK_{dicty}$ as being an adenylate kinase. $UMPK_{dicty}$ has been found to have a narrow specificity for nucleoside monophosphates but a broader specificity for nucleoside triphosphate, therefore the relative activity with UTP and UMP as the substrates is only 4% as that of ATP and UMP as the substrates.(Wiesmuller *et al.*, 1990) However, this shows at least the recognition of uridine in adenosine binding site is not stringent and that is part of reason why it can still catalyse the reverse reaction when excess of UDP is present.

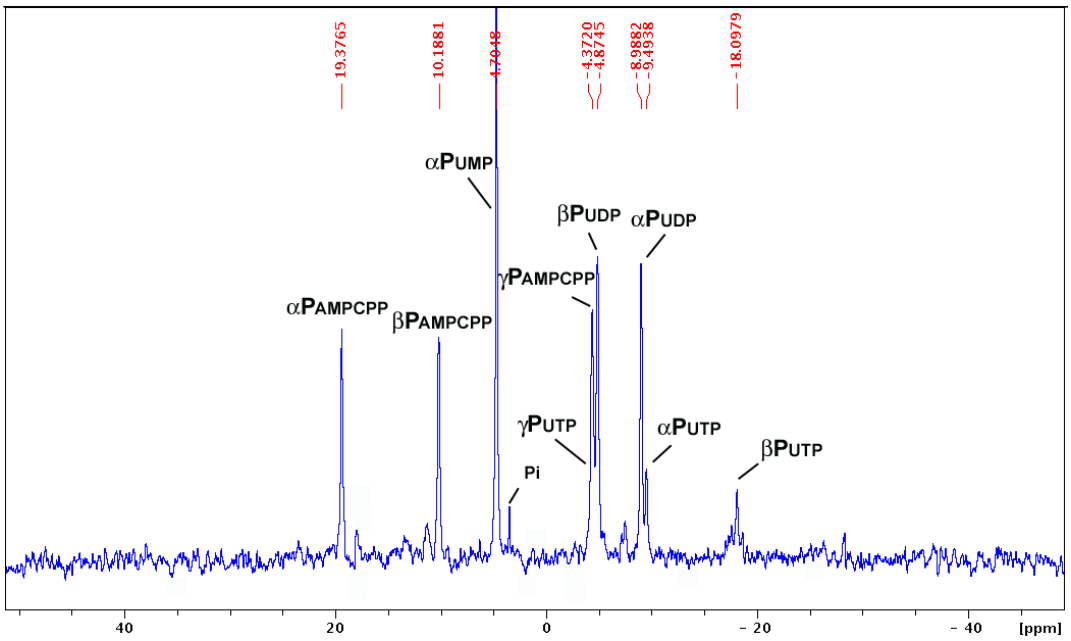


Figure 7.23 ^{31}P NMR spectrum of sample designed to make a UMPK-AMPCP-AlF_3^0-UDP complex. Recorded in buffer with 0.5 mM $UMPK_{dicty}$, 50 mM Tris-HCl, 5 mM $MgCl_2$, 1 mM $AlCl_3$, 50 mM DTT, 10 mM NH_4F, 2 mM UDP, and 1 mM AMPCP at pH 8.0. All the phosphorus species are labelled with abbreviations, *e.g.*, αP_{UTP}.

Reaction (**b**) is also the reverse reaction of the phosphorylation of UMP by ATP. Since the AMPCP is stable and so cannot be a phosphate donor, and also UDP is in large excess in the solution, the enzyme instead catalyses the phosphorylation of AMPCP using UDP as the phosphate donor. The reaction was so effective that by the time the ^{31}P NMR was measured, there was no AMPCP left. This can be seen from chemical shift changes of α- and β-phosphates in which β-phosphate moves 13.5 ppm upfield (from 23.5 ppm to 10 ppm) after it is attached to γ-phosphate. The α-phosphate moves 6.2 ppm downfield (from 13.2 ppm to 19.4ppm) (**Figure 7.23**).

In both these cases, there were no observable ^{19}F resonances corresponding to the three downfield peaks between pH 7.0 ~ 9.0, and no other new metal fluoride complex was formed. The conclusion is that UMPK-ADP-AlF_3^0-UDP or UMPK-AMPCP-AlF_3^0-UDP cannot be constituted as TSA complexes by $UMPK_{dicty}$. Even though AlF_3^0 and BeF_2^0 have the same zero charge, the geometrical demand of the two metals is different. In the structure of UMPK-UDP-BeF_2^0-ADP (PDB: 4UKD), beryllium is tetrahedral coordinating to two fluorines and two β-oxygens from ADP and UDP. The distance between the two oxygens is 2.56 Å and the O–Be–O angle is 109.01°. Whereas in all the structures assigned to AlF_3^0, the O–Al–O angle is always close to 180° and the distances between the two oxygens are always in the range 4.0 to 4.5 Å. This is because

the axial bond length for Al–O is always around 2.0 to 2.3 Å. We conclude that after $UMPK_{dicty}$ binds its two nucleotide substrates, it closes up to make a short distance for the γ-phosphate of ATP to transfer to UMP. As a result, there is only enough space for four phosphates in the TS. The five phosphates in the stable bisubstrate analogue bind in a distorted conformation in order to accommodate the additional charged moiety. UMPK-UDP-AlF_3^0-ADP has the same 6 negative charges as the TS, but it has one more 'phosphate' than the true TS. This cannot be accommodated by the active site as the positions of the phosphates are well defined by the arginines, lysine, and catalytic magnesium that bridge the β- and γ-oxygens of ATP in all the structures studied, *e.g.* UMPK-ADP-BeF_2-UDP (4UKD), UMPK-ADP-AlF_3^0-CMP (3UKD and 5UKD), and UMPK-Up_5A (1UKD).

7.3.6 A UMPK-UMP-AlF_4^--AMPCP TSA complex and attempts to form BeF_3^- GSA using stable ADP analogues

The attempts to make a UMPK-ADP-BeF_2^0-UDP complex from ADP and UMP, described above, showed that inhibition of the adenylate kinase activity of $UMPK_{dicty}$ was not strong enough to stop the side reaction 2ADP→AMP + ATP even using the effective inhibitor BeF_x (K_d = 5 μM) at 5 mM $BeCl_2$ and 10 mM NH_4F or Ap_5A at 50 μM. If the ADP can be converted into AMP and ATP, and ATP in turn phosphorylates UMP into UDP, the situation becomes too complex and difficult to analyse. To prevent this, it appeared that other stable ADP analogues like AMPCP and $AMPCF_2P$ should be explored.

The α,β-bridging CH_2 in AMPCP is less electronegative than the bridging oxygen in ADP, giving pK_{a3} of β-phosphonate in AMPCP of 8.1. This pK_{a3} difference effect is expected to show when AMPCP is used for the complex. To setup the experiment for a UMPK-AMPCP-MF_x-UMP complex, 2 mM UMP and 2 mM AMPCP was added to 0.5 mM $UMPK_{dicty}$ in buffer of 50 mM Tris-HCl, 5 mM $MgCl_2$, 50 mM DTT, 10 mM NH_4F at pH 7.0. Unlike our results using ADP, ^{19}F NMR shows that at pH 7.0 there is weak UMPK-AMPCP-MgF_3^--UMP complex formation with chemical shifts of the three peaks at –139.3, –146.7, and –167.5 ppm **(Figure 7.24a)**.

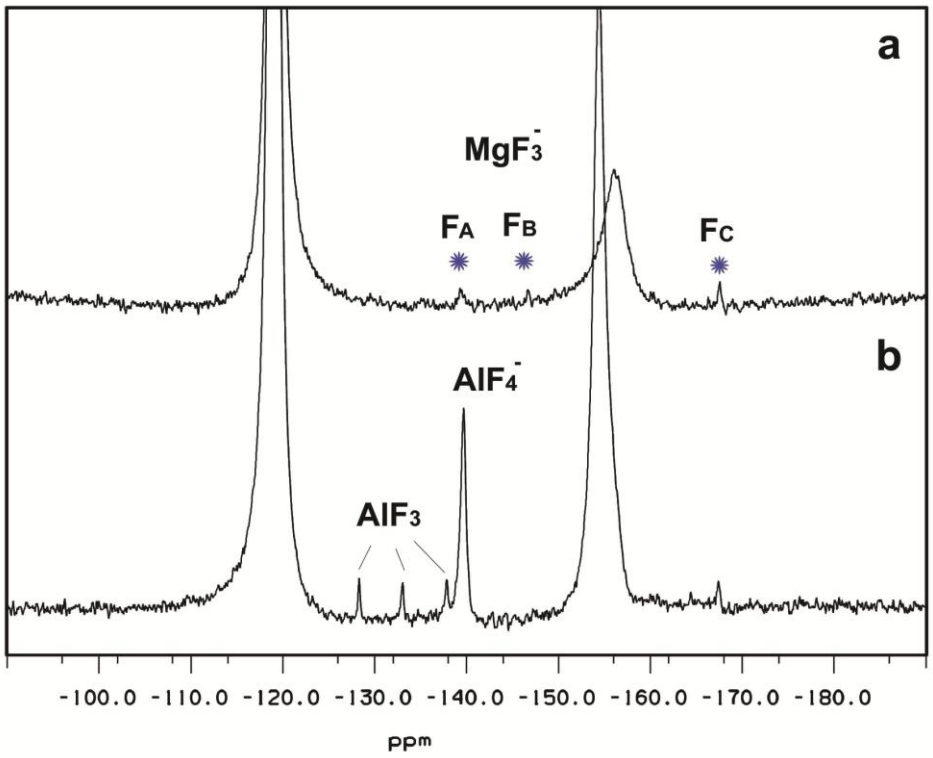


Figure 7.24 ^{19}F NMR spectra of (**a**) a weak UMPK-UMP-MgF_3^--AMPCP, in 0.5 mM $UMPK_{dicty}$, 50 mM Tris-HCl, 5 mM $MgCl_2$, 50 mM DTT, 10 mM NH_4F, 2 mM UMP, and 2 mM AMPCP at pH 7.0. The protein-bound MgF_3^- TSA complex signals F_A = –139.3 ppm, F_B = –146.7 ppm, and F_C = –167.5 ppm. (**b**) UMPK-UMP-AlF_4^--AMPCP TSA complex in the same buffer as (**a**) but added 2 mM $AlCl_3$ at pH 7.4. The signal from protein-bound AlF_4^- TSA complex is rotationally averaged at –139.6 ppm. Three downfield peaks F_A = –128.3, F_B = –133.0, and F_C = –137.8 ppm are assigned as protein-bound AlF_3OH^- TSA complex.

Apart from the molecular conformation difference, the low concentration of MgF_3^- in solution may also be responsible if it is not stable enough to overcome the energy barrier of deprotonation of the β-phosphonate, when the pH of the sample is lower than the pK_{a3}. The pH is therefore raised to 9.0, but there is still hardly any MgF_3^- complex (data not shown). This means the subtle structural difference between AMPCP and ADP such as bond angles and lengths can also affect the formation of a weakly bound MgF_3^- complex.

By contrast, when 2 mM $AlCl_3$ was added at pH 7.4, the ^{19}F NMR shows a UMPK-AMPCP-AlF_4^--UMP complex is formed with a rotationally averaged signal at –139.6 ppm, 0.6 ppm upfield compared to the signal from UMPK-ADP-AlF_4^--UMP complex (**Table 7.5**). At the same time, three downfield peaks F_A = –128.3, F_B = –133.0, and F_C = –137.8 ppm in 1:1:1 ratio are also present **(Figure 7.24b)**, among which F_A and F_B both 1.6 ppm downfield relative to the corresponding peaks observed in the spectrum

when ADP is used (**Figure 7.11**, **Table 7.5**). A ^{31}P NMR spectrum of sample was also recorded (**Figure 7.25**) which shows the two nucleotides are both stable so that there is no interference from other nucleotide species.

Table 7.5 ^{19}F chemical shifts (ppm) of the UMPK-UMP-AlF_4^--ADP TSA and UMPK-UMP-AlF_4^--AMPCP TSA complexes.

Complex	F_A	F_B	F_C	$F_{rot\ ave}$
UMPK-UMP-AlF_4^--ADP	–129.9	–134.6	–137.3	–139.0
UMPK-UMP-AlF_4^--AMPCP	–128.3	–133.0	–137.8	–139.6

* The three downfield signals F_A, F_B and F_C have yet to be identified but are likely to be AlF_3OH^-. The rotationally averaged fluorine signal is from an AlF_4^- complex.

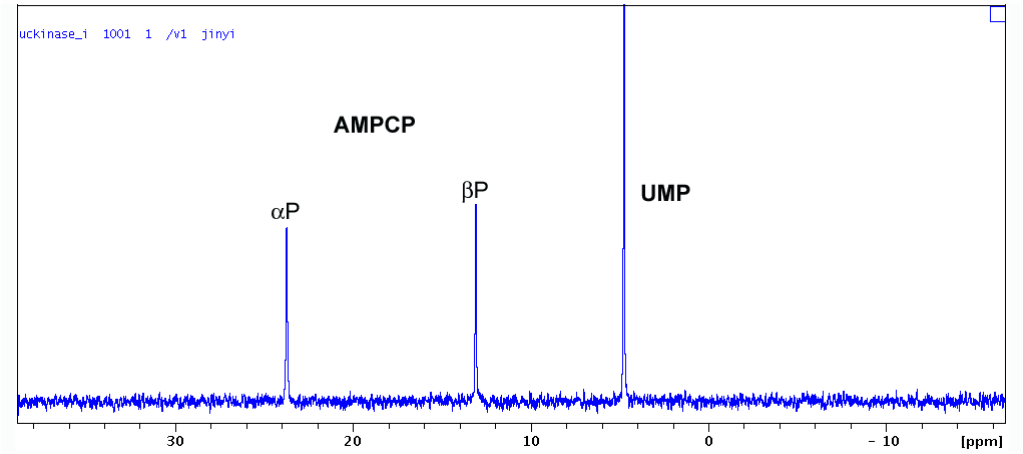


Figure 7.25 ^{31}P NMR of UMPK-UMP-AlF_4^--AMPCP TSA complex in 0.5 mM $UMPK_{dicty}$, 50 mM Tris-HCl, 5 mM $MgCl_2$, 50 mM DTT, 10 mM NH_4F, 2 mM $AlCl_3$, 2 mM UMP, and 2 mM AMPCP at pH 9.0. (Protein has largely aggregated when this spectrum was done, so binding peaks cannot be seen, but it shows both AMPCP and UMP were stable and no other nucleotides were produced in the sample.)

The success of using AMPCP to make the AlF_4^- TSA complex led us to think of using AMPCP to replace ADP to make a UMPK-UMP-BeF_3^--AMPCP GSA complex. It is noteworthy that pK_{a3} of AMPCP is much higher than for ADP, as discussed earlier in this chapter. The survey of most of the BeF_3^- complexes in the PDB indicates BeF_3^- has to coordinate to an ionised oxygen to make a stable O–Be ionic bond. Therefore trials were done at pH 7.3, 8.0 and 8.4 in the buffer of in 50 mM Tris-HCl, 5 mM $MgCl_2$, 50 mM DTT, 5 mM $BeCl_2$, 10 mM NH_4F, 0.5 mM $UMPK_{dicty}$, 2 mM AMPCP, and 2 mM UMP. However, no BeF_3^- was observed under any of the conditions investigated (**Figure 7.26**).

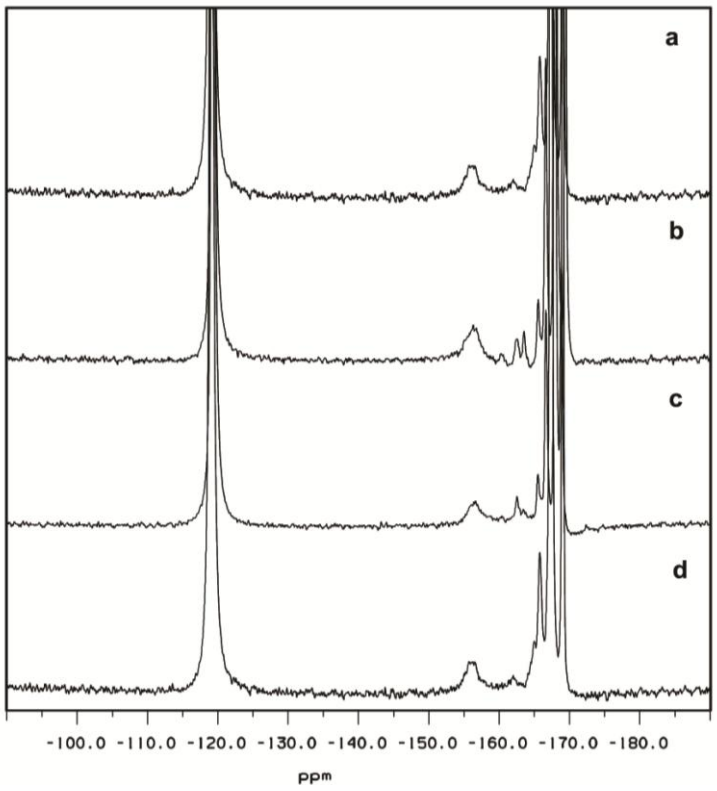


Figure 7.26 ^{19}F NMR spectra showing no BeF_3^- GSA complex is formed in 50 mM Tris-HCl, 5 mM $MgCl_2$, 50 mM DTT, 10 mM NH_4F, 5 mM $BeCl_2$, 0.5 mM $UMPK_{dicty}$, 2 mM AMPCP, 2 mM UMP at (**a**) pH 7.3; (**b**) pH 8.0; (**c**) pH 8.4; and (**d**) a control at pH 8.0 with no UMP present.

This negative outcome is in line with our previous observations that, although BeF_3^- is a more stable metal fluoride species in water than is MgF_3^-, MgF_3^- TSA and BeF_3^- GSA complexes are both very susceptible to small changes in the intended axial oxygen host compared to the more electrophilic AlF_4^-. Any subtle differences on the nature of the oxygen host such as bond length, electron density, pK_a can affect the formation of MgF_3^- TSA, and BeF_3^- GSA complexes. This has also been seen in D8E mutant of βPGM when the catalytically nucleophilic residue Asp 8 was replaced by glutamic acid: only AlF_4^- can be formed but not MgF_3^- or BeF_3^- (**Section 3.3.7**). Instead, a mutation on the residue which play a role as a catalytic GAB does not necessarily affect the formation of a BeF_3^- GSA complex but always destroys the formation of a MgF_3^- TSA complex. This has also been observed for the D10N mutant of βPGM and the D13N mutant of PSP, in which Asp 10 in βPGM and Asp 13 in PSP are both identified as catalytic general acid/bases.

7.3.7 Attempts to form metal fluoride complexes with fluorinated nucleotide analogues

Even when all the ingredients were added to the sample to an accurate final concentration, it is still difficult to compare the relative integrals of signals from different samples in the ^{19}F NMR because their intensity may vary depending on whether the active sites of the protein are saturated with substrates or if the protein has aggregated during a relatively long period of acquisition time (for ^{19}F NMR, it usually takes 1~2 d for each spectrum). Particularly in the case of the rotationally averaged peak for the UMPK-UMP-AlF_4^--ADP TSA complex, the integration of this peak compared to that of each peak in UMPK-UMP-MgF_3^--ADP was not exactly 4:1. This situation can be caused by the consumption of the substrate by a contaminative activity, the instability of $UMPK_{dicty}$ itself at rt, and the different binding constants for MgF_3^- and AlF_4^- species in the complexes. In addition to these considerations, the total amount of protein was decreasing during the lengthy pH titration course because of aggregation and denaturation. Therefore, a strongly binding substrate with a fluorine atom in it which can be used as an internal integration standard for the signals in ^{19}F NMR is really necessary.

5-Fluoro-2-deoxyuridine 5'-phosphate (5FdUMP)

Among all fluorinated UMP analogues, 5-fluorouridine 5'-phosphate (5F-UMP) should be the best candidate. It has one proton on the uracil ring replaced by fluorine with no other structural changes in the molecule. The size of fluorine is similar to hydrogen (covalent radius of 1.35 Å and 1.2 Å, respectively), so that fluorine substitution has little steric effect. The crystal structure PDB: 1QF9 with a TSA complex in the active centre shows there is plenty space around C-5 to accommodate a slightly bigger 5-fluorine atom. More importantly, when F is substituted on a double bond as in 5F-UMP, the fluorine electron density is significantly reduced by feeding back to the π-bond which makes it less polar to accept a hydrogen bond. Therefore, the substitution of fluorine for hydrogen is perhaps the most subtle type of 'single-atom mutagenesis' in the substrate for our benefit. Unfortunately, it is not commercially available. The closest commercially available candidate is 5-fluoro-2'-deoxyuridine 5'-phosphate (5FdUMP, Sigma Cat. No. F3503). This has no hydroxyl group on the 2'-carbon. In the crystal structure with wild type UMP (PDB: 1QF9), this 2'-hydroxyl has a 2.8 Å hydrogen bond with a twisted angle to the backbone carbonyl oxygen on Glu 63.

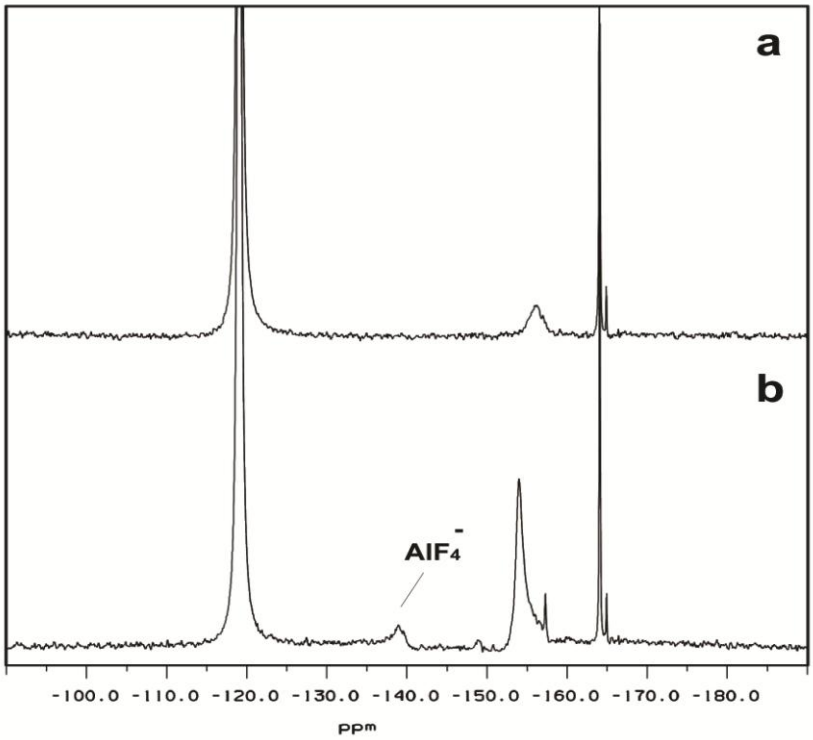


Figure 7.27 ^{19}F NMR spectra showing: (**a**) no MgF_3^- complex is formed in 50 mM Tris-HCl, 5 mM $MgCl_2$, 50 mM DTT, 10 mM NH_4F, 0.5 mM $UMPK_{dicty}$, 1 mM ADP, 1 mM 5FdUMP at pH 8.0; (**b**) a weak complex is generated at –139.0 ppm, when 2 mM $AlCl_3$ was added to sample (**a**). The small sharp peak at –165 ppm is the fluorine signal from compound 5FdUMP.

To setup the complex, 2 mM 5FdUMP and 1 mM ADP were added to 0.5 mM $UMPK_{dicty}$ in 50 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM DTT, 20 mM NH_4F, at pH 8.0. ^{19}F NMR analysis shows the signal for the 5-fluorouracil at −165 ppm but 5F-dUMP does not bind strongly enough to give a stable MgF_3^- complex (**Figure 7.27a**). Addition of aluminium chloride results in formation of a weak aluminium fluoride complex with almost the same chemical shift as the UMPK-ADP-AlF_4^--UMP complex, though the binding affinity is about 10 times reduced, judged by the relative intensity of the peaks for AlF_x and 5FdUMP (**Figure 7.27b**). This is mainly due to the missing 2'–OH and so it cannot form a hydrogen bond to the Glu 63 backbone. Evidently, this interaction is very important for $UMPK_{dicty}$ to discriminate UMP and dUMP and this could be tested in a control experiment using dUMP. Because of the time restriction, this hasn't been done and it is not a priority for us for the time being since 5FdUMP is not a good candidate as a UMP analogue for the purpose of internal integration reference.

α,β-Difluoromethylene ADP (AMPCF$_2$P)

Since 5FUMP is not available and 5FdUMP failed to deliver an internal integration standard, other fluorinated alternatives of ADP analogues were explored. Adenosine-5'-*O*-α,β-difluoromethylenediphosphate (AMPCF$_2$P) was one. AMPCF$_2$P is a stable ADP analogue with a –CF$_2$– group bridging α,β-phosphates instead of oxygen. The fluorines in the –CF$_2$– group are potentially available to form hydrogen bonds with hydrogen donors. The only adverse factor with this analogue is that the distance from the bridging carbon to the hydrogen donor is extended 1.3~1.4 Å longer because of the C–F bond. As an ADP analogue, the β-phosphate of AMPCF$_2$P has pK_{a3} 5.7, close to that of ADP at 6.4.(Hong *et al.*, 2009)

AMPCF$_2$P in its sodium salt form was generously provided by Prof. G. M. Blackburn. This compound has been tested on other systems that show it has some weak inhibitory effect on amoebae of cellular *D. discoideum*.(Rogers *et al.*, 1992) To setup the MgF_3^- complex, 2 mM AMPCF$_2$P and 2 mM UMP were added to 0.5 mM UMPK$_{dicty}$, 50 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM DTT, 20 mM NH_4F at pH 7.3, subsequently followed by adding 2 mM $AlCl_3$ to test the possibility of AlF_4^- formation.

The ^{19}F NMR shows no MgF_3^- complex was formed and only a very weak UMPK-UMP-AlF_4^--AMPCF$_2$P complex was present after addition of $AlCl_3$ (**Figure 7.28**). It is not clear why AMPCF$_2$P doesn't form a TSA complex. From the structure of UMPK-CMP-AlF_3^0-ADP, there isn't any hydrogen bond directed to the bridging oxygen, the P–O–P bond angle is 137.3°, and the P–O bond length is 1.6 Å. When AMPCF$_2$P is used, the only potential hydrogen bond partners with the fluoromethylene are the guanine of the side chain of Arg 131 which is 3.7 Å away and the backbone amide of Gly 18 (3.4 Å) which may change the binding of the phosphate moiety. According to structures of other enzymes with ADPCF$_2$P as the ligand (PDB: 3GDX and 2VR1), because the tetrahedral carbon is sp^3 hybridised forming two polarised covalent bonds with two fluorine, the P–C–P bond angle is smaller (around 110 to 120° in both structures), and the P–C bond is always 1.8 Å. These differences in bond angles and lengths must adversely influence the binding of ADPCF$_2$P.

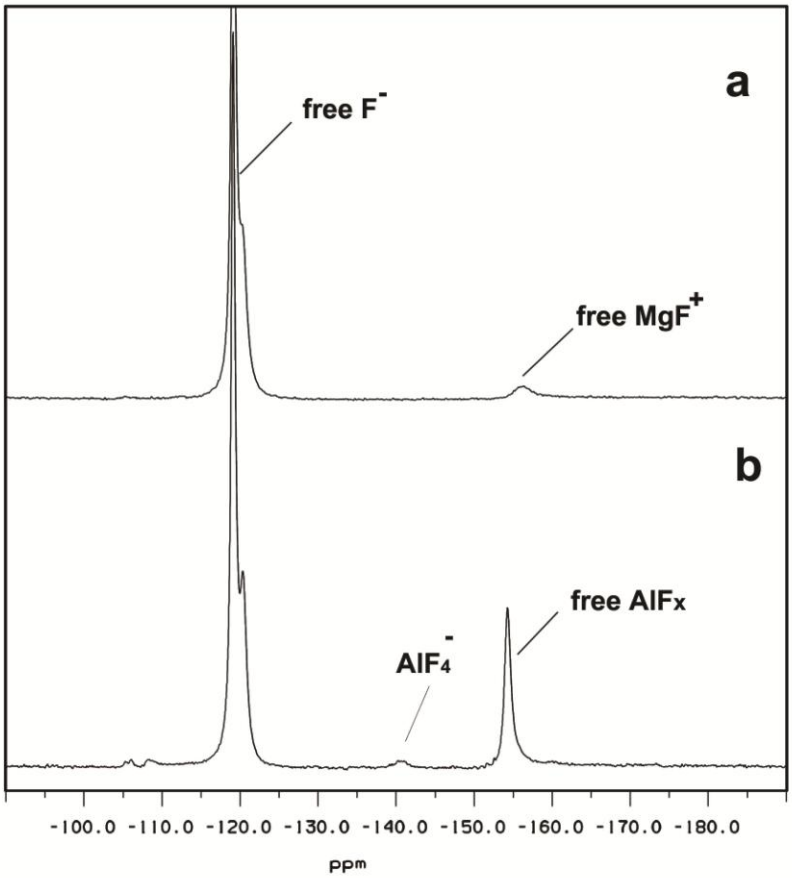


Figure 7.28 ^{19}F NMR spectra showing (**a**) no MgF_3^- complex is formed in 50 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM DTT, 20 mM NH_4F, 0.5 mM $UMPK_{dicty}$, 2 mM $AMPCF_2P$, 2 mM UMP at pH 7.3; (**b**) only a very weak AlF_4^- complex was generated at –141.0 ppm when 2 mM $AlCl_3$ was added to sample (**a**). The shoulder peak to the right of free F^- at –120 ppm is the fluorine signal from compound $AMPCF_2P$. The FP coupling is not seen because line broadening was used for data processing.

β-Fluoro-adenosine 5'-diphosphate (ADPβF)

Adenosine-5'-*O*-(2-fluorodiphosphate) (ADPβF) in its monosodium salt form was also generously provided by Prof. G. M. Blackburn. ADPβF has a non-bridging oxygen on the terminal phosphate replaced by a fluorine. Some studies have reported ADPβF as a poor ADP analogue but a good adenosine 5'-phosphosulfate (APS) analogue because ADPβF and APS both have one charge less than ADP at neutral pH.(Satishchandran *et al.*, 1992) Even so, we chose to use it for the advantage of its fluorine NMR signal.

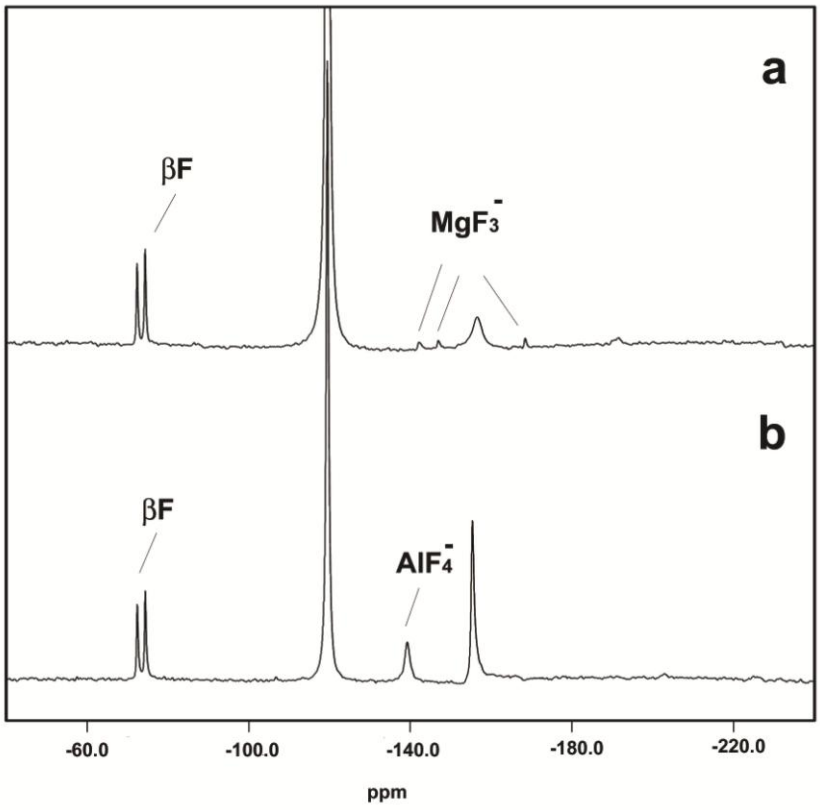


Figure 7.29 ^{19}F NMR spectra show (**a**) a small amount of UMPK-ADP-MgF_3^--UMP TSA complex was formed as highlighted during the first 7 h in the buffer of 50 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM DTT, 20 mM NH_4F, 0.5 mM $UMPK_{dicty}$, 2 mM ADPβF, 2 mM UMP at pH 7.7. (**b**) 2 mM $AlCl_3$ was added to sample (**a**) and a UMPK-ADP-AlF_4^--UMP TSA complex was formed with higher occupancy because more ADP was made during 14 h acquisition time by ADPβF hydrolysis.

^{19}F NMR shows ADPβF is not a successful substrate for $UMPK_{dicty}$ because it can be hydrolysed into ADP and AMP (supported by ^{31}P NMR, **Figure 7.30**) and eventually makes UMPK-ADP-MgF_3^--UMP and UMPK-ADP-AlF_4^--UMP TSA complexes (**Figure 7.29**).

A discussion of the possible reasons why this side reaction took place is outside the scope of this thesis. But it is noteworthy that the same hydrolytic activity was also observed in other enzymes such as small G protein Ras (Stumber *et al.*, 2002) and also for RhoA, which has been discussed in this thesis (**Chapter 6**).

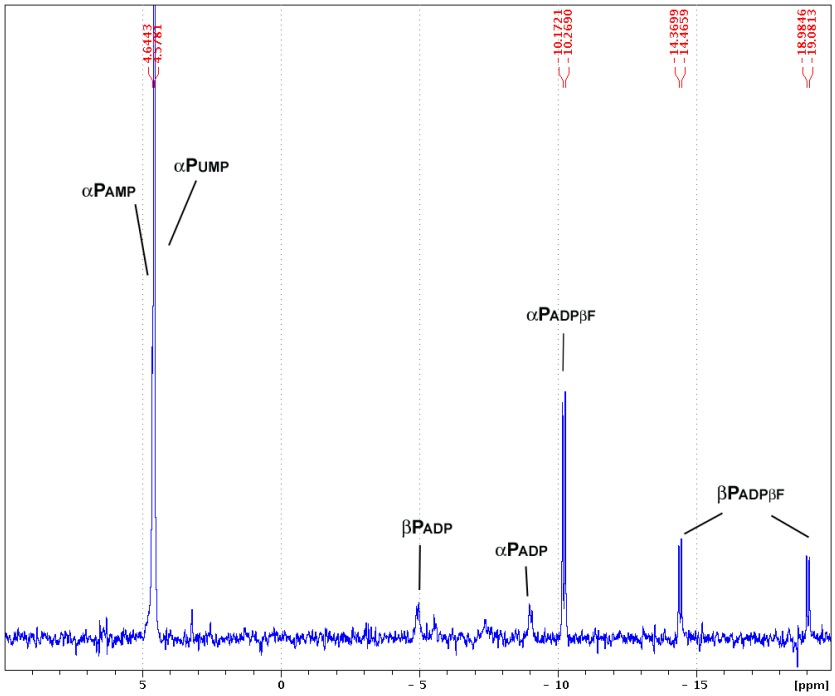


Figure 7.30 ^{31}P NMR shows some ADPβF was hydrolysed into ADP and AMP. The spectrum is recorded in the buffer of 50 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM DTT, 20 mM NH_4F, 0.5 mM $UMPK_{dicty}$, 2 mM ADPβF, 2 mM UMP at pH 7.7.

7.4 Conclusions

Based on the useful availability of published structures, $UMPK_{dicty}$ has become a good model to study phosphoryl transfer reactions between two nucleotides and broaden the list of small molecule kinases we have studied using NMR. The ^{19}F NMR results have disproved the assignment of tbp metal fluoride moiety as AlF_3^0 in structure PDB: 3UKD and in 5UKD crystallised at pH 8.5 (see **Section 7.3.1**). It has been established that MgF_3^- is the real TSA at high pH and further established that 'pH influences coordination number of the metal fluoride TSA but the charge maintains'. The pH titration revealed a new intermediate aluminium species assigned as $AlF_3(OH)^-$ not hitherto identified by crystallography. Experiments suffered from the promiscuity of $UMPK_{dicty}$ in several experiments that could not be effectively inhibited by Ap_5A or by metal fluoride complexes (see **Section 7.3.4**). A beryllium fluoride inhibitor complex was made and its NMR analysis agrees with published observations in the solid state (see **Section 7.3.3**). During the process of finding a stable ADP analogue, AMPCP proved to be a good substitute for ADP, but neither $AMPCF_2P$ or ADPβF could form trifluoromagnesate complexes and only weakly give fluoroaluminate complexes (see **Section 7.3.7**). Finally, it is shown that UMP cannot be replaced by 5FdUMP (**Section 7.3.7**).

Chapter 8. Conclusions of Thesis

^{19}F NMR as a tool for studying TSs has been further developed on all five proteins studied in this thesis. Once the protein-bound metal fluoride complexes have been generated in the active site of the enzyme, the changes of conformation and charges can be identified by changes of the chemical shift and solvent-induced isotope shifts. In the study of metal fluoride complexes as transition state analogues and ground state analogues, ^{19}F NMR is extremely valuable to examine the nature of the metal and provides the accurate number for the fluorines in the metal fluoride complexes.

This thesis also corrected errors in published assignments of trigonal bipyramidal TSA complexes designated as AlF_3 in crystal structures formed at high pH, such as the cAPK-ADP-'AlF_3^0'-SP20 TSA complex (**Section 5.3.2**) and UMPK-ADP-'AlF_3^0'-UMP TSA complex (**Section 7.3.1**). The work in this thesis established that these tbp metallofluoride moieties observed in the solid state structures are indeed MgF_3^-, irrespective of pH.

Based on the pH titration analysis of the switch between trifluoromagnesate and tetrafluoroaluminate TSAs in cAPK (**Section 5.3.4**) and $UMPK_{dicty}$ (**Section 7.3.2**), this thesis generates more sound evidence to support the Charge Balance Hypothesis which endorses the prioritisation of charge over geometry in TSAs for these phosphotransfer enzymes.

All the evidence obtained in this thesis shows that phosphoryl transfer requires apical oxygens that either are anionic or are hydrogen bonded to a catalytic base. This is reflected by the capability of forming the TSAs, especially MgF_3^- TSAs, which requires a more stringent setup in the active site of the enzyme than do fluoroaluminate complexes. So far, this work has identified no exception for wild type proteins to form MgF_3^- and AlF_4^- TSA complexes, including PSP, βPGM, hPGK, cAPK, UMP/CMP kinase, RhoA/RhoGAP. In terms of mutants, only a few MgF_3^- complexes have been found in PSP_{D13N}-MgF_3^--H_2O (**Section 4.3.2.2**) and RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ (**Section6.3.4.1**). Aluminium tetrafluoride has a much greater intrinsic stability in water, therefore TSAs containing it can be formed in wild type enzymes as well as in some of the mutants, as long as there is at least an anionic carboxylate as apical ligand, such as PSP_{D13N}-AlF_4^--H_2O complex (**Section 4.3.2.3**), βPGM_{D10N}-AlF_4^--G6P complex

(**Section 3.3.4**), βPGM_{D10N}-AlF_4^--βG1P complex (**Section 3.3.5.1**), and βPGM_{D8E}-AlF_4^--G6P complex (**Section 3.3.7**).

In this thesis, the ground state analogue trifluoroberyllate has successfully been formed in $cAPK_{wt}$ as cAPK-ADP-BeF_3^--SP20 complex (**Section 5.3.7**), in βPGM_{D10N} as βPGM_{D10N}-BeF_3^--G6P complex and βPGM_{D10N}-BeF_3^--H_2O complex (**Section 3.3.6**), and in PSP_{D10N} as PSP_{D10N}-BeF_3^--H_2O complex (**Section 4.3.3**). In $UMPK_{dicty}$, beryllium fluoride in UMPK-UDP-BeF_2^0-ADP complex mimics the γ-bridging phosphate in the bisubstrate analogue Up_5A with high affinity (**Section 7.3.3**). It is therefore concluded that the trifluoroberyllate GSAs that have been tested can only attach onto an anionic oxygen from a carboxylate or a phosphate in all these cases.

Although the attempt in this thesis to use GTPγF as the substrate for RhoA to test the Charge Balance Hypothesis failed (**Section 6.3.5**), the discovery of the different structure features of small G protein RhoA in the RhoA-GDP-MgF_3^--$RhoGAP_{R85A}$ and RhoA-GDP-AlF_4^--$RhoGAP_{R85A}$ TSA complexes has drawn our attention away from the importance of the arginine finger. This may have been over-valued in the literature for its function in terms of its positive charge contribution to the TS (**Section 6.3.4**). The hydrolysis mechanism of RhoA functioning as a GTPase may be fundamentally different as opposed to other phosphoryl transfer enzymes so far studied.

Chapter 9. Forward Look

After the study of these five phosphotransferase enzymes, we surprisingly find each protein has a special structural feature to make its metal fluoride transition state analogues behave differently. For example, MgF_3^- is more competitive with AlF_4^- for cAPK than for other proteins that have been studied. In UMP/CMP kinase, a new intermediate AlF_3OH^- species is observed which has not been seen before. In RhoA/RhoGAP, the aluminium fluoride moiety doesn't response to the deletion of positively charged arginine finger in RhoGAP. Based on the discoveries made in this thesis, we can look forward to further experiments to be attempted in the near future to round out these projects.

For UMP/CMP kinase, it would be very useful to determine the identity of the three downfield aluminium-associated peaks in ^{19}F NMR spectra. Because even high resolution crystallography cannot distinguish F^- and OH^- from electron density, a NMR method is still preferred. The use of 10-fold F^- (100 mM NH_4F) instead of 20-fold (200 mM) should be tried to see whether a larger concentration of fluorine can compete with OH^- and change the equilibrium between AlF_3OH^- and AlF_4^- complexes. An alternative approach would be to use gallium to form a metal fluoride TSA complex, as has been successfully done with βPGM. Because the radius of the Ga^{3+} ion is different from Al^{3+}, it may stop the averaged rotation of AlF_4^- resonances observed not only in UMP/CMP kinase but also in RhoA-RhoGAP and F_1ATPase. This has not been tried before.

For RhoA-RhoGAP, the unexpected absence of reciprocal response by chemical mutants to the negative charge of the fluoroaluminate moiety as a consequence of the R85A RhoGAP mutation is still mysterious. We believe it is caused by the charge distribution on Arg 85 of RhoGAP making it less positive on the εNH_2 so that the difference between Arg and Tyr is not very large. This can be examined by using ^{13}C and ^{15}N double labelled protein to acquire a 2D (H)NCO NMR spectrum. Although there are 7 arginines in the complex, 5 of them are on the surface, 1 arginine has no aromatic ring nearby, only arginine 85 stacks against the aromatic ring of Tyr 34 making the chemical shifts of the hydrogens move downfield where they are readily observable. A new mutant K18A of RhoA will be used to test the hydrolysis activity towards GTPγF.

For cAPK, the challenge now is to accomplish the whole structure of the reaction coordinates using the trifluoroberyllate ground state analogue, trifluoromagnesate transition state analogue, and a novel phosphonate analogue on Ser 17 of the SP20 substrate peptide. That is a challenge that demands novel chemical synthesis. Another approach will be to use the formation of TSA as a sensitive probe to examine the binding of inhibitory compounds at the allosteric site of protein kinases.

For βPGM, the binding constant of the two compounds βCH_2G1P and (*S*)-βCHFG1P will be evaluated by ITC so that the bonding interactions between fluorine and its surrounding hydrogens can be quantified.

Appendix I

The equations that have been used for the calculation of the binding constants of MgF_3^- and AlF_4^- (**Section 5.3.3**) by Dr.Matthew Cliff (TUOS) are:

$[MgF] = 3\times [Mg] \times [F] / K_{FMg1}$
$[MgF_2] = 3\times [MgF] \times [F] / K_{FMg2}$
$[MgF_3] = [MgF_2] \times [F] / K_{FMg3}$
$[E:MgF_3] = [MgF_3] \times [E] / K_{MgF}$

$[AlF] = [Al] \times [F] / K_{FAl1}$
$[AlF_2] = [AlF] \times [F] / K_{FAl2}$
$[AlF_3] = [AlF_2] \times [F] / K_{FAl3}$
$[AlF_4] = [AlF_3] \times [F] / K_{FAl4}$
$[AlF_5] = [AlF_4] \times [F] / K_{FAl5}$
$[E:AlF_4] = [AlF_4]\times[E] / K_{AlF}$

In which,
Peak intensity Sig_E:MgF = $SIG1\times[E:MgF_3]$
Peak intensity Sig_E:AlF = $SIG2\times[E:AlF_4]$

Fixed parameters are [Mg] = 30 mM; [E] = 0.6 mM; [F] = 30 mM

The dissociation constants of magnesium fluoride species are estimated from past literature (Baxter *et al.*, 2006) and behaviour of other metals.
$K_{FMg1} = 10$;
$K_{FMg2} = 100$;
$K_{FMg3} = 3000$;

The dissociation constants of aluminium fluoride species are estimated with considering the binding of aluminium to the excess ADP in solution (Bruce Martin, 1996):
$K_{FAl1} = 0.0004$;
$K_{FAl2} = 0.006$;
$K_{FAl3} = 0.16$;
$K_{FAl4} = 0.5$;
$K_{FAl5} = 50$

The fitted parameters as calculation results:
SIG1 = 18.43 ± 0.64264 (conversion factor from intensity to concentration)
SIG2 = 24.59 ± 0.64387 (conversion factor from intensity to concentration)

K_{MgF3} = 3.7 fM ±1.3 μM (big error)
K_{AlF4} = 840 fM ± 300uM

Appendix II

Reanalysis of the deposited electron density data for PDB: 1L3R, cAPK (Madhusudan *et al*., 2002) carried out by Dr. Matthew Bowler, ESRF, Grenoble, France.

Deposited coordinates and structure factors for 1L3R were downloaded from the PDB. The model was subjected to 10 cycles of restrained refinement using REFMAC5 (Murshudov *et al.*, 1997) with resulting *R* factors that were similar to those reported before maps were calculated. In order to compare the refinement of the proposed trigonal planar aluminium fluoride with the alternative interpretation of an MgF_3^- complex, a model containing no TSA was refined against the structure factors. MgF_3^- was then built into the resulting difference Fourier maps showing a trigonal planar species and the model subjected to a further 10 cycles of refinement. The atoms of the MgF_3^- moiety were not restrained during refinement. The resulting electron density maps are shown in **Figure 5.4**. While trifluoromagnesate fits well to the density, with bond lengths refining to those expected of this species, two additional peaks are visible in the difference maps, with peak heights of 4.5σ. These peaks correspond well to the expected positions of two fluoride atoms from a tetrafluoroaluminate complex. AlF_4^- was then placed in the density and the mixture of metal fluoride species refined against the data. By comparing atomic *B*-factor values for the metal fluoride moieties with the bound nucleotide, the occupancy of the active site was judged to contain 70% tbp MgF_3^- to 30% octahedral AlF_4^- (**Figure 5.5**)

Appendix III

Crystallisation and structure analysis of various RhoA-RhoGAP TSA complexes by Erika Pellegrini, ESRF, Grenoble, France.

Human recombinant RhoA and RhoGAP were both expressed as GST fusion proteins in *E. coli*. Rosetta2 cells and purified by affinity chromatography followed by removal of the GST tag by thrombin cleavage and size-exclusion chromatography. Part of the proteins are made in Sheffield and delivered to Erika Pellegrini. The RhoA-GDP-AlF_4-$RhoGAP_{R85A}$ complex was obtained mixing the two proteins in a 1:1 ratio, supplemented the protein buffer (50 mM Bis-Tris-HCl pH 7, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT) with 20 mM NaF and 2 mM $AlCl_3$. After overnight incubation the metallofluoride complex was concentrated to 200 uM. The same protocol was repeated for the RhoA-GDP-MgF_3-$RhoGAP_{R85A}$ complex, supplemented with 10 mM $MgCl_2$ instead of $AlCl_3$. Initially, plate crystals were obtained in sitting drop vapour diffusion experiments using equal amounts of reservoir solution (100 mM Bis Tris-HCl pH6, 21-24% PEG 3350) and a stock solution of the protein complex. These crystals were used to make a microseed stock and used to seed crystals in the same conditions in microbatch experiments under paraffin oil. The crystals obtained by evaporation were prepared for flash cooling by immersion in the cryosolution made up with 25% PEG 400. Diffraction data from RhoA-GDP-AlF_4-$RhoGAP_{R85A}$ crystals at 3 Å resolution were collected at 293 K, at the ESRF beamline ID14-4. The dataset was integrated with XDS (Kabsch, 2010). All subsequent calculations were carried out with the ccp4 program suite (MolRep and REFMAC5; Vagin *et al.*, 1997; Murshudov *et al.*, 1997). The structure of the complex was solved by molecular replacement using the previous structure (1TX4, Rittinger *et al.* 1997) as a search model with the bound ligands and water molecules removed. Ligands were included in the final rounds of refinement. The structure was refined until no peaks above 4σ in the difference Fourier map remained.

Crystallisation of RhoA-GDP was as previously described (Pellegrini *et al*., 2011). Data processing performed as described for the complex. 1FTN was selected as the search model (Graham *et al*., 2002). As aluminum fluoride was not detected in this structure, solved at 1.3 Å resolution, the structure is not described.

The RhoA-GDP-AlF_4-$RhoGAP_{R85A}$ complex crystallised in different conditions those previously described (Rittinger *et al.* 1997). Moreover, several optimizations steps were required to obtain thick plate crystals, diffracting to high resolution in both vertical and horizontal directions. The crystals used for solving the structure described here were obtained through a combination of seeding and microbatch experiments under paraffin oil. At least two or three seeding generations were required for separating needles from plates and single plates from plate clusters. The crystals grown by evaporation instead of vapor diffusion increased their size in all three dimensions and diffracted X-rays well in all directions. The metallofluoride complex crystallized in the centered orthorhombic space group $C222_1$ with one molecule in the asymmetric unit. The structure was solved to 3 Å resolution, with a final R factor of 28% and free *R* factor of 34% (**Table1**).

Table 1

Structure	**RhoA-RhoGAP-AlF_3-GDP**
Space group	$C222_1$
Wavelength	0.933 Å
Unit cell dimensions (Å) *a,b,c*	102.7, 111.9, 66.8,
Resolution range (Å)	75.6 - 3.0
Number of unique reflections	7286
Multiplicity	3.3 (3.6)
Completeness (%)	96.9 (99.4)
R_{merge}	0.11 (0.30)
$\langle I/\sigma I \rangle$	7.1 (3.7)
Wilson B factor	44.6 $Å^2$
Water molecules	202
R-factor (%)	28
Free R-factor (%)	34
RMS deviations:	
Bonds (Å)	0.01
Angles (°)	1.3

The two proteins assumed the same relative orientation as in 1TX4 with rmsd between Cα atoms of 0.5 Å. The difference Fourier maps showed clear peaks for the expected ligands (GDP and aluminum fluoride) and the absence of electron density for the $Arg85_{RhoGAP}$ side chain. Surprisingly, density for the back bone of residues 31-33 in RhoA effector loop was clearly visible. These residues are disordered in 1TX4 structure.

For RhoA-GDP-$AlF_3(H_2O)$ complex, we suggest that while in solution the protein is binding and unbinding the AlF_3, in the solid state the open conformation of the effector loop is stabilized by crystal contacts. RhoA-GDP crystallised in the orthorhombic space group $P2_12_12_1$, with one molecule in the asymmetric unit. The structure demonstrates that due to contacts between symmetry related molecules the effector loop has been

stabilized in the conformation of RhoA in 1FTN structure instead of the one described for the complex making the binding of the TSA impossible.

References

Abbott, S. J., Jones, S. R., Weinman, S. A. and Knowles, J. R. (1978) Chiral [^{16}O, ^{17}O, ^{18}O]phosphate monoesters. 1. Asymmetric synthesis and stereochemical analysis of [1(*R*)-^{16}O, ^{17}O, ^{18}O]phospho-(*S*)-propane-1,2-diol. *J. Am. Chem. Soc.* **100**: 2558-2560.

Adams, J. A. and Taylor, S. S. (1993) Divalent metal ions influence catalysis and active-site accessibility in the cAMP-dependent protein kinase. *Protein Sci.* **2**: 2177-2186.

Admiraal, S. J. and Herschlag, D. (1995) Mapping the transition state for ATP hydrolysis: implications for enzymatic catalysis. *Chem. Biol.* **2**: 729-739.

Admiraal, S. J. and Herschlag, D. (2000) The Substrate-Assisted General Base Catalysis Model for Phosphate Monoester Hydrolysis: Evaluation Using Reactivity Comparisons. *J. Am. Chem. Soc.* **122**: 2145-2148.

Akamine, P., Madhusudan, Wu, J., Xuong, N. H., Ten Eyck, L. F. and Taylor, S. S. (2003) Dynamic features of cAMP-dependent protein kinase revealed by apoenzyme crystal structure. *J. Mol. Biol.* **327**: 159-171.

Allen, K. N. (2003) Response to Comment on "The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction". *Science* **301**: 1184d.

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

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

Bao, Z. Q., Jacobsen, D. M. and Young, M. A. (2011) Briefly bound to activate: transient binding of a second catalytic magnesium activates the structure and dynamics of CDK2 kinase for catalysis. *Structure* **19**: 675-690.

Baryshnikova, O. K., Williams, T. C. and Sykes, B. D. (2008) Internal pH indicators for biomolecular NMR. *J. Biomol. NMR* **41**: 5-7.

Batiz-Hernandez, H., Bernheim, R.A. (1967) Chapter 2 The isotope shift. *Progress in NMR Spec.* **3**: 63-85.

Baumeister, M. A., Rossman, K. L., Sondek, J. and Lemmon, M. A. (2006) The Dbs PH domain contributes independently to membrane targeting and regulation of guanine nucleotide-exchange activity. *Biochem. J.* **400**: 563-572.

Baxter, N. J., Blackburn, G. M., Marston, J. P., Hounslow, A. M., Cliff, M. J., Bermel, W., Williams, N. H., Hollfelder, F., Wemmer, D. E. and Waltho, J. P. (2008) Anionic charge is prioritized over geometry in aluminum and magnesium fluoride transition state analogs of phosphoryl transfer enzymes. *J. Am. Chem. Soc.* **130**: 3952-3958.

Baxter, N. J., Bowler, M. W., Alizadeh, T., Cliff, M. J., Hounslow, A. M., Wu, B., Berkowitz, D. B., Williams, N. H., Blackburn, G. M. and Waltho, J. P. (2010) Atomic details of near-transition state conformers for enzyme phosphoryl transfer revealed by MgF_3^- rather than by phosphoranes. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 4555-4560.

Baxter, N. J., Hounslow, A. M., Bowler, M. W., Williams, N. H., Blackburn, G. M. and Waltho, J. P. (2009) MgF(3)(-) and alpha-galactose 1-phosphate in the active site of beta-phosphoglucomutase form a transition state analogue of phosphoryl transfer. *J. Am. Chem. Soc.* **131**: 16334-16335.

Baxter, N. J., Olguin, L. F., Golicnik, M., Feng, G., Hounslow, A. M., Bermel, W., Blackburn, G. M., Hollfelder, F., Waltho, J. P. and Williams, N. H. (2006) A Trojan horse transition state analogue generated by MgF_3^- formation in an enzyme active site. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 14732-14737.

Benkovic, S. J. and Schray, K. (1973) *The enzymes*.Vol 8, 3rd Edition. Academic Press, New York.

Berges, J. A., Fisher, A.E., and Harrison, P.J. (1993) A comparison of Lowry, Bradford and Smith protein using different protein standards and protein isolated from the marine diatom Thalassiosira pseudonana. *Marine Biol.* **115**: 187-193.

Bernasconi, C. F. (1992) The principle of non-perfect synchronization. *Adv. Phys. Org. Chem.* **27**: 119-238

Bernheim, R. A. and Batiz Hernandez, H. (1964) Indirect Nuclear Spin-Spin Coupling and Isotope Shifts in the Nuclear Magnetic Resonance of $^{15}NH_3$, $^{15}NH_2D$, and $^{15}NHD_2$. *J. Chem. Phys.* **40**: 3446.

Bjarnadottir, U. and Nielsen, J. E. (2010) Calculating pK_a values in the cAMP-dependent protein kinase: the effect of conformational change and ligand binding. *Prot. Sci.* **19**: 2485-2497.

Blackburn, G. M. and Brown, M. J. (1969) Mechanism of hydrolysis of diethyl 2-carboxyphenylphosphonate. *J. Am. Chem. Soc.* **91**: 525-526.

Blackburn, G. M., England, D. A. and Kolkmann, F. (1981) Monofluoro- and difluoro-methylenebisphosphonic acids: isopolar analogues of pyrophosphoric acid. *J. Chem. Soc., Chem. Commun.* 930-932.

Blackburn, G. M., Williams, N. H., Gamblin, S. J. and Smerdon, S. J. (2003) Comment on "The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction". *Science* **301**: 1184.

Blattler, W. A. and Knowles, J. R. (1979) Stereochemical course of phosphokinases. The use of adenosine [γ-(S)-16O,17O,18O]triphosphate and the mechanistic consequences for the reactions catalysed by glycerol kinase, hexokinase, pyruvate kinase, and acetate kinase. *Biochem.* **18**: 3927-3933.

Boguski, M. S. and McCormick, F. (1993) Proteins regulating Ras and its relatives. *Nature* **366**: 643-654.

Bos, J. L., Rehmann, H. and Wittinghofer, A. (2007) GEFs and GAPs: critical elements in the control of small G proteins. *Cell* **129**: 865-877.

Boureux, A., Vignal, E., Faure, S. and Fort, P. (2007) Evolution of the Rho family of ras-like GTPases in eukaryotes. *Mol. Biol. Evol.* **24**: 203-216.

Bowler, M. W., Cliff, M. J., Waltho, J. P. and Blackburn, G. M. (2010) Why did Nature select phosphate for its dominant roles in biology? *New J. Chem.* **34**: 784-789.

Britt, B. M. (1997) For enzymes, bigger is better. *Biophys. Chem.* **69**: 63-70.

Bruce Martin, R. (1996) Ternary complexes of Al^{3+} and F^- with a third ligand. *Coord. Chem. Rev.* **149**: 23-32.

Burley, S. K. and Petsko, G. A. (1986) Amino-aromatic interactions in proteins. *FEBS Lett.* **203**: 139-143.

Byeon, L., Shi, Z., *et al.* (1995) Mechanism of adenylate kinase. The "essential lysine" helps to orient the phosphates and the active site residues to proper conformations. *Biochemistry* **34**: 3172-3182.

Cavanagh, J., Fairbrother, W. J., Palmer, A. G., Rance, M. and Skelton, N. J. (2007) Protein NMR Spectroscopy, 2nd Edition. Elsevier Academic Press Publications. ISBN: 978-0-12-164491-8

Caverzasio, J., Imai, T., *et al.* Caverzasio, J., Imai, T., Ammann, P., Burgener, D. and Bonjour, J. P. (1996) Aluminum potentiates the effect of fluoride on tyrosine phosphorylation and osteoblast replication in vitro and bone mass in vivo. *J. Bone Miner. Res.* **11**: 46-55.

Caverzasio, J., Palmer, G. and Bonjour, J. P. (1998) Fluoride: Mode of Action. *Bone* **22**: 585-589.

Chabre, M. (1990) Aluminofluoride and beryllofluoride complexes: new phosphate analogs in enzymology. *Trends in Biochem.l Sci.* **15**: 6-10.

Chen, C. C., Smith, D. L., Bruegger, B. B., Halpern, R. M. and Smith, R. A. (1974) Occurrence and distribution of acid-labile histone phosphates in regenerating rat liver. *Biochemistry* **13**: 3785-3789.

Cheng, Y., Zhang, Y. and McCammon, J. A. (2005) How does the cAMP-dependent protein kinase catalyze the phosphorylation reaction: an ab initio QM/MM study. *J. Am. Chem. Soc.* **127**: 1553-1562.

Cherepanov, A. V., Doroshenko, E. V., Matysik, J., de Vries, S. and de Groot, H. J. (2008) The associative nature of adenylyl transfer catalyzed by T4 DNA ligase. *Proc. Natl. Acad. Sci. U. S. A.* **105**: 8563-8568.

Cherfils, J. and Chardin, P. (1999) GEFs: structural basis for their activation of small GTP-binding proteins. *Trends Biochem. Sci.* **24**: 306-311.

Cleland, W. W. and Hengge, A. C. (2006) Enzymatic Mechanisms of Phosphate and Sulfate Transfer. *Chem. Rev. (Washington, DC, United States)* **106**: 3252-3278.

Cleland, W. W. (2007) Use of isotope effects to determine enzyme mechanisms. *J. Label Compd Radiopharm.* **50**: 1006-1015.

Cliff, M. J., Bowler, M. W., Varga, A., Marston, J. P., Szabo, J., Hounslow, A. M., Baxter, N. J., Blackburn, G. M., Vas, M. and Waltho, J. P. (2010) Transition state analogue structures of human phosphoglycerate kinase establish the importance of charge balance in catalysis. *J. Am. Chem. Soc.* **132**: 6507-6516.

Coleman, J. E. (1992) Structure and Mechanism of Alkaline Phosphatase. *Annu. Rev. Biophys. Biomol. Struct.* **21**: 441-483.

Coleman, D. E., Berghuis, A. M., Lee, E., Linder, M. E., Gilman, A. G. and Sprang, S. R. (1994) Structures of active conformations of $G_{i\alpha1}$ and the mechanism of GTP hydrolysis. *Science* **265**: 1405-1412.

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

Dahnke, T. and Tsai, M. D. (1994) Mechanism of adenylate kinase. The conserved aspartates 140 and 141 are important for transition state stabilization instead of substrate-induced conformational changes. *J. Biol. Chem.* **269**: 8075-8081.

Danielson, M. A. and Falke, J. J. (1996) Use of 19F NMR to probe protein structure and conformational changes. *Annu. Rev. Biophys. Biomol. Struct.* **25**: 163-195.

Das, R., Esposito, V., Abu-Abed, M., Anand, G. S., Taylor, S. S. and Melacini, G. (2007) cAMP activation of PKA defines an ancient signaling mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **104**: 93-98.

Denu, J. M., Stuckey, J. A., Saper, M. A. and Dixon, J. E. (1996) Form and Function in Protein Dephosphorylation. *Cell* **87**: 361-364.

Di Sabato, G. and Jencks, W. P. (1961) Mechanism and Catalysis of Reactions of Acyl Phosphates. I. Nucleophilic Reactions1. *J. Am. Chem. Soc.* **83**: 4393-4400.

Dunitz, J. D. and Taylor, R. (1997) Organic Fluorine Hardly Ever Accepts Hydrogen Bonds. *Chem. –Eur. J.* **3**: 89-98.

Dzeja, P. P. and Terzic, A. (2003) Phosphotransfer networks and cellular energetics. *J. Exp. Biol.* **206**: 2039-2047.

Edelman, A. M., Blumenthal, D. K. and Krebs, E. G. (1987) Protein Serine/Threonine Kinases. *Annu.l Rev. Biochem.* **56**: 567-613.

Evers, A., Hancock, R. D., Martell, A. E. and Motekaitis, R. J. (1989) Metal ion recognition in ligands with negatively charged oxygen donor groups. Complexation of iron(III), gallium(III), indium(III), aluminum(III), and other highly charged metal ions. *Inorg. Chem.* **28**: 2189-2195.

Fastrez, J. and Fersht, A. R. (1973) Demonstration of the acyl-enzyme mechanism for the hydrolysis of peptides and anilides by chymotrypsin. *Biochemistry* **12**: 2025-2034.

Feldhaus, P., Frohlich T., Goody, R. S., Isakov, M., and Schirmer, R. H. (1975) Synthetic inhibitors of adenylate kinases in the assays for ATPases and phosphokinases. *Eur. J. Biochem.* 57: 197-204.

Ferguson, K. M., Higashijima, T., Smigel, M. D. and Gilman, A. G. (1986) The influence of bound GDP on the kinetics of guanine nucleotide binding to G proteins. *J. Biol. Chem.* **261**: 7393-7399.

Fersht, A. (1985) *Enzyme structure and mechanism*, 2nd Edition, W.H. Freeman and Company: ISBN 0-7167-1614-3.

Fersht, A. (1999) *Structure and mechanism in protein science*, 1st Edition, W.H. Freeman and Company: ISBN 0-7167-3268-8.

Fidyk, N. J. and Cerione, R. A. (2002) Understanding the catalytic mechanism of GTPase-activating proteins: demonstration of the importance of switch domain stabilization in the stimulation of GTP hydrolysis. *Biochemistry* **41**: 15644-15653.

Fischer, E. (1894) Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der deutschen chemischen Gesellschaft* **27**: 2985-2993.

Fothergill-Gilmore, L. A. and Watson, H. C. (2006) The Phosphoglycerate Mutases. *Advances in Enzymology and Related Areas of Molecular Biology*, John Wiley & Sons, Inc.**:** 227-313.

Frey, P. A. and Hegeman, A. D. (2007) *Enzymatic reaction mechanisms*. Oxford University Press, Oxford. ISBN: 0-19-512258-5.

Findlay, D., Herries, D. G., Mathias, A. P., Rabin, B. R. and Ross, C. A. (1961) The active site and mechanism of action of bovine pancreatic ribonuclease. *Nature* **190**: 781-784.

Gazzano, E., Bergandi, L., Riganti, C., Aldieri, E., Doublier, S., Costamagna, C., Bosia, A. and Ghigo, D. (2010) Fluoride effects: the two faces of janus. *Curr. Med. Chem.* **17**: 2431-2441.

Gerig, J. T. (1989) Fluorine nuclear magnetic resonance of fluorinated ligands. *Methods Enzymol.* **177**: 3-23.

Glynn, I. M. (1985) The Na^+,K^+-transporting adenosine triphosphatase. *The Enzymes of Biological Membranes.* Plenum Press, New York. 35–114

Golicnik, M., Olguin, L. F., Feng, G., Baxter, N. J., Waltho, J. P., Williams, N. H. and Hollfelder, F. (2009) Kinetic analysis of beta-phosphoglucomutase and its inhibition by magnesium fluoride. *J. Am. Chem. Soc.* **131**: 1575-1588.

Goudreau, P. N., Lee, P. J. and Stock, A. M. (1998) Stabilization of the phospho-aspartyl residue in a two-component signal transduction system in Thermotoga maritima. *Biochemistry* **37**: 14575-14584.

Graham, D. L., Eccleston, J. F., Chung, C. W. and Lowe, P. N. (1999) Magnesium fluoride-dependent binding of small G proteins to their GTPase-activating proteins. *B Biochemistry* **38**: 14981-14987.

Graham, D. L., Eccleston, J. F. and Lowe, P. N. (1999) The conserved arginine in rho-GTPase-activating protein is essential for efficient catalysis but not for complex formation with Rho.GDP and aluminum fluoride. *Biochemistry* **38**: 985-991.

Graham, D. L., Lowe, P. N., Grime, G. W., Marsh, M., Rittinger, K., Smerdon, S. J., Gamblin, S. J. and Eccleston, J. F. (2002) MgF_3^- as a transition state analog of phosphoryl transfer. *Chem. Biol.* **9**: 375-381.

Graves, J. D. and Krebs, E. G. (1999) Protein Phosphorylation and Signal Transduction. *Pharmacol. Ther.* **82**: 111-121.

Grasby,J.A. (1998) Ribozymes. Comprehensive Biological Catalysis. Vol. 1. Reaction of Electrophilic Carbon, Phosphorus and Sulfur. Academic Press, London. pp.563-571.

Gutowsky, H. S. (1959) Isotope Effects in High Resolution NMR Spectroscopy, *J. Chem. Phys.* **31**: 1683.

Haldane, J. B. S. (1930) *Enzymes*, MIT Press: ISBN 0-262-58003-9.

Hanks, S., Quinn, A. and Hunter, T. (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241**: 42-52.

Hansen, P. E., Dettman, H. D. and Sykes, B. D. (1985) Solvent-induced deuterium isotope effects on ^{19}F chemical shifts of some substituted fluorobenzenes. Formation of inclusion complexes. *J. Magn. Reson.* **62**: 487-496.

Harms, M. J., Schlessman, J. L., Sue, G. R. and Garcia-Moreno, E. B. (2011) Arginine residues at internal positions in a protein are always charged. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 18954-18959.

Herberg, F. W., Bell, S. M. and Taylor, S. S.. (1993) Expression of the catalytic subunit of cAMP-dependent protein kinase in *Escherichia coli*: multiple isozymes reflect different phosphorylation states. *Protein Eng.* **6**: 771-777.

Hengge,A.C. (1998) Transfer of the PO_3^{2-} group. Comprehensive Biological Catalysis. Vol. 1. Reaction of Electrophilic Carbon, Phosphorus and Sulfur. Academic Press, London. pp.517-542.

Hengge, A. C. (2002) Isotope effects in the study of phosphoryl and sulfuryl transfer reactions. *Acc. Chem. Res.* **35**: 105-112.

Higashijima, T., Ferguson, K. M., Sternweis, P. C., Ross, E. M., Smigel, M. D. and Gilman, A. G. (1987) The effect of activating ligands on the intrinsic fluorescence of guanine nucleotide-binding regulatory proteins. *J. Biol. Chem.* **262**: 752-756.

Higashijima, T., Graziano, M. P., Suga, H., Kainosho, M. and Gilman, A. G. (1991) ^{19}F and ^{31}P NMR spectroscopy of G protein alpha subunits. Mechanism of activation by Al^{3+} and F^-. *J. Biol. Chem.* **266**: 3396-3401.

Hoffman, G. R., Nassar, N., Oswald, R. E. and Cerione, R. A. (1998) Fluoride activation of the Rho family GTP-binding protein Cdc42Hs. *J. Biol. Chem.* **273**: 4392-4399.

Hong, J. A., Bhave, D. P. and Carroll, K. S. (2009) Identification of critical ligand binding determinants in Mycobacterium tuberculosis adenosine-5'-phosphosulfate reductase. *J. Med. Chem.* **52**: 5485-5495.

Hunter, T. and Cooper, J. A. (1985) Protein-tyrosine kinases. *Annu. Rev. Biochem.* **54**: 897-930.

Huse, M. and Kuriyan, J. (2002) The conformational plasticity of protein kinases. *Cell* **109**: 275-282.

Hutter, M. C. and Helms, V. (1999) Influence of key residues on the reaction mechanism of the cAMP-dependent protein kinase. *Protein Sci.* **8**: 2728-2733.

Hutter, M. C. and Helms, V. (2000) Phosphoryl transfer by a concerted reaction mechanism in UMP/CMP-kinase. *Protein Sci.* **9**: 2225-2231.

Ihara, K., Muraguchi, S., Kato, M., Shimizu, T., Shirakawa, M., Kuroda, S., Kaibuchi, K. and Hakoshima, T. (1998) Crystal structure of human RhoA in a dominantly active form complexed with a GTP analogue. *J. Biol. Chem.* **273**: 9656-9666.

Ihnat, P. M., Vennerstrom, J. L. and Robinson, D. H. (2002) Solution equilibria of deferoxamine amides. *J. Pharm. Sci.* **91**: 1733-1741.

Imai, Y. N., Inoue, Y., Nakanishi, I. and Kitaura, K. (2009) Amide–π interactions between formamide and benzene. *J. Comput. Chem.* **30**: 2267-2276.

Jencks, W. P. (1987) *Catalysis in chemistry and enzymology*. Dover publication, Inc. New York. ISBN 0-486-65460-5.

Jones, P. G., Kirby, A. J., and Pilkington, M. (2002) 2,2,2-Tris(cyclohexyloxy)-4,5-(2',2''-biphenylo)-1,3,2-dioxaphospholene. *Acta Cryst.* **E58**, o268-o269

Jin, Y., Cliff, M. J., Baxter, N. J., Dannatt, H. R. W., Hounslow, A. M., Bowler, M. W., Blackburn, G. M. and Waltho, J. P. (submitted) Trifluoromagnesate and tetrafluoroaluminate transition state analogue complexes support charge balance operating in cyclic AMP-dependent protein kinase. *J. Am. Chem. Soc.* **submitted**.

Jones, S. R., Kindman, L. A. and Knowles, J. R. (1978) Stereochemistry of phosphoryl group transfer using a chiral [^{16}O, ^{17}O, ^{18}O] stereochemical course of alkaline phosphatase. *Nature* **275**: 564-565.

Johnson, K. A. (2008) Role of induced fit in enzyme specificity: a molecular forward/reverse switch. *J. Biol. Chem.* **283**: 26297-26301.

Johnson, L. N. (2009) The regulation of protein phosphorylation. *Biochem. Soc. Trans.* **37**: 627-641.

Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F. and Holmes, K. C. (1990) Atomic structure of the actin:DNase I complex. *Nature* **347**: 37-44.

Kalckar, H. M. (1943) The role of myokinase in transphosphorylations: ii. The enzymatic action of myokinase on adenine nucleotides. *J. Biol. Chem.* **148**: 127-137.

Kanazawa, Y., Baldeschwieler, J. D. and Craig, N. C. (1965) NMR and double resonance spectra of the deuterodifluoroethylenes. *J. Mol. Spec.* **16**: 325-348.

Kaplan, J. H. (2002) Biochemistry of Na^+,K^+-ATPase. *Annu. Rev. Biochem.* **71**: 511-535.

Kaur, T., Bijarnia, R. K. and Nehru, B. (2009) Effect of concurrent chronic exposure of fluoride and aluminum on rat brain. *Drug Chem. Toxicol.* **32**: 215-221.

Kennedy, E. J., Yang, J., Pillus, L., Taylor, S. S. and Ghosh, G. (2009) Identifying critical non-catalytic residues that modulate protein kinase A activity. *PLoS One* **4**: e4746.

Knighton, D. R., Zheng, J. H., Ten Eyck, L. F., Xuong, N. H., Taylor, S. S. and Sowadski, J. M. (1991) Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**: 414-420.

Knowles, J. R. (1980) Enzyme-catalyzed phosphoryl transfer reactions. *Annu. Rev. Biochemistry* **49**: 877-919.

Krebs, H. A. and Hems, R. (1955) Phosphate-transfer reactions of adenosine and inosine nucleotides. *Biochem. J.* **61**: 435-441.

Lad, C., Williams, N. H. and Wolfenden, R. (2003) The rate of hydrolysis of phosphomonoester dianions and the exceptional catalytic proficiencies of protein and inositol phosphatases. *Proc. Natl. Acad. Sci. U. S. A.* **100**: 5607-5610.

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

Laidler, K. J., and M. Christine King (1983) Development of transition-state theory. *J. Phys. Chem.* **87**: 2657-2664

Lassila, J. K., Zalatan, J. G. and Herschlag, D. (2011) Biological phosphoryl-transfer reactions: understanding mechanism and catalysis. *Annu. Rev. Biochem.* **80**: 669-702

Li, L. (2003) The biochemistry and physiology of metallic fluoride: action, mechanism, and implications. *Crit. Rev. Oral Biol. Med.* **14**: 100-114.

Lowe, G., Potter, B. V. L., Sproat, B. S. and Hull, W. E. (1979) The effect of ^{17}O and the magnitude of the ^{18}O-isotope shift in ^{31}P nuclear magnetic resonance spectroscopy. *J. Chem. Soc., Chem. Commun.,*.733-735.

Lowe, G., Cullis, P. M., Jarvest, R. L., Potter, B. V. L. and Sproat, B. S. (1981) Stereochemistry of Phosphoryl Transfer. *Philos. T. Roy. Soc. B, Biol. Sci.* **293**: 75-92.

Ma, J. C. and Dougherty, D. A. (1997) The Cationminus signpi Interaction. *Chem. Rev.* **97**: 1303-1324.

Madhusudan, Akamine, P., Xuong, N. H. and Taylor, S. S. (2002) Crystal structure of a transition state mimic of the catalytic subunit of cAMP-dependent protein kinase. *Nat. Struct. Biol.* **9**: 273-277.

Madhusudan, Trafny, E. A., Xuong, N. H., Adams, J. A., Ten Eyck, L. F., Taylor, S. S. and Sowadski, J. M. (1994) cAMP-dependent protein kinase: crystallographic insights into substrate recognition and phosphotransfer. *Protein Sci.* **3**: 176-187.

Manning, G., Whyte, D. B., Martinez, R., Hunter, T. and Sudarsanam, S. (2002) The protein kinase complement of the human genome. *Science* **298**: 1912-1934.

Masterson, L. R., Cheng, C., Yu, T., Tonelli, M., Kornev, A., Taylor, S. S. and Veglia, G. (2010) Dynamics connect substrate recognition to catalysis in protein kinase A. *Nat. Chem. Biol.* **6**: 821-828.

Masterson, L. R., Mascioni, A., Traaseth, N. J., Taylor, S. S. and Veglia, G. (2008) Allosteric cooperativity in protein kinase A. *Proc. Natl. Acad. Sci. U. S. A.* **105**: 506-511.

Mildvan, A. S. (1997) Mechanisms of signaling and related enzymes. *Proteins* **29**: 401-416.

Miles, R. D., Gorrell, A. and Ferry, J. G. (2002) Evidence for a transition state analog, MgADP-aluminum fluoride-acetate, in acetate kinase from *Methanosarcina thermophila*. *J. Biol. Chem.* **277**: 22547-22552.

Moré, J. J., (1978) “The Levenberg-Marquardt algorithm: Implementation and theory,” Reference Lecture Notes in Mathematics, Springer, **630**/1978, 105-116. DOI: 10.1007/BFb0067700)

Muller, K., Faeh, C. and Diederich, F. (2007) Fluorine in pharmaceuticals: looking beyond intuition. *Science* **317**: 1881-1886.

Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta. Cryst. Sect. D* **53**: 240-255.

Nassar, N., Hoffman, G. R., Manor, D., Clardy, J. C. and Cerione, R. A. (1998) Structures of Cdc42 bound to the active and catalytically compromised forms of Cdc42GAP. *Nat. Struct. Biol.* **5**: 1047-1052.

Neves, S. R., Ram, P. T. and Iyengar, R. (2002) G Protein Pathways. *Science* **296**: 1636-1639.

"The Nobel Prize in Physiology or Medicine 1992". Nobelprize.org. 14 Dec 2011 http://www.nobelprize.org/nobel_prizes/medicine/laureates/1992/

Oksanen, E., Ahonen, A. K., Tuominen, H., Tuominen, V., Lahti, R., Goldman, A. and Heikinheimo, P. (2007) A complete structural description of the catalytic cycle of yeast pyrophosphatase. *Biochemistry* **46**: 1228-1239.

Orr, G. A., Simon, J., Jones, S. R., Chin, G. J. and Knowles, J. R. (1978) Adenosine 5'-O-([γ-18O]gamma-thio)triphosphate chiral at the γ-phosphorus: stereochemical consequences of reactions catalyzed by pyruvate kinase, glycerol kinase, and hexokinase. *Proc. Natl. Acad. Sci. U. S. A.* **75**: 2230-2233.

Pan, J. Y. and Wessling-Resnick, M. (1998) GEF-mediated GDP/GTP exchange by monomeric GTPases: a regulatory role for Mg^{2+}? *Bioessays* **20**: 516-521.

Pauling, L. (1948) Nature of forces between large molecules of biological interest. *Nature* **161**: 707-709.

Pfeffer, S. R., Dirac-Svejstrup, A. B. and Soldati, T. (1995) Rab GDP dissociation inhibitor: putting rab GTPases in the right place. *J. Biol. Chem.* **270**: 17057-17059.

Praefcke, G. J., Geyer, M., Schwemmle, M., Robert Kalbitzer, H. and Herrmann, C. (1999) Nucleotide-binding characteristics of human guanylate-binding protein 1 (hGBP1) and identification of the third GTP-binding motif. *J. Mol. Biol.* **292**: 321-332.

Ramsey, N. F. (1952) Vibrational and centrifugal effects on nuclear interactions and rotational moments in molecules. *Phys. Rev.* **87**: 1075.

Re, S., Imai, T., Jung, J., Ten-No, S. and Sugita, Y. (2011) Geometrically associative yet electronically dissociative character in the transition state of enzymatic reversible phosphorylation. *J. Comput. Chem.* **32**: 260-270.

Reinstein, J., Schlichting, I. and Wittinghofer, A. (1990) Structurally and catalytically important residues in the phosphate binding loop of adenylate kinase of *Escherichia coli*. *Biochemistry* **29**: 7451-7459.

Rice, W. J., Young, H. S., Martin, D. W., Sachs, J. R. and Stokes, D. L. (2001) Structure of Na^+,K^+-ATPase at 11-Å resolution: comparison with Ca^{2+}-ATPase in E1 and E2 states. *Biophys. J.* **80**: 2187-2197.

Ridder, I. S. and Dijkstra, B. W. (1999) Identification of the Mg^{2+}-binding site in the P-type ATPase and phosphatase members of the HAD (haloacid dehalogenase) superfamily by structural similarity to the response regulator protein CheY. *Biochem. J.* **339 (Pt 2)**: 223-226.

Ridley, A. J. and Hall, A. (1992) The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. *Cell* **70**: 389-399.

Rigas, J. D., Hoff, R. H., Rice, A. E., Hengge, A. C. and Denu, J. M. (2001) Transition State Analysis and Requirement of Asp-262 General Acid/Base Catalyst for Full Activation of Dual-Specificity Phosphatase MKP3 by Extracellular Regulated Kinase. *Biochemistry* **40**: 4398-4406.

Rittinger, K., Walker, P. A., Eccleston, J. F., Smerdon, S. J. and Gamblin, S. J. (1997) Structure at 1.65 Å of RhoA and its GTPase-activating protein in complex with a transition-state analogue. *Nature* **389**: 758-762.

Rogers, M. J., Russell, R. G. G., Blackburn, G. M., Williamson, M. P. and Watts, D. J. (1992) Metabolism of halogenated bisphosphonates by the cellular slime mould *Dictyostelium discoideum*. *Biochem. Biophys. Res. Comm.* **189**: 414-423.

Rossman Rossman, K. L., Worthylake, D. K., Snyder, J. T., Siderovski, D. P., Campbell, S. L. and Sondek, J. (2002) A crystallographic view of interactions between Dbs and Cdc42: PH domain-assisted guanine nucleotide exchange. *Embo. J.* **21**: 1315-1326.

Salopek-Sondi, B. and Luck, L. A. (2002) ^{19}F NMR study of the leucine-specific binding protein of *Escherichia coli*: mutagenesis and assignment of the 5-fluorotryptophan-labeled residues. *Protein Eng.* **15**: 855-859.

Saraste, M., Sibbald, P. R. and Wittinghofer, A. (1990) The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* **15**: 430-434.

Satishchandran, C., Myers, C. B. and Markham, G. D. (1992) Adenosine-5'-O-(2-fluorodiphosphate) (ADPβF), an analog of adenosine-5'-phosphosulfate *Bioorg. Chem.* **20**: 107-114.

Scheffzek, K., Ahmadian, M. R., Kabsch, W., Wiesmuller, L., Lautwein, A., Schmitz, F. and Wittinghofer, A. (1997) The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* **277**: 333-338.

Scheffzek, K., Kliche, W., Wiesmuller, L. and Reinstein, J. (1996) Crystal structure of the complex of UMP/CMP kinase from *Dictyostelium discoideum* and the bisubstrate inhibitor P1-(5'-adenosyl) P5-(5'-uridyl) pentaphosphate (UP_5A) and Mg^{2+} at 2.2 Å: implications for water-mediated specificity. *Biochemistry* **35**: 9716-9727.

Schlichting, I. and Reinstein, J. (1997) Structures of active conformations of UMP kinase from *Dictyostelium discoideum* suggest phosphoryl transfer is associative. *Biochemistry* **36**: 9290-9296.

Schlichting, I. and Reinstein, J. (1999) pH influences fluoride coordination number of the AlF_x phosphoryl transfer transition state analog. *Nat. Struct. Biol.* **6**: 721-723.

Schramm, V. L. (1998) Enzymatic transition states and transition state analog design. *Annu. Rev. Biochem.* **67**: 693-720.

Schramm, V. L. (2007) Enzymatic transition state theory and transition state analogue design. *J. Biol. Chem.* **282**: 28297-28300.

Shad, K. F. (2006) How Aluminium Fluoride causes Dementia. The annual meeting of the Society for Neuroscience. Atlanta, USA.

Shankar, S., Ye, R. W., Schlictman, D. and Chakrabarty, A. M. (1995) Exopolysaccharide alginate synthesis in Pseudomonas aeruginosa: enzymology and regulation of gene expression. *Adv. Enzymol. R.A.M.B.* **70**: 221-255.

Smtih, D. L., Bruegger, B. B., Halpern, R. M. and Smith, R. A. (1973) New histone kinases in nuclei of rat tissues. *Nature* **246**: 103-104.

Sondek, J., Lambright, D. G., Noel, J. P., Hamm, H. E. and Sigler, P. B. (1994) GTPase mechanism of Gproteins from the 1.7-Å crystal structure of transducin alpha-GDP-AlF_4^-. *Nature* **372**: 276-279.

Sosnicki, J. G., Langaard, M. and Hansen, P. E. (2007) Long-range deuterium isotope effects on ^{13}C chemical shifts of intramolecularly hydrogen-bonded N-substituted 3-(cycloamine)thiopropionamides or amides: a case of electric field effects. *J. Org. Chem.* **72**: 4108-4116.

Spotswood, T., Evans, J. M. and Richards, J. H. (1967) Enzyme-substrate interaction by nuclear magnetic resonance. *J. Am. Chem. Soc.* **89**: 5052-5054.

Sträter, N., Lipscomb, W. N., Klabunde, T. and Krebs, B. (1996) Two-Metal Ion Catalysis in Enzymatic Acyl- and Phosphoryl-Transfer Reactions. *Angew. Chem. Int. Ed.Engl.* **35**: 2024-2055.

Steitz,T.A. (1993) DNA dependent and RNA dependent DNA polymerases. *Curr. Opin. Struct. Biol.* **3**: 31-38.

Caldwell, S. R., Newcomb, J. R., Schlecht, K. A. and Raushel, F. M. (1991) Limits of diffusion in the hydrolysis of substrates by the phosphotriesterase from Pseudomonas diminuta. *Biochemistry* **30**: 7438-7444.

Stewart, L., Redinbo, M. R., Qiu, X., Hol, W. G. and Champoux, J. J. (1998) A model for the mechanism of human topoisomerase I. *Science* **279**: 1534-1541.

Stockbridge, R. B. and Wolfenden, R. (2009) The Intrinsic Reactivity of ATP and the Catalytic Proficiencies of Kinases Acting on Glucose, *N*-Acetylgalactosamine, and Homoserine. *J. Biol. Chem.* **284**: 22747-22757.

Stumber, M., Herrmann, C., Wohlgemuth, S., Kalbitzer, H. R., Jahn, W. and Geyer, M. (2002) Synthesis, characterization and application of two nucleoside triphosphate analogues, GTPgammaNH(2) and GTPgammaF. *Eur. J. Biochemistry* **269**: 3270-3278.

Sudom, A. M., Prasad, L., Goldie, H. and Delbaere, L. T. (2001) The phosphoryl-transfer mechanism of *Escherichia coli* phosphoenolpyruvate carboxykinase from the use of AlF_3. *J. Mol. Biol.* **314**: 83-92.

Tan, E., Besant, P. G. and Attwood, P. V. (2002) Mammalian histidine kinases: do they REALLY exist? *Biochemistry* **41**: 3843-3851.

Taylor, S. S. and Kornev, A. P. (2011) Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem. Sci.* **36**: 65-77.

Thompson, P. R. and Cole, P. A. (2001) Probing the mechanism of enzymatic phosphoryl transfer with a chemical trick. *Proc. Natl. Acad. Sci. U. S. A.* **98**: 8170-8171.

Tian, G. C., Yan, H. G., Jiang, R. T., Kishi, F., Nakazawa, A. and Tsai, M. D. (1990) Mechanism of adenylate kinase. Are the essential lysines essential? *Biochemistry* **29**: 4296-4304.

Titball, R. W. and Rubidge, T. (1990) The role of histidine residues in the alpha toxin of Clostridium perfringens. *FEMS Microbio. Lett.* **56**: 261-265.

Tock, M. R., Frary, E., Sayers, J. R. and Grasby, J. A. (2003) Dynamic evidence for metal ion catalysis in the reaction mediated by a flap endonuclease. *EMBO J.* **22**: 995-1004.

Toney, M. D. and Kirsch, J. F. (1989) Direct Bronsted analysis of the restoration of activity to a mutant enzyme by exogenous amines. *Science* **243**: 1485-1488.

Toney, M. D. and Kirsch, J. F. (1992) Bronsted analysis of aspartate aminotransferase *via* exogenous catalysis of reactions of an inactive mutant. *Protein Sci.* **1**: 107-119.

Toxicology, B. E. S. a. (2006) Fluoride in Drinking Water: A Scientific Review of EPA's Standards. Washington, D. C., The National Academies Press.

Tremblay, L. W., Zhang, G., Dai, J., Dunaway-Mariano, D. and Allen, K. N. (2005) Chemical confirmation of a pentavalent phosphorane in complex with beta-phosphoglucomutase. *J. Am. Chem. Soc.* **127**: 5298-5299.

Tuchsen, E. and Hansen, P. E. (1991) Hydrogen bonding monitored by deuterium isotope effects on carbonyl ^{13}C chemical shift in BPTI: intra-residue hydrogen bonds in antiparallel beta-sheet. *Int. J. Biol. Macromol.* **13**: 2-8.

Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**: 1022-1025.

Van Dyke Tiers, G. (1957) A novel deuterium isotope effect in nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.* **79**: 5585-5585.

Vonrhein, C., Schlauderer, G. J. and Schulz, G. E. (1995) Movie of the structural changes during a catalytic cycle of nucleoside monophosphate kinases. *Structure* **3**: 483-490.

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

Wang, H., Falck, J. R., Hall, T. M. and Shears, S. B. (2011) Structural basis for an inositol pyrophosphate kinase surmounting phosphate crowding. *Nat. Chem. Biol.* **8**: 111-116.

Warshel, A. and Aqvist, J. (1991) Electrostatic energy and macromolecular function. *Annu. Rev. Biophys. Biophys. Chem.* **20**: 267-298.

Webster, C. E. (2004) High-energy intermediate or stable transition state analogue: theoretical perspective of the active site and mechanism of beta-phosphoglucomutase. *J. Am. Chem. Soc.* **126**: 6840-6841.

Wei, Y., Zhang, Y., Derewenda, U., Liu, X., Minor, W., Nakamoto, R. K., Somlyo, A. V., Somlyo, A. P. and Derewenda, Z. S. (1997) Crystal structure of RhoA-GDP and its functional implications. *Nat. Struct. Biol.* **4**: 699-703.

Wiesmuller, L., Noegel, A. A., Barzu, O., Gerisch, G. and Schleicher, M. (1990) cDNA-derived sequence of UMP-CMP kinase from *Dictyostelium discoideum* and expression of the enzyme in *Escherichia coli*. *J. Biol. Chem.* **265**: 6339-6345.

Wiesmuller, L., Scheffzek, K., Kliche, W., Goody, R. S., Wittinghofer, A. and Reinstein, J. (1995) Crystallization and preliminary X-ray analysis of UMP/CMP-kinase from *Dictyostelium discoideum* with the specific bisubstrate inhibitor P1-(adenosine 5')-P5-(uridine 5')-pentaphosphate (UP_5A). *FEBS Lett.* **363**: 22-24.

Williams, D. M., Jakeman, D. L., Vyle, J. S., Williamson, M. P. and Blackburn, G. M. (1998) Synthesis and binding of stable bisubstrate ligands for phosphoglycerate kinase. *Bioorg. Med. Chem. Lett.* **8**: 2603-2608.

Williams, N. H. (1998) Phosphate diesterases and triesterases. Comprehensive Biological Catalysis. Vol. 1. Reaction of Electrophilic Carbon, Phosphorus and Sulfur. Academic Press, London. pp.543-557.

Wolfenden, R. (1969) Transition state analogues for enzyme catalysis. *Nature* **223**: 704-705.

Wolfenden, R. and Snider, M. J. (2001) The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res.* **34**: 938-945.

Xiaoxia, L., Marston, J. P., Baxter, N. J., Hounslow, A. M., Yufen, Z., Blackburn, G. M., Cliff, M. J. and Waltho, J. P. (2011) Prioritization of charge over geometry in Transition State Analogues of a dual specificity protein kinase. *J. Am. Chem. Soc.* **133**: 3989-3994.

Xu, Y. W., Morera, S., Janin, J. and Cherfils, J. (1997) AlF_3 mimics the transition state of protein phosphorylation in the crystal structure of nucleoside diphosphate kinase and MgADP. *Proc. Natl. Acad. Sci. U. S. A.* **94**: 3579-3583.

Yang, J., Kennedy, E. J., Wu, J., Deal, M. S., Pennypacker, J., Ghosh, G. and Taylor, S. S.. (2009) Contribution of non-catalytic core residues to activity and regulation in protein kinase A. *J. Biol. Chem.* **284**: 6241-6248.

Yang, J., Ten Eyck, L. F., Xuong, N. H. and Taylor, S. S. (2004) Crystal structure of a cAMP-dependent protein kinase mutant at 1.26 Å: new insights into the catalytic mechanism. *J. Mol. Biol.* **336**: 473-487.

Yonemoto, W., McGlone, M. L., Grant, B. and Taylor, S. S. (1997) Autophosphorylation of the catalytic subunit of cAMP-dependent protein kinase in *Escherichia coli*. *Protein Eng.* **10**: 915-925.

Yoza, N., Ueda, N. and Nakashima, S. (1994) pH-Dependence of ^{31}P-NMR spectroscopic parameters of monofluorophosphate, phosphate, hypophosphate, phosphonate, phosphinate and their dimers and trimers *Fresenius J. Anal. Chem.* **348**: 633-638.

Zhang, B., Zhang, Y., Wang, Z. and Zheng, Y. (2000) The role of Mg^{2+} cofactor in the guanine nucleotide exchange and GTP hydrolysis reactions of Rho family GTP-binding proteins. *J. Biol. Chem.* **275**: 25299-25307.

Zhang, G., Dai, J., Wang, L., Dunaway-Mariano, D., Tremblay, L. W. and Allen, K. N. (2005) Catalytic cycling in beta-phosphoglucomutase: a kinetic and structural analysis. *Biochemistry* **44**: 9404-9416.

Zwaig, N. and Milstein, C. (1966) The phosphorylated intermediate in the phosphoglyceromutase reaction. *Biochem. J.l* **98**: 360-368.