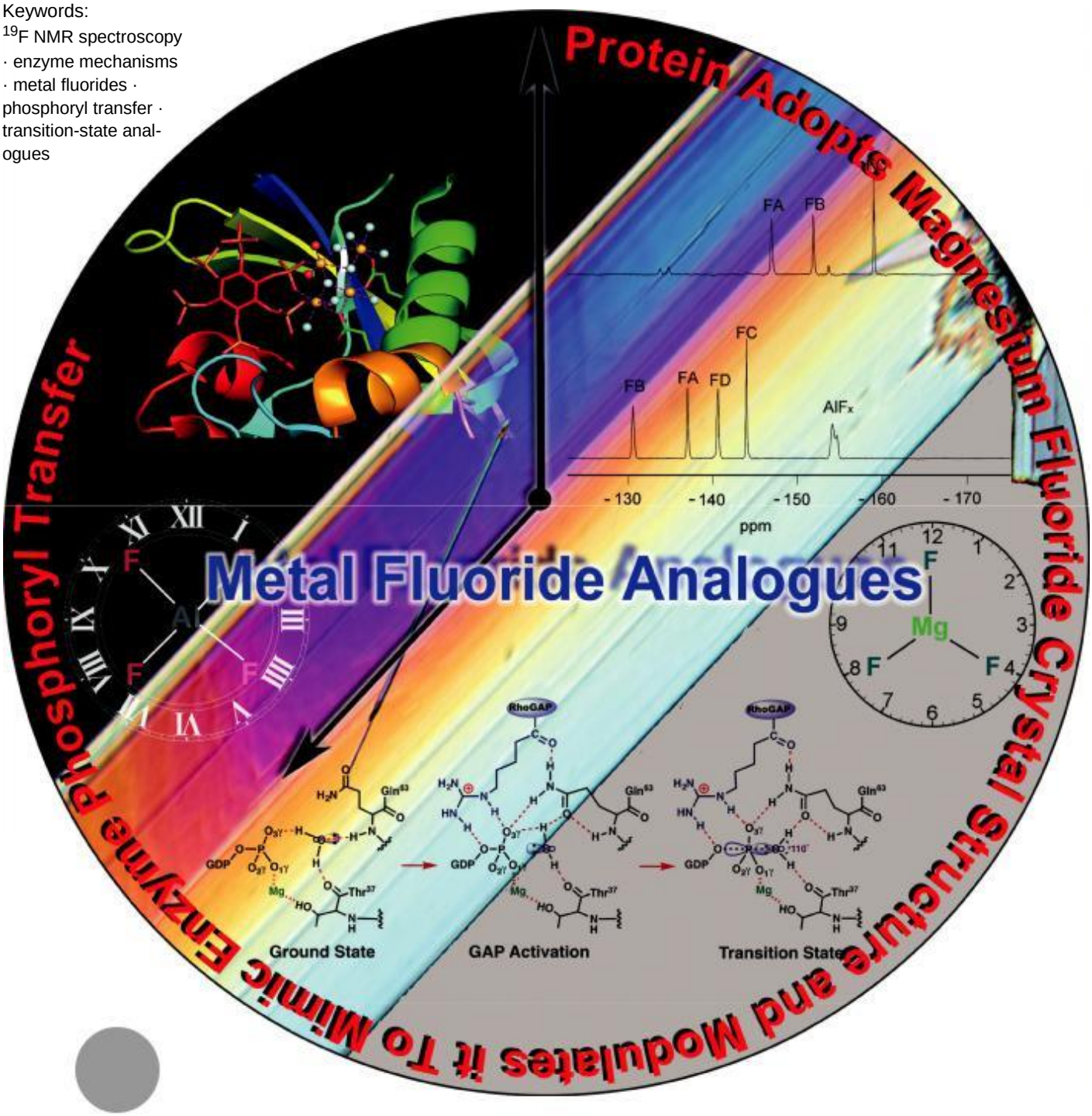
# Metal Fluorides as Analogues for Studies on Phosphoryl Transfer Enzymes

Yi Jin, Nigel G. Richards, Jonathan P. Waltho, and G. Michael Blackburn*

Keywords:
$^{19}$F NMR spectroscopy · enzyme mechanisms · metal fluorides · phosphoryl transfer · transition-state analogues



The 1994 structure of a transition-state analogue with $AlF_4^{@}$ and GDP complexed to G1a, a small G protein, heralded a new field of research into the structure and mechanism of enzymes that manipulate the transfer of phosphoryl ($PO_3^{@}$) groups. The number of enzyme struc-tures in the PDB containing metal fluorides ($MF_x$) as ligands that imitate either a phosphoryl or a phosphate group was 357 at the end of 2016. They fall into three distinct geometrical classes: 1) Tetrahedral complexes based on $BeF_3^{@}$ that mimic ground-state phosphates;
2) octahedral complexes, primarily based on $AlF_4^{@}$, which mimic "in-line" anionic transition states for phosphoryl transfer; and 3) trigonal bipyramidal complexes, represented by $MgF_3^{@}$ and putative $AlF_3^{0}$ moieties, which mimic the geometry of the transition state. The inter-pretation of these structures provides a deeper mechanistic under-standing into the behavior and manipulation of phosphate monoesters in molecular biology. This Review provides a comprehensive overview of these structures, their uses, and their computational development.

## 1. Introduction

There are now over 500 metal fluoride ($MF_x$) structures in the Protein Data Bank (PDB; Figure 1). Molecular analysis of these structures has established a simple, logical, and rational understanding of the chemical constitution of tran-sition state analogue (TSA) and ground state analogue (GSA) structures of $MF_x$ complexes. For a decade following their discovery in 1994, the atomic structures of proteins containing a metal fluoride ($MF_x$) species were based primarily on geometric considerations. From 2003 onwards, this resulted in a growing uncertainty about their chemical constitution. Recently, $^{19}F$ NMR analysis of these complexes has been used firstly to analyze and identify their atomic composition, secondly to establish their significance in solution, and thirdly to deliver experimental measurements on the electronic environment provided by the protein in conformations close to the transition state (TS). This has led to the identification of a significant number of misassignments, thus providing a corrective critique for past errors and future uncertainties.

The validity of trigonal bipyramidal (tbp) $MF_x$ structures as analogues of the phosphoryl group for analysis of "true" transition states has been endorsed by many computational studies. Several of these structures have been starting points for multiple studies on enzyme mechanisms using QM/MM and DFTanalysis. They provide a firm base for understanding enzymatic mechanisms for the catalysis of phosphate mono-esters and anhydrides, notably ATPases, GTPases, kinases, mutases, phosphohydrolases, and phosphatases.[1] Thus, phos-phoryl transfer reactions employ "in-line" geometry, are

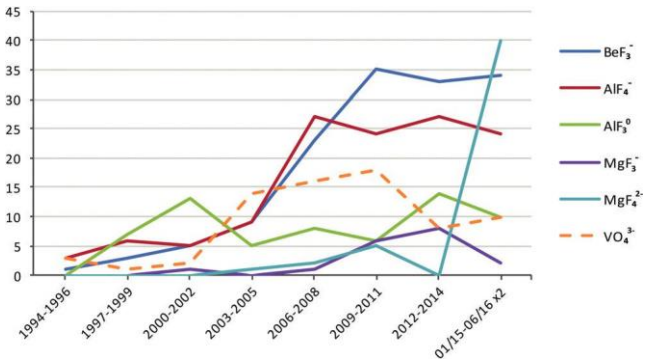
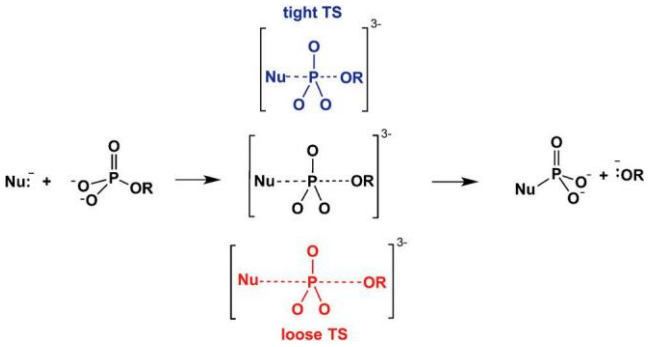



Figure 1. Number of $MF_x$ structures published in the PDB triennially. Vanadate data included for reference. The data for 01/15 to 06/16 are multiplied by a factor of 2 to represent a triennial figure.

[*] Dr. Y. Jin
Department of Chemistry, University of York York, YO10 5DD (UK)
Prof. Dr. N. G. Richards
School of Chemistry, Cardiff University Cardiff, CF10 3AT (UK)
Prof. Dr. J. P. Waltho
Manchester Institute of Biotechnology Manchester, M1 7DN (UK)
Prof. Dr. G. M. Blackburn
Department of Molecular Biology and Biotechnology University of Sheffield Sheffield, S10 2TN (UK)
E-mail: g.m.blackburn@sheffield.ac.uk

concerted, and utilize tight control of hydrogen bonds in the active-site complex to disfavor the formation of hydrogen bonds that would inhibit the chemical step in catalysis. In some cases, this hydrogen bonding includes interactions with residues, historically ascribed to provide classical general acid/base catalysis, which orientate the nucleophile for correct orbital overlap with the phosphorus center. Perhaps contro-versially, the analysis of $MF_x$ structures also suggests that any simple extrapolation of physical organic model studies to understand enzyme-catalyzed phosphoryl transfer is not possible.

Phosphoric acid ($H_3PO_4$), its esters, amidates, and anhydrides share a common tetrahedral geometry based on a phosphorus(V) core linked near-symmetrically to four oxygen or nitrogen atoms. Biological phosphoryl transfer (PTx) reactions call for the relocation of a phosphoryl group, $PO_3^@$, from a donor to an acceptor atom, typically N, O, or S, and more rarely C or F. There are many reviews of this reaction,[1] but there is no consensus on whether the reactions are more associative (tight TS) or more dissociative (loose TS) in character (Scheme 1). In either case, the phosphorus center will have trigonal bipyramidal (tbp) geometry during the PTx reactions, with the axial dimensions defined by the tight or loose nature of the transition state. A fully associative reaction would have a five-coordinate phosphorus center in the form of a covalent pentaoxyphosphorane, a putative, stable intermediate. The boundary between associative and dissociative geometries has been assigned an axial O-P-O value of 4.9 &, based on van der Waals considerations.[1a] As the primary database for $MF_x$ complexes is structurally



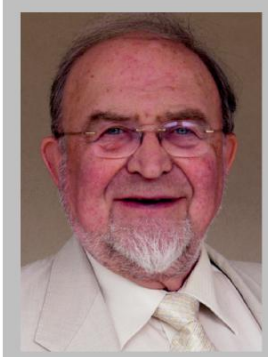

Scheme 1. Mechanism of a concerted PTx reaction. Top: bond making precedes bond breaking (blue); center: bond breaking balanced by bond making (black); bottom: bond breaking in advance of bond making (red).

driven, we review the separate groups of $MF_x$ protein complexes in terms of their geometry. This has the additional advantage of overriding ambiguities in the assignment of atomic composition, as shown in Sections 4.2 and 4.3.

## 2. Tetrahedral Phosphate Mimics, $BeF_3^@$

Beryllium(II) forms stable fluorides in water that exist as a mixture of tetrahedral species including $BeF_2{\cdot}2H_2O$, $BeF_3^@{\cdot}H_2O$, and $BeF_4^{2@}$.[2] Early NMR studies on fluoroberyllate complexes with ADP led to the analysis of mixed fluoroberyllate·ADP species with myosin, and the first X-ray


Yi Jin studied chemistry at Xiamen University, China, under Prof. Yufen Zhao. She was awarded her doctorate in 2012 under the guidance of Prof. J. P. Waltho and Prof. G. M. Blackburn at the University of Sheffield, UK. She is currently a postdoctoral researcher in the group of Prof. G. J. Davies at YSBL, University of York, UK. Her research interests involve mechanistic studies of disease-relevant phosphoryl transfer enzymes and carbohydrate-processing enzymes by using chemical, NMR spectro-scopic, molecular biology, and crystallo-graphic approaches.



Michael Blackburn is Emeritus Professor of Biomolecular Chemistry at Sheffield University and a founding member of the Krebs Institute. His undergraduate and postdoctoral career in Cambridge University led him in 1961 into the biological chemistry of phosphorus under Alexander Todd. He has worked for over 50 years at the interface of chemistry and molecular biology on nucleotides, their analogues, and the enzymes that use them. Most recently, he has focused on the unique paradox between their structural stability and kinetic lability.

Nigel Richards is Professor of Biological Chemistry at Cardiff University and a Research Fellow at the Foundation for Applied Molecular Evolution in Gainesville, Florida. After undergraduate studies at Imperial College, London, a PhD at Cambridge University, and postdoctoral research at Columbia University, where he was an author of the MacroModel software package, he held academic positions in the UK and USA. His research interests are the design of enzyme inhibitors to investigate cellular metabolism in sarcomas and the catalytic mechanisms of Fe- and Mn-dependent enzymes.

Jonathan Waltho is Gibson Professor of Biophysics at the University of Sheffield, and Professor of Structural Biology at the University of Manchester. He received his undergraduate degree from Durham University and his PhD from Cambridge University. Following postdoctoral studies at SmithKline French and The Scripps Research Institute, he has tackled a wide range of problems in protein folding, misfolding, and protein–ligand interactions. His current research interests focus on structure, electronics, and dynamics in phosphoryl, methyl, and hydride transfer enzymes.

analysis of a fluoroberyllate protein structure was reported in 1995 for an ADP·$BeF_3^{@}$ complex with myosin (PDB: 1 mmd).[3] Since then, 122 trifluoroberyllate complexes have been described, with 3 solved by NMR spectroscopy and 119 X-ray structures with resolutions of 1.2 & or lower. The vast majority of these structures have a tetrahedral trifluorober-yllate bonded to an anionic oxygen atom. The structures can be divided into two principal groups: over 70 are coordinated to an aspartate carboxylate group (including the 3 NMR-derived structures) and around 50 are coordinated to a terminal phosphate group of a nucleotide. Only 2 are coordinated to the nitrogen atom of a histidine ring.

## 2.1. Aspartyl Trifluoroberyllates

These structures share a common core, with bidentate coordination to an essential metal ion, generally $Mg^{2+}$ or rarely $Mn^{2+}$. The coordination occurs through fluorine F1 and the second carboxylate oxygen atom, OD2, to give a near-planar six-membered ring (Figure 2). [Here, and throughout, the naming of atoms in phosphates and their analogues

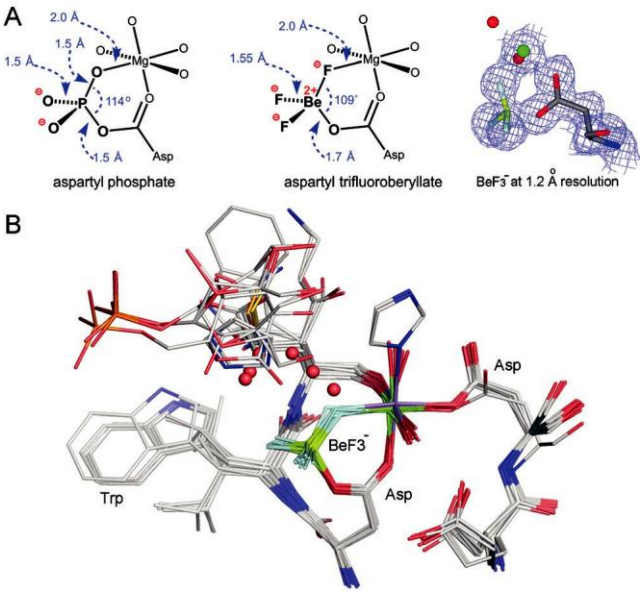



Figure 2. A) Typical aspartyl trifluoroberyllate structure with a catalytically active magnesium atom (center). Left: Aspartyl phosphate complex with a catalytically active magnesium center from phosphoserine phosphatase (PDB: 1j97) for comparison of the geometry. Right: Electron-density map for the 1.2 & resolution structure of b-phosphoglucomutase (PDB: 2wf8). B) 17 Aligned aspartyl trifluoroberyllate structures with $BeF_3^{@}$ locked in a six-membered ring (center). A catalytically active $Mg^{2+}$ ion (rarely $Mn^{2+}$) and an aspartate (usually Asp) fuse a 13-atom ring to the fluoroberyllate ring with atoms from the adjacent two amino acids downstream (rear center). The octahe-dral coordination to the Mg center is completed by an additional aspartate residue (right) and 1 or 2 water molecules, but only in two structures by histidine (top right). Atom colors: fluorine, light blue; beryllium, yellow-green; nitrogen, blue; oxygen, red. In 7 structures, an isolated water molecule (red spheres) is found close to one fluorine atom. (Electron densities presented in CCP4MG from mtz data in EDS and contoured at 1s.)

conforms to IUPAC recommendations of 2016.][4] Beryllium is difficult to locate by X-ray diffraction because it has a low electron density. This results in uncertainty in its location, and hence considerable variation in the attributed geometry (Figure 2 A). Linus Pauling assigned predominantly ionic character to the Be@F bond (80%), thereby leading to the expectation that the trifluoroberyllate function would be solvated by water.[5] However, only 10 of the 30 best-resolved structures show such an isolated water proximate to the $BeF_3^{@}$ moiety, which is not “in line” with the O@Be bond (155.3 : 9.28), and is at widely variable distances from the beryllium atom (3.8 : 0.5 &; Figure 2 B, as well as Table S1 in the Supporting Information).

## 2.2. ADP·$BeF_3^{@}$ Structures

There are 42 X-ray structures of $BeF_3^{@}$ complexes with ADP and 6 with GDP, which constitute isosteric mimics of ATP and GTP respectively. They are distributed among kinases, hydrolases, mutases, helicases, and small G proteins. Of the ADP·$BeF_3^{@}$ structures, 25 are resolved at 2.5 & and 20 align remarkably well (Figure 3). The beryllium atom is bonded to O3B, and a catalytic $Mg^{2+}$ ion is coordinated to F1 and to O1B in a six-membered ring. There is remarkable consistency in neighboring amino acids: an arginine and a lysine coordinate b- and g-phosphates and balance the anionic charge of the nucleotide. By contrast, the adenine base occupies a range of conformations (Figure 3, as well as Table S2 in the Supporting Information). A very significant feature is that 12 of the 20 structures have a water molecule hydrogen bonded to one of the three fluorine atoms. These water molecules lie well within the $BeF_3^{@}$ “cone” with their oxygen atoms being about 3.4 & from the beryllium center, with a median “in-line” angle of 158**8**, and forming a hydrogen bond to one of the fluorine atoms (2.8 : 0.3 &). As the axial O-Be-O distance is close to 5.1 &, these water molecules are part of a Near Attack Conformation (NAC) that is inter-mediate between a ground state (GS) and a TS.[6] The 6 GDP structures are very similar to structures of ADP complexes, but at rather lower resolution (Table S3).

The $BeF_3^{@}$ complex for human phosphoglycerate kinase (hPGK) raises the question: “Where is the beryllium located in the case of two oxyanion acceptors?” The structure of the complex hPGK·ADP·$BeF_3^{@}$·3PG (PDB: 4axx, 1.74 & reso-lution) places the Be atom 1.73 & from the carboxylate oxygen atom and 2.85 & from the ADP oxygen atom O3B. However, the three fluorine atoms are on average 2.75 & from the carboxylate oxygen atom and 2.96 & from the ADP oxygen atom (Figure 4 A). Since the sum of van der Waals radii for Be@O is 3.26 &, these data suggest mixed occupancy, with beryllium closer on average to the carboxylate group.[7]

## 2.3. Histidine Trifluoroberyllates

Various approaches to analogues of t-phosphohistidine have been explored. Studies on nicotinamide phosphoribo-syltransferase (NAMPT) has structurally mimicked the

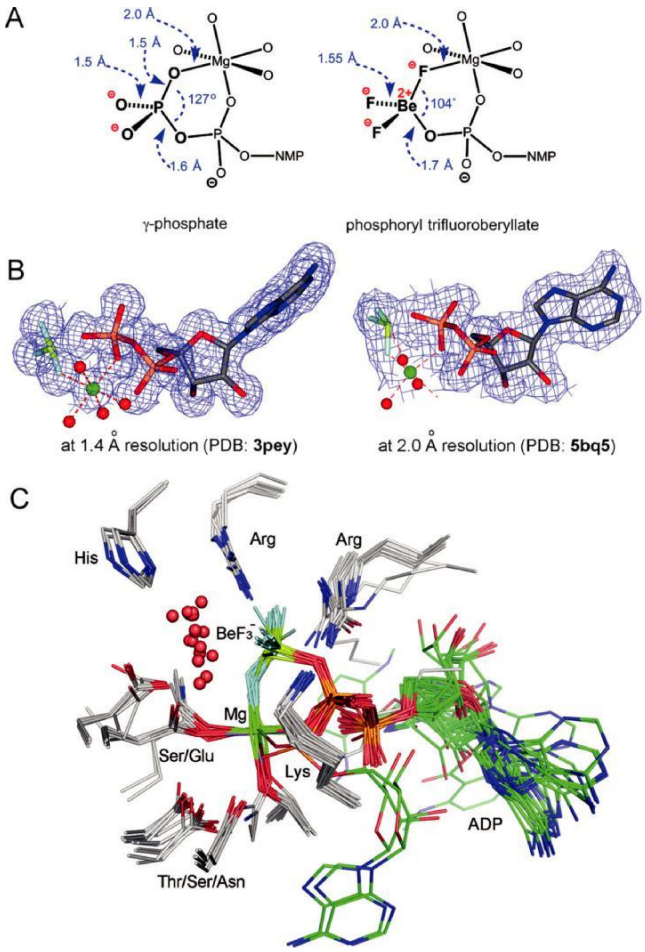



Figure 3. A) Typical nucleoside diphosphate trifluoroberyllate structure (right) with a coordinatively bound catalytically active magnesium ion for comparison of the geometry with the nucleoside triphosphate (left). B) Electron-density map for the 1.4 & resolution structure for ATP-dependent RNA helicase DNP5 (PDB: 3pey, left) compared with the 2.0 & resolution structure for a regulatory AAA + ATPase domain (PDB: 5bq5, right). C) In 20 aligned ADP·trifluoroberyllate structures, $BeF_3^{@}$ is locked in a six-membered ring (center) with the $Mg^{2+}$ ion coordinating F1 and O3B. The octahedral coordination to the Mg ion is completed by OB1, 2 trans-arranged water molecules (not shown), a Ser/Glu side-chain oxygen atom, and a Ser/Thr/Asn side-chain oxygen atom. g-Phosphate coordination to an Arg and a Lys is also common. The location of adenine residues is very variable (in green). In 12 structures, an isolated water molecule (red spheres) is located close to the $BeF_3^{@}$ "cone". Atom colors: fluorine, light blue; beryllium, yellow-green; nitrogen, blue; oxygen, red; protein residues gray. Note: It is possible that two of these structures (PDB: 1w0j and 4znl) may really be trifluoromagnesate because a) their tbp geometry is "in line" with a short O—M—O distance, and b) their crystallization solutions contained 100 mm citrate or EDTA buffer, each of which has a high affinity for beryllium.

phosphorylation of an active-site histidine with trifluoro-beryllate. Crystal structures of reactant and product com-plexes of NAMPT (PDB: 3dhf, Figure 4 B) show a covalent His247·$BeF_3^{@}$, although, in contrast to all other trifluoro-beryllate structures, the magnesium center is coordinated to one fluorine atom without any direct linkage to His247.[8]

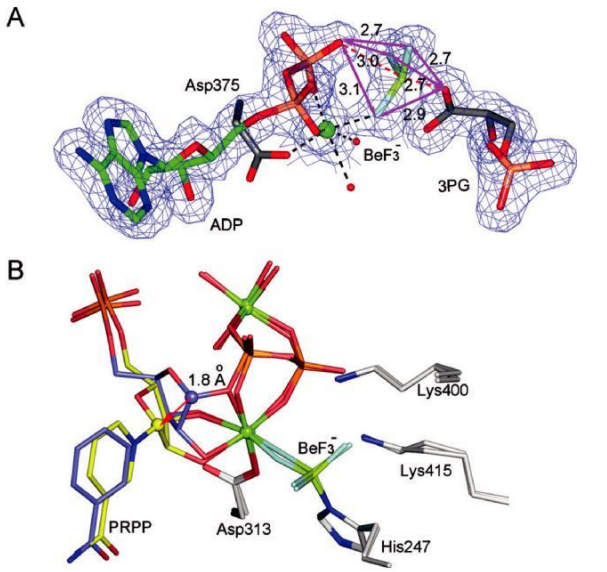



Figure 4. A) Structure of the $BeF_3^{@}$ complex for hPGK (PDB: 4axx). The beryllium center (lime green) is "in line" between O3B of ADP and 3PG. The nonbonding fluorine–oxygen distances (magenta arrows) are shorter to the carboxylate group than to the ADP oxygen atom. B) Nicotinamide phosphoribosyl transferase (PDB: 3dhf) catalyzes displacement of pyrophosphate from C1 of ribose 5-phosphate (reac-tants in purple, products in silver, the red arrow shows departure of the phosphoryl oxygen atom). The structures of two overlaid com-plexes show $BeF_3^{@}$ bonded to Ne of His247 and one fluorine atom coordinating to the octahedral $Mg^{2+}$ ion (green sphere). C1' of PRPP in the reactant (purple sphere) moves 1.8 & to bond to the nicotin-amide N1 (yellow sphere; reactant: purple sticks, product: yellow sticks, Be: lime green).

## 2.4. A Nucleotide–Beryllium Difluoride Structure

A solitary example of beryllium difluoride bridging ADP and UDP illuminates the activity of UMP/CMP kinase (PDB: 4ukd).[9] The 2.0 & structure (Figure 5 A) shows a tetrahedral beryllium center bridging O3B of ADP and O1B of UDP. An essential $Mg^{2+}$ ion coordinates one fluorine atom as well as O1B of ADP. The two diastereotopic fluorine atoms show well-separated resonances in the $^{19}F$ NMR spectrum (Fig-ure 5 B). This stable mimic of $Ap_5U$ is strongly coordinated to four arginine moieties and one lysine, and thus endorses the observation that nucleotide kinases are more strongly inhib-ited by $Ap_5Nuc$ than by $Ap_4Nuc$ on account of their additional negative charge.[10]

## 2.5. Conclusions

The significant ability of beryllium(II) fluorides to com-plete tetrahedral coordination by binding to an anionic oxygen atom makes them effective isosteric and electrostatic GS analogues of phosphate in a wide range of situations.[11] The Be@F and Be@O bond lengths are close to those of P@O (1.6 : 0.5 &), and the dominant ionic character of the Be@F bond means that the fluorine atoms readily form hydrogen bonds with a range of donors and/or coordinate to Group 2

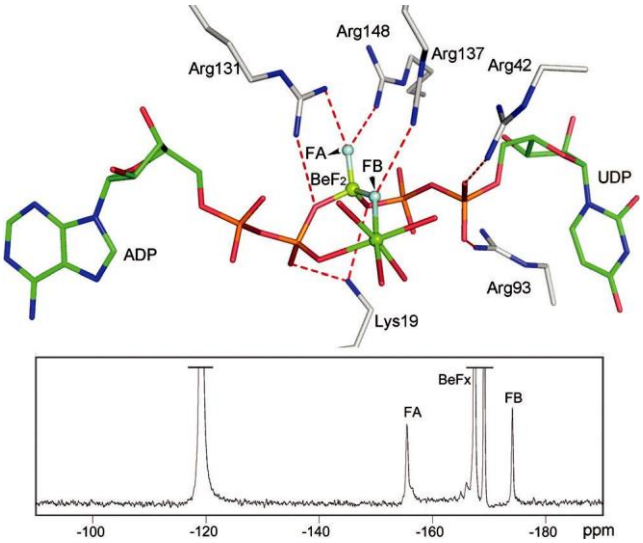



Figure 5. A) Structure of $BeF_2$ complexed to two nucleotides in UMP/ CMP kinase (PDB: 4ukd). Beryllium (green sphere) is bonded to oxygen atoms of ADP (green) and UDP (green), with one fluorine atom (light blue) coordinating to an octahedral $Mg^{2+}$ ion (lime sphere). The tetrahedral complex is coordinated by five hydrogen bonds to four amino acids (gray sticks). B) $^{19}F$ NMR spectrum for the ADP·$BeF_2$·UDP complex.

metal ions.[5] These mimics have been used to study changes in the major conformations of proteins by crystallography, NMR spectroscopy, and electron microscopy (EM), whereas studies on ADP·$BeF_3^{@}$ have supported investigations of ATPases, which drive various mechanical processes at a molecular level, particularly for myosin.[12] They have proved especially valuable for the identification of NACs in enzyme mecha-nisms, especially for b-phosphoglucomutase (bPGM).[13]

## 3. Octahedral $MF_x$ Complexes

Aluminum(III) forms stable fluorides in water that exist as a mixture of octahedral species including $AlF_2^{+}·4H_2O$, $AlF_3·3H_2O$, $AlF_4^{@}·2H_2O$, and $AlF_5^{2@}·H_2O$ depending on the fluoride concentration.[14] Their stability is a function of the pH value because aluminum forms insoluble $Al(OH)_3$ above pH 7.5.[14] Aluminum and fluoride were discovered to stim-ulate the activity of small G proteins in the presence of GDP,[15] and the proposal that they could mimic the active GTP-bound state[16] was endorsed by $^{19}F$ NMR analysis, which identified the formation of a GDP·$AlF_x$ complex for $G_1a$.[17] In 1994, crystal structures for tetrahedral GDP·$AlF_4^{@}$ com-plexes of transducin a and a heterotrimeric G protein sub-unit, $G_ia_1$, appeared almost simultaneously, and were soon followed by an ADP·$AlF_4^{@}$ structure of a myosin frag-ment.[3a,18] Since then, the number of such $AlF_4^{@}$ complex structures in the PDB (PDB ligand: ALF) determined by crystallography has grown steadily to reach 109 by March 2016 (Figure 1, see also Table S4 in the Supporting Informa-tion).

### 3.1.1. Aspartyl Tetrafluoroaluminates

The PDB has 14 structures with a tetrafluoroaluminate bonded to an aspartyl oxygen atom. This mimics an aspartyl phosphate, known to be a transient species in the catalytic activity of these enzymes. They contain a $Mg^{2+}$ ion enclosed in a six-membered ring, as seen for the corresponding $BeF_3^{@}$ structure (Section 2.1), and all align very well on PDB: 2wf7 (Figure 6, see also Table S6 in the Supporting Information),

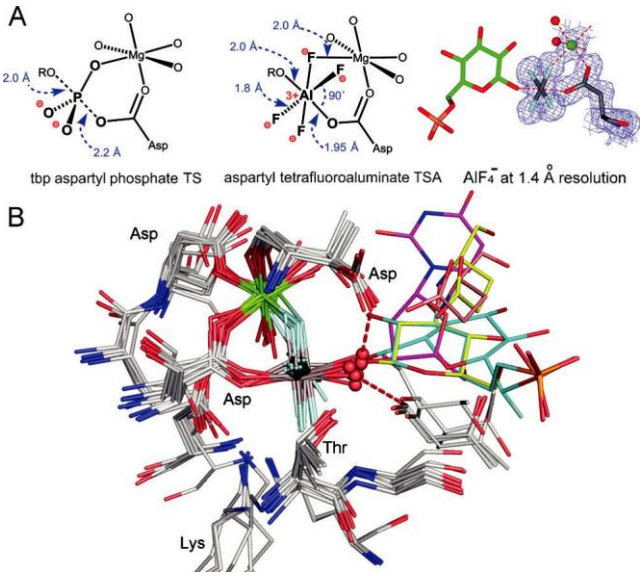



Figure 6. A) Typical aspartyl tetrafluoroaluminate structure with a coor-dinated catalytically active magnesium ion (center). Left: Aspartyl phosphate complex with catalytic magnesium from phosphoserine phosphatase (PDB: 1j97) for comparison of the geometry. Right: Elec-tron-density map for the 1.2 & resolution structure of b-phosphoglu-co-mutase (PDB: 2wf8). B) Structures of 14 aspartyl tetrafluoroaluminates superposed by Ca alignment. The aluminum ion has octahedral coordination to Asp-O4 (gray), with formation of a six-membered ring with a catalytic magnesium ion and “in line” with the acceptor oxygen atom, water molecule (red sphere), or the hydroxy group of a nucleo-side or hexose reactant (colored). Atom colors: fluorine, light blue; aluminum, gray; nitrogen, blue; oxygen, red; magnesium, green.

thus showing commonality of the additional four ligands coordinating to the catalytic $Mg^{2+}$ center. These structures fall into two subsets: six members of the first group have a second aspartate residue next-but-one to the first, and it coordinates the oxygen atom of the sixth aluminum ligand. The O@Al@O bonds are “in line” (167.58 : 7.08) with the aluminum center midway between the two oxygen atoms (separation 3.9 : 0.1 &). The Al@F bonds are 1.78 : 0.02 & (for the 6 best-resolved structures), independent of coordination to the Mg center. Three of the six structures are for bPGM, with the other three being a human mitochondrial deoxyribonucleo-tidase, a phosphoserine phosphatase (PSP), and a C-terminal domain phosphatase that operates on RNA polymerase II. In all of these, a catalytic aspartate accepts a short hydrogen bond from the apical water/hydroxy group (2.59 : 0.05 &) to complete the orientation of this oxygen atom for nucleophilic attack on the aspartyl phosphate group.[19]

The second subset comprises ATPases involved in pumping Ca, Cu, and Zn. They use an aspartyl phosphate intermediate, whose TS for hydrolysis is mimicked by the octahedral $AlF_4^{@}$. These complexes have "in-line" O@Al@O bonds (163.88 : 8.18) with aluminum midway between the two oxygen atoms (O-O separation 3.92 : 0.14 &) and Al@F bonds of 1.78 : 0.02 &. An axial water oxygen atom forms short hydrogen bonds to an invariant glutamate (2.5 : 0.1 &) and to a threonine carbonyl group (2.57 : 0.05 &). These residues clearly orientate and polarize the water for "in-line" attack on the aspartyl phosphate (see Section 8.3).[20]

### 3.1.2. Nucleotide Tetrafluoroaluminates, GDP

There are 46 X-ray structures of $AlF_4^{@}$ complexes with GDP that constitute isoelectronic but non-isosteric mimics of GTP in small G proteins, dynamins, ribosomal factors, kinases, ATPases, mutases, ion pumps, and helicases. Of these structures, 25 are resolved at 2.7 & and align remarkably well (Figure 7, see also Table S5 in the Supporting Information). The aluminum center is bonded to GDP through O3B, and the catalytic $Mg^{2+}$ ion is coordinated to F1 and O1B in a six-membered ring. There is remarkable consistency in neighboring amino acids, notably by a hepta-peptide near the N-terminus with the sequence XXXXGKS-(T), whose serine hydroxy group coordinates magnesium trans to a fluorine atom. The guanosine base and ribose occupy a common conformation (Figure 7), with the excep-tion of Atlastin (PDB: 4ido). The geometry of the $AlF_4$ moiety is well-defined, being regularly octahedral at 2.7 & resolution, with an average "in-line" O-Al-O angle of 172.8**8** : 7.18 and with aluminum midway between the axial oxygen atoms that are 4.07 : 0.23 & apart (Table 1). The Al@ F bonds are 1.77 : 0.28 & in length. All the structures have an axial oxygen ligand (Figure 7, red spheres) to aluminum that is trigonal planar with respect to two hydrogen-bond accept-ors (y-dihedral angle 4.98 : 2.98), whose angle to the axial oxygen atom averages at 102 : 68 (Figure 7 and Figure 8).

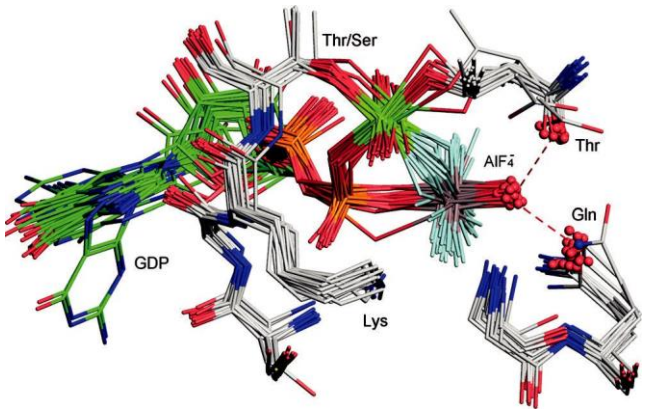



Figure 7. GDP tetrafluoroaluminate structures. 25 Structures are superposed on Ca (PDB: 2gj8). $AlF_4^{@}$ is locked in a six-membered ring (center) with catalytic $Mg^{2+}$ coordinating F1 and O3B. The octahedral coordination to $Mg^{2+}$ is provided by OB1, 2 trans-arranged water molecules, a Thr hydroxy group (top right), and a Ser/Thr hydroxy group (top center). The coordination of a phosphate oxygen atom to a Lys (center) is standard. The location of the guanine residues is regular (left, green) with two exceptions. Atom colors: fluorine, light blue; aluminum, gray; nitrogen, blue; oxygen, red; magnesium, green.

One is the backbone carbonyl group of a threonine, whose OG atom coordinates the magnesium center (Figure 7, upper right). The second is a glutamine side-chain carbonyl group or a water molecule (Figure 7, lower right, red spheres).

### 3.1.3. Nucleotide Tetrafluoroaluminates, ADP

The 45 octahedral structures that have $AlF_4^{@}$ bonded to a terminal oxygen atom of ADP (O3B) include kinases, hydrolases, isomerases, myosins, helicases, transporter pumps, and nitrogenases. They mimic ATP and are relatively diverse in conformation. The 24 that are resolved at 2.5 & have an axial O-Al-O distance of 4.05 : 0.03 & with an "in-line" angle of 1708 : 88. The majority of the 45 have a water molecule as

Table 1: Triple structure overlays for ten proteins in the PDB.[a]

| Protein | $PO_3^{@}$ donor | $PO_3^{@}$ acceptor | PDB1 reactant complex | PDB2 TSA complex | PDB3 product complex | $P_r \cdot\cdot P_p$ | $O_r \cdot\cdot O_p$ OG1[e] | $O_r \cdot\cdot O_p$ OG2[e] | $O_r \cdot\cdot O_p$ OG3[e] | $O_r \cdot\cdot O_p$ global | Od··Oa reactant | Od··Oa TSA | Od··Oa product | Od··Oa global | O-P-O TSA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ecoAcid Pase | AspP | water | 2heg | 2hf7 | 1rmy | 1.43 | 0.48 | 0.54 | 0.45 | 0.49 | 5.0 | 4.21 | 4.50 | 4.57 | 170.23 |
| AK | ATP | AMP | 1ank | 3sr0 | 4cf7 | 1.24 | 0.59 | 1.00 | 0.66 | 0.75 | *4.53* | 4.17 | 4.71 | 4.47 | 173.20 |
| cAPK | ATP | SerOH | 1rdq | 1l3r[b] | 1rdq | 1.06 | @0.50 | 0.51 | 0.26 | 0.09 | *4.52* | 4.28 | 4.33 | 4.30 | 162.18 |
| hPGK | ATP | 3PG | 4axx | 2wzb | 2x15 | 1.21 | 0.23 | 0.58 | 0.59 | 0.15 | *4.55* | 4.27 | 4.54 | 4.58 | 170.91 |
| bPGM | AspP | G1P | tbp | 2wf5 | 2wf8 | 1.30 | 0.55 | 0.58 | 0.22 | 0.45 | n/a | 4.20 | 4.41 | 4.30 | 176.45 |
| hPPIP5K2 | ATP | InsP7 | 3t9c | 3t9e | 3t9f | 1.36 | 0.40 | 0.50 | 0.58 | 0.49 | *4.66* | 4.20 | 4.66 | 4.84 | 167.13 |
| PSP | AspP | SerOH | 1l7p | 1l7n[b] | 1j97 | 0.98 | 0.18 | @0.48 | 0.28 | 0.00 | *5.07* | 4.24 | 5.45 | 4.79 | 173.93 |
| Rab11a | GTP | water | 1oiw | 1grn | 1oix | 1.10 | 0.43 | @0.48 | *0.76*[c] | 0.24 | n/a [c] | 4.39 | 4.68 | 4.55 | 157.49 |
| Ras | GTP | water | 1ctq | 1wq1 | 1xd2 | 1.39 | 0.65 | 0.81 | 1.15 | 0.73 | *6.22* | 4.45 | 4.67 | 4.61 | 165.13 |
| RhoA·GAP | GTP | water | 1a2b | 1ow3 | 5xxx[d] | 0.93 | @0.66 | 0.38 | 0.53 | 0.08 | *5.24* | 4.19 | 4.44 | 4.62 | 172.38 |
| mean : SD | | | | | | 1.20 : 0.18 | 0.24 : 0.46 | 0.39 : 0.49 | 0.48 : 0.19 | 0.37 : 0.41 | 4.80 : *0.30* | 4.26 : 0.09 | 4.55 : 0.14 | 4.65 : 0.51 | 170.2 : 4.6 |

[a] Distances given in &, angles in 8. [b] Re-refined (by Dr. Matt Bowler) as $MgF_3^{@}$ on the basis of $^{19}F$ NMR analysis. [c] Data in italics are 2 SD from the mean, thus omitted from analysis. [d] In preparation. [e] Clockwise order for the three O-O distances (with $Mg_{cat}$ behind) and O1G coordinated to magnesium.[4] n/a =not applicable.

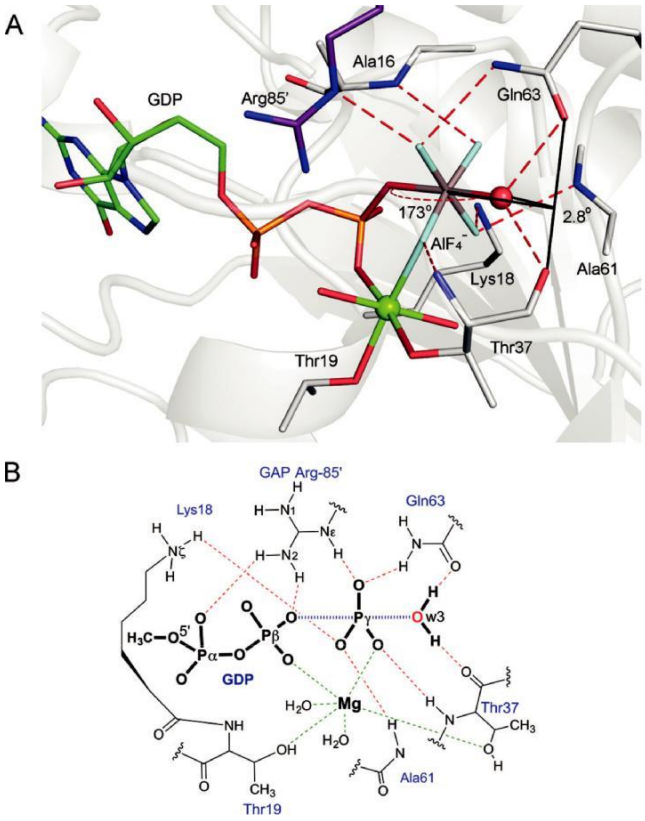



Figure 8. A) RhoA/RhoGAP·GDP·$AlF_4^{@}$ complex (PDB: 1tx4) showing hydrogen bonds from a nucleophilic water molecule to carbonyl oxygen atoms of Gln63 and Thr37 with a y-dihedral angle of 2.88 and in-line angle of 173.08. Atom colors: carbon, silver; aluminum, gray; nitrogen, blue; oxygen, red; fluorine, light blue; magnesium, green. B) Hydrogen-bond network for the RhoA/GAP·GTP·wat TS complex.

the second oxygen ligand, with the catalytic $Mg^{2+}$ ion also coordinated to one b-oxygen and a fluorine atom. This is illustrated for F1ATPase (PDB: 1h8e; Figure 9 A). Three complexes have the magnesium center triply coordinated to OA, OB, and F.

Overall, the aluminum atom is closer to O3B (1.95 : 0.09 &) than to the second oxygen atom (2.08 : 0.12 &), and the Al@F bond lengths (for the 12 best-resolved structures) are 1.77 : 0.04 & (Figure 10). The variable general position of the fluorine atoms relative to the catalytic $Mg^{2+}$ ion suggests

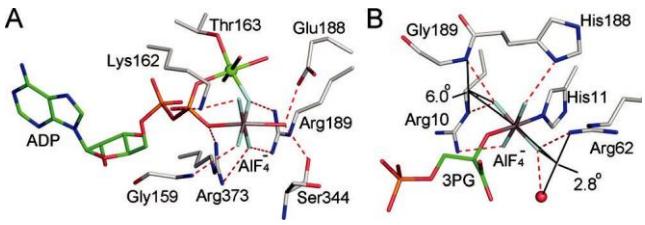



Figure 9. A) F1ATPase TSA complex (PDB: 1h8e) with ADP·$AlF_4^{@}$·wat. B) Phosphoglycerate mutase (PDB: 2f90) with an $AlF_4^{@}$ TSA complex mimicking a PTx reaction from His11 to OH-2 of 3PG. Aluminum coordinates four fluorine atoms, with His11 Ne and PGA OH-2 as axial ligands. Atom colors: ADP and 3PG, green; fluorine, light blue; amino acids, silver.

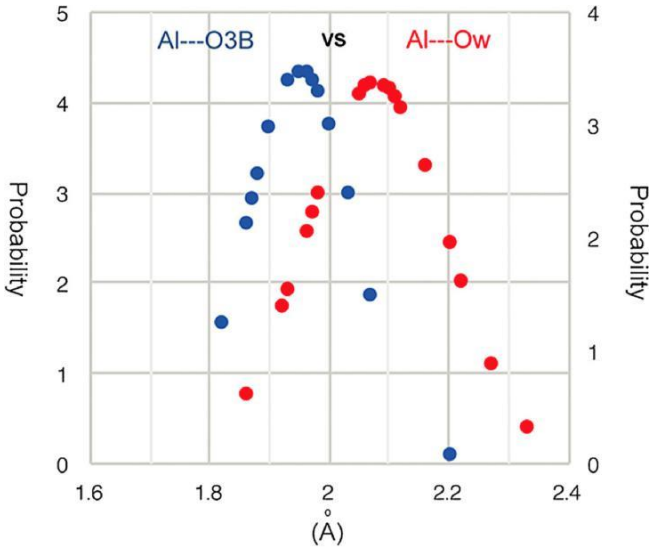



Figure 10. Normal distribution of O3B-Al (blue) and Al-Ow (red) bond lengths in 21 ADP·$AlF_4^{@}$ TSA complexes at 2.4 & resolution. Mean values and standard deviations: 1.95 :0.09 and 2.08 :0.12 &.

that some compromise has been reached in fitting four fluorine atoms into protein loci that have evolved to accommodate three electronegative oxygen atoms.

### 3.1.4. Other Tetrafluoroaluminates

Two structures have $AlF_4^{@}$ bonded to a histidine nitrogen atom, as illustrated for phosphoglycerate mutase (PDB: 2f90). This mimics the PTx reaction of His11 with OH-2 of 3PG (Figure 9 B).

## 3.2. Octahedral Trifluoroaluminates, $AlF_3^0$

There are three examples of octahedral complexes where an aluminum trifluoride core is expanded to octahedral, six-coordination by having three oxygen ligands (Table S7). For the small G protein Rab5a, the mutation A30P results in the addition of the side-chain hydroxy group of Ser29 to the aluminum center. For hPGK, the mutation K219A results in the addition of water to the aluminum center. For a bacterial dUTPase, $AlF_3^0$ takes the place of the b-phosphoryl group in dUTP, with coordination to O3A, O3B, and to the water nucleophile completing the octahedral array (Figure 11). This structure provides a unique example where nucleophilic attack is directed at a nonterminal NTP phosphorus center.[21]

# 4. Trigonal Bipyramidal $MF_x$ Complexes

## 4.1. Trifluoromagnesate, $MgF_3^{@}$

Magnesium does not form stable fluorides in water. Magnesium fluoride is moderately soluble (2 mm) with an estimated dissociation constant for $MgF_{2(aq)}$ of $10^{@5}$ m.[22] Trifluoromagnesate–protein complexes were first anticipated

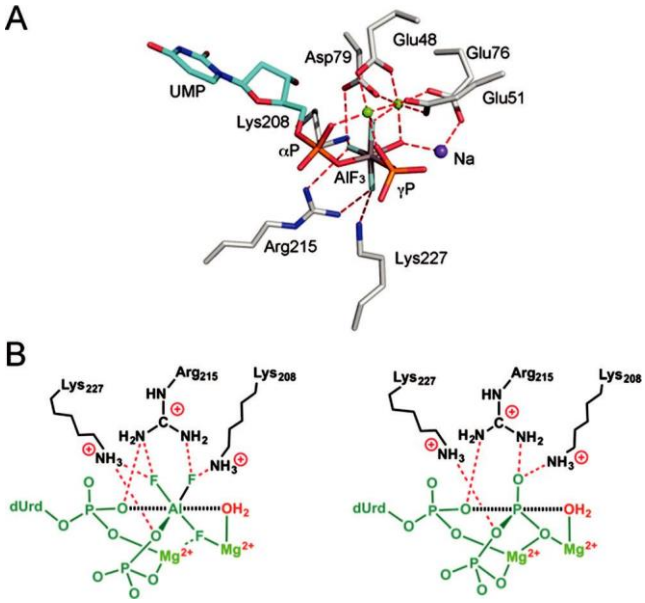



Figure 11. A) Trifluoroaluminate structure for dUTPase (PDB: 4di8). $AlF_3$ coordinates GMP (green bonds) with an in-line water molecule coordinated to sodium (purple sphere) and with $PO_4^{2@}$ adjacent to the leaving O3A. Two magnesium atoms (green spheres) are coordinated to the reactants and to four carboxylate residues. B) Trigonal bipyramidal of octahedral $AlF_3$ as a TS mimetic for a phosphoryl group (charges on phosphate moieties omitted for clarity).

on the basis of magnesium-dependent fluoride-inhibition studies, which led to the first identification of $MgF_3^{@}$ in a tbp crystalline TSA complex for the small G protein RhoA/RhoGAP (Figure 12A).[23] The PDB now lists 16 entries for this ligand (PDB ligand: MGF), while a further 3 entries assigned as tbp $AlF_3^0$ have been shown by $^{19}F$ NMR spectroscopy to be $MgF_3^{@}$ complexes (Table S8).[24] Magnesium is normally six-coordinate and gives octahedral complexes with oxygen ligands. By contrast, trifluoromagnesate is five-coordinate, and has ideal characteristics to mimic the phosphoryl group as it is isoelectronic with $PO_3^{@}$ and has the same tbp geometry. Examples of its use as phosphate mimetics include small and large molecule kinases, mutases, phosphatases, and hydrolases. Their complexes invariably involve coordination to one catalytic $Mg^{2+}$ ion (two for some protein kinases) and usually involves a cyclic six-membered ring structure, as shown for aspartyl phosphate mimics (Figure 12B). They have an axial O-Mg-O distance of 4.19 : 0.08 &, with an in-line angle of 171.4**8** : 3.98. The axial Mg@O bonds are 2.13 : 0.10 &, with Mg@F bonds of 1.83 : 0.06 &, In contrast, the computed lengths of nonbridg-ing P@O bonds are 1.52 : 0.02 &.[25]

## 4.2. Aluminum Trifluoride, $AlF_3^0$

The first example of an aluminum trifluoride complex was presented in 1997 for a tbp complex in the active site of a dinucleotide kinase (PDB: 1kdn), shortly to be followed by a study on Ras/RasGAP in complex with GDP.[26] There are now 56 structures that report an $AlF_3^0$ core. Of these, three are octahedral (see Section 3.2), and four have been shown by $^{19}F$ NMR spectroscopy to be $MgF_3^{@}$ (see Sections 4.1 and 7.2). Of the remainder, only two alkaline phosphatase structures may be identified confidently as having a tbp $AlF_3^0$ core (Figure 13). In mutant P300A (PDB: 1kh5), two catalytically active $Zn^{2+}$ ions share one fluorine atom, while Ser102 and a zinc-coordinated water molecule provide the axial ligands

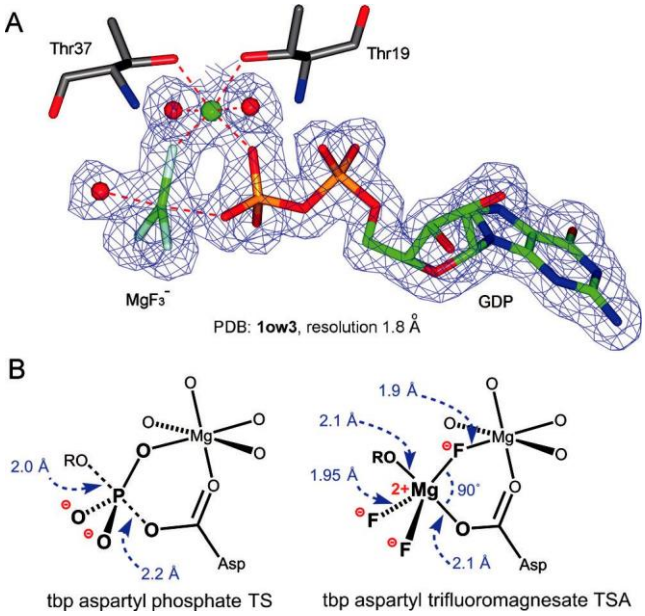



Figure 12. A) $MgF_3^{@}$ complex with GDP for RhoA (PDB: 1ow3) showing the electron density. B) Typical $MgF_3^{@}$ complexes with aspartate residues in a six-membered ring complex with the catalytic $Mg^{2+}$ center.

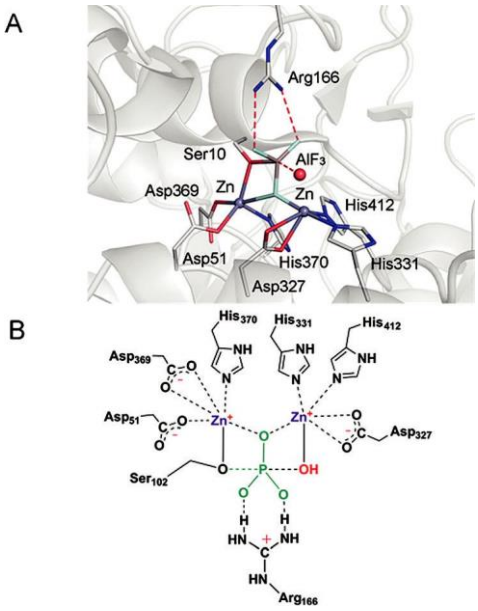



Figure 13. A) Structure of the catalytic center for alkaline phosphatase complexed to $AlF_3$ (PDB: 1kh5). B) Organization of the coordination in the active site with the transferring phosphoryl group (green) and nucleophilic water molecule (red).

for the tbp aluminum center. It has an apical O-Al-O distance of 3.80 & and Al@F bonds of 1.75 &, which are characteristic of the $AlF_4^{@}$ complexes described above (Section 3.1, Table S7). What is the situation for the remaining 48 $AlF_3^0$ complexes?

The influence of the pH value on the transition between octahedral and tbp structures of $AlF_x$ complexes in protein crystal structures of PTx enzymes was proposed to involve a switch from $AlF_4^{@}$ to $AlF_3^0$ at elevated pH values.[27] However, studies on the pH dependence of the solubility of the aluminum ion supported an alternative interpretation.[14] $Al(OH)_3$ precipitates at pH 8, thereby resulting in replace-ment of aluminum by magnesium in the protein complexes, with a consequent change to the tbp geometry. That conclusion has now been validated by pH-dependent $^{19}F$ NMR analyses for several enzymes (Section 7.2).[24b,28] In some borderline cases, for example, protein kinase A (cAPK) and PSP, there is partial dual occupancy of the active site by tbp and octahedral complexes in the crystal.[19, 24b,c] In structural terms, the dimensions of the tbp complexes closely reflect those of known trifluoromagnesates: axial O@M@O bonds of 4.29 : 0.39 &, and M@F bonds of 1.75 : 0.12 & (see Section 7.2 and Figure 17). It is, therefore, likely that $^{19}F$ NMR analysis or crystallization in an aluminum-free medium will justify reassignment of some, or many, of these complexes as trifluoromagnesates (Table S9).

Taken together with trifluoromagnesates, a common gen-eral pattern of axial ligands emerges. The $MF_3$ species requires at least one anionic oxygen atom. ADP (25) and GDP (10) phosphates provide the overwhelming majority of examples, while aspartate (11) is also significant. Water (27) is the dominant neutral axial ligand, while serine and threonine hydroxy groups appear infrequently. There is no example of both axial ligand positions occupied by two neutral ROH groups. As was observed for octahedral complexes (Sec-tion 3.1.4), there is only one example with histidine as a ligand (PDB: 1kdn). (NB: protein tyrosine phosphatases use a cys-teine–histidine ion-pair mechanism.[29])

### 4.3. Tetrafluoromagnesate, $MgF_4^{2@}$

A group of structures for the $Ca^{2+}$ ion pump ATPase contain tetrahedral moieties that have been assigned as $MgF_4^{2@}$ without further experimental validation. Magnesium is only rarely four coordinate and then usually has sterically bulky ether oxygen atoms as ligands.[30] In all the examples in the PDB, the tetrahedral $MgF_4^{2@}$ moiety is remote from ADP, is coordinated to magnesium, and has one or more of its atoms in contact with a backbone carbonyl oxygen atom (e.g. PDB: 1wpg).[31] Subsequent work has described the same tetrahedral moiety for the Na/K pump ATPase (PDB: 2zxe).[32] However, this "$MgF_4^{2@}$" is proximate to a magnesium atom that has an aspartate ligand that closely resembles the tbp structure with a six-membered ring that is common for complexes of aspartate with $MgF_3^{@}$ (Section 4.1 and Fig-ure 16C). Indeed, crystallographic refinement with $MgF_3^{@}$ in place of $MgF_4^{2@}$ produces an equally valid structure (Sec-tion 7.3). This leads to the conclusion that, unless established with further measurements, a more consistent chemical interpretation for all such "$MgF_4^{2@}$" situations is that they are trifluoromagnesates that mimic the TS for the hydrolysis of an aspartyl phosphate.

Finally, the most remarkable $MF_x$ structure is that of a human diphosphoinositol phosphatase, cocrystalized with myo-inositol hexakisphosphate and then soaked with sodium fluoride (PDB: 2q9p).[33] The resulting complex has four octahedral magnesium atoms with nine fluorine ligands. This complex embraces $MgF_2$, $MgF_3$, $MgF_4$, and $MgF_5$ species in a single complex and offers the first example of an octahedral $MgF_x$ (Figure 14). Its core appears related to the rutile structure of $MgF_2$ which has an octahedral magnesium center and trigonal planar fluorine.[34]

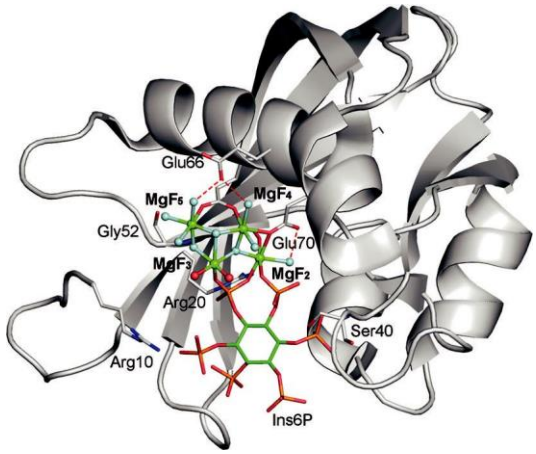



Figure 14. Structure of hPPIP5K2 (PDB: 2q9p) to show the "$Mg_4F_9$" cluster adjacent to phosphates 4 and 5 of Ins6P.

## 5. $^{19}F$ NMR Studies on $MF_x$

The inclusion of metal fluoride moieties within protein complexes has opened up the opportunity to use $^{19}F$ NMR measurements to examine the environment in which phos-phate groups reside within the protein. The $^{19}F$ isotope has 100% natural abundance and a very high gyromagnetic ratio (25.18 0 $10^7$ $T^{@1}$ $s^{@1}$), which leads to NMR spectra of very high sensitivity. Hence, metal fluoride species can be detected at low protein concentrations, and in large-molecular-weight complexes.[20,24b,c,35]

### 5.1. Chemical Shifts

The chemical shifts of $^{19}F$ resonances provide a key measure of interactions between $MF_x$ moieties and their protein hosts. They are reliable reporters of the electronic environment in the vicinity of the fluorine nuclei. When combined with calculations (Section 6.3), they can also act as indirect reporters of the changes in the electronic environ-ment experienced by phosphoryl oxygen atoms at the TS for the transfer reaction.[20,36] $^{19}F$ resonances display a high degree of dispersion and can be predicted with good precision from quantum calculations of electronic distribution.[37] The aver-

age chemical shifts of resonances from $AlF_x$, $MgF_x$, and $BeF_x$ species in aqueous solution differ (@154, @156, and @169 ppm, respectively), but a wide spread of individual shifts is observed in complexes with proteins. In cognate bPGM complexes, for example, the average chemical shifts are @138 ($AlF_4^{@}$, Figure 15C), @153 ($MgF_3^{@}$, Figure 15B),

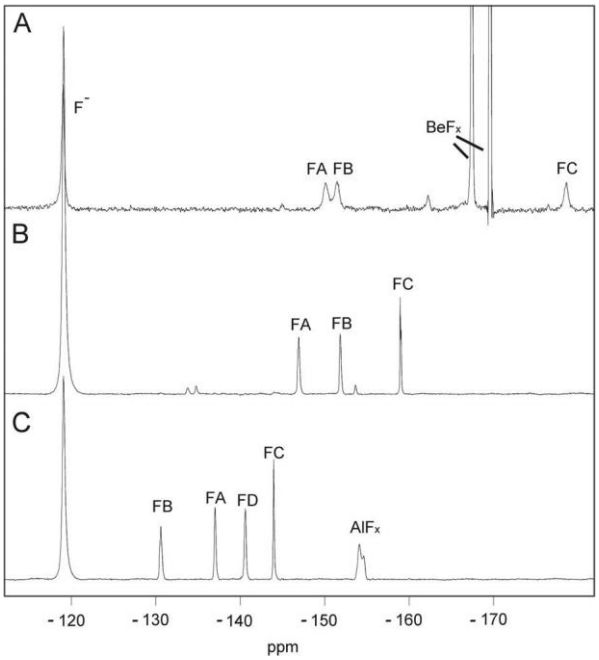



Figure 15. 1D $^{19}F$ NMR spectra of bPGM complexes with A) $BeF_3^{@}$, B) $MgF_3^{@}$ plus G6P, and C) $AlF_4^{@}$ plus G6P. The $^{19}F$ resonance at @119 ppm in each spectrum is from free $F^{@}$ ions, while those between @160 and @170 ppm (upper spectrum) are from unbound $BeF_x$ species and those between @150 and @160 ppm (lower spectrum) are from unbound $AlF_x$ species. The middle spectrum contains 3 small signals from a second $MgF_3^{@}$-bound protein conformation.

and @160 ppm ($BeF_3^{@}$, Figure 15A).[13,24c] This distribution is strongly affected by the vicinity of hydrogen-bond donors, as shown clearly by a comparison of the G6P and the 2-deoxyG6P complexes of bPGM.[28] In the bPGM·$MgF_3^{@}$·2deoxyG6P TSA complex, one fluorine atom loses its hydrogen-bond partner and its resonance moves substantially upfield (@18.1 ppm). (NB: $^{19}F$ chemical shifts are quoted relative to trifluoroacetic acid as a reference.)

The high sensitivity of $^{19}F$ chemical shifts to the surrounding environment can be used to show how enzymes control the influence of changes in the protonation state. Thus, for bPGM, it was observed that $^{19}F$ chemical shifts are invariant over the pH range 6.5–9.5, thus indicating that any changes in the protonation state of the protein has no detectable influence on the environment of the TS complex. Character-istic average chemical shift values for different $MF_x$ species have identified that millimolar concentrations of fluoride are sufficient to leach $Al^{3+}$ ions from glass, including borosilicate glass, and transform $MgF_3^{@}$ complexes into $AlF_4^{@}$ complexes unless an aluminum chelator such as deferoxamine is present.

### 5.2. Chemical Exchange

It is observed, particularly in the $AlF_4^{@}$ complexes of some enzymes (including many early NMR studies of these complexes), that individual $^{19}F$ resonances coalesce to a single resonance as a result of the rapid exchange of fluorine atoms between sites.[23b, 38] Resolved resonances of similar complexes have chemical shift differences of up to 10 kHz, which shows that the interchange of fluorine atoms greatly exceeds this rate in some $AlF_4^{@}$ complexes. All the $MgF_3^{@}$ complexes of wild-type enzymes reported to date have resolved $^{19}F$ resonances, and hence much slower rates of fluorine interchange. For $BeF_3^{@}$ complexes, the spectra show evidence of faster exchange rates than for $MgF_3^{@}$ com-plexes.[13]

### 5.3. NOE Effects

Proton distribution in the vicinity of fluorine nuclei in the $MF_x$ moiety can be assessed through the quantitation of $^{19}F$-$^{1}H$ NOE effects. This approach has been used to determine solution structures of bPGM·$MgF_3^{@}$·G6P TSA and bPGM·$AlF_4^{@}$·G6P TSA complexes, and so resolve a contro-versy concerning a reported pentaoxyphosphorane for this enzyme (Section 7.1).[24a,c] Traditionally, $^{19}F$-$^{1}H$ NOE effects are difficult to quantify because of the effects of spin diffusion between $^{1}H$ nuclei as the $^{19}F$-$^{1}H$ NOE builds but, for $MF_x$ complexes, the primary NOE effects are to exchangeable protons. Hence, $^{1}H$-$^{1}H$ spin diffusion can be suppressed by using a perdeuterated enzyme in a protonated buffer. Resonance assignment of the exchangeable $^{1}H$ nuclei in the protein allows unambiguous assignment of individual $^{19}F$ resonances.

### 5.4. Solvent-Induced Isotope Shifts (SIIS)

Proton distributions in the vicinity of fluorine nuclei can be assessed independently of $^{19}F$-$^{1}H$ NOE effects on the basis of solvent-induced hydrogen/deuterium primary isotope shifts (SIIS) of the $^{19}F$ resonances. For hydrogen bonds to $MF_x$ moieties, F····H@N and F····H@O, the magnitudes of the isotope shifts reflect local proton densities because of the through-space transmission of differences in the electric field between X@H and X@D bonds.[39] For example, in the bPGM·$MgF_3^{@}$·G6P TSA complex (Figure 15B), $F_A$ is coordinated by three protons (in a distorted tetrahedral arrangement), $F_B$ is coordinated by two protons (in a trigonal arrangement), and $F_C$ is coordinated by one proton, thus giving sum SIIS values of 1.6 ppm, 1.4 ppm, and 0.9 ppm, respectively. Comparing the G6P and the 2-deoxyG6P TSA complexes of bPGM, the sum SIIS value of one fluoride ion for the latter complex falls to close to zero (0.2 ppm), which indicates that loss of the hexose 2-OH group leaves this fluorine atom virtually devoid of hydrogen bonds.[28] The consequence of the removal of this hydroxy group on the whole TSA complex is also observable by the other two fluorine atoms moving closer to their hydrogen-bond part-

ners, as shown by small increases in their sum SIIS values (to 1.7 ppm and 1.5 ppm).

### 5.5. Scalar Couplings across Hydrogen Bonds

Details of the coordination of the $MF_x$ moiety by the protein is further shown in scalar couplings between nuclei involved in N@H···F hydrogen bonds. $^1J_{HF}$ and $^2J_{NF}$ couplings have been reported for individual $H^N$···F pairs, with values up to 59 and 36 Hz, respectively.[36b] The magnitudes of both scalar couplings correlate closely with distances measured from crystal-structure analysis. Hence, as well as reporting on the interaction across individual hydrogen bonds, scalar couplings provide an independent means of assigning $^{19}F$ resonances, and cross-validating solution and crystal behav-ior.

### 5.6. Conclusions

NMR measurements of $^{19}F$ nuclei in the active site of $MF_x$ TSA complexes provide a picture of the relationship between charge distribution in the mimic for phosphoryl group transfer and the enzyme. The good relationship between $^{19}F$ chemical shifts and SIIS values illustrates the dominant influence that very local hydrogen bonds have on shaping the charge density on $MF_x$ moieties. Moreover, the strong correlation between observed NMR parameters and the coordinates determined for numerous proteins in the crystalline state is a vital link that shows atomic positions determined at high resolution in the solid phase very closely reflect solution behavior.

## 6. Computational Analyses of $MF_X$ Complexes

There have been almost no direct computational studies of $MF_x$ complexes within protein binding sites. Instead, these GSA and TSA structures have been widely used as starting points for a very large number of calculations by replacing the $MF_x$ moiety by $PO_3$ while retaining the tbp geometry. The resulting structures have then permitted computations aimed at delineating the molecular mechanisms of a variety of enzymes that catalyze PTx reactions,[40] particularly the small GTPases, which play critical roles in cell signaling and regulation, and to cAPK.[41] Theoretical methods provide considerable insight into the distribution of electrons within molecules, and the energies of protein/ligand interactions that mediate binding and TS stabilization.[42] Calculations have also been used to obtain accurate structures that were used to resolve the nature of $MF_x$ species in X-ray crystal structures of relatively low resolution.[43] More recently, computational methods have also validated the idea that tbp $MF_x$ structures are analogues of the phosphoryl group in the "true" transition states for enzyme-catalyzed reactions, and provide useful information on the extent to which $MF_x$ moieties resemble ground states or transition states in enzyme-catalyzed PTx reactions.[20] Although the covalent character of P@O and M@ F bonds is very different in the GS and (most likely) the TS, these calculations demonstrate that differences in $^{19}F$ chemical shifts do provide insights into the environments experienced by the oxygen atoms in the "true" TS for the reaction.

### 6.1. Computational Methods.

The principal approach to elucidating the properties of $MF_x$ complexes has been the use of DFT, because of the ability of this method to yield accurate structural proper-ties.[44] Numerous reviews are available that detail the theoretical principles underlying DFT, together with its limitations, which include problems in modeling dispersion interactions and activation energy barriers in chemical reactions.[45] One important advantage of DFT, is that molecular systems composed of relatively large numbers of atoms can be treated completely quantum mechanically, thus allowing considerable insight into the electrostatic properties of $MF_x$ complexes and how these might be perturbed by being in a protein environment. The general strategy has been to build active-site models composed of the $MF_x$ complex and residues that interact directly with the complex and surround-ing molecules, such as ADP and GDP.[46] Larger models can also be built that include "second-shell" residues, which form hydrogen bonds to the initial set of inner residues.[20] In an alternative approach, which avoids the need to place artificial coordinate restraints on atoms in the QM region, the complete system is modeled by using QM/MM methods.[47] Here, the QM region is embedded in the rest of the protein and solvent, with the additional atoms (in an MM region) being described by classical potential energy functions that depend on "force-field" parameters. Various methods can then be used to "couple" the QM and MM regions.[48] The advantage of the QM/MM approach, which also permits the inclusion of electrostatic effects arising from the protein and solvent environment, lies in the elimination of "edge effects" at the boundaries of the QM region arising from coordinate restraints. In addition, the relatively simple potentials used to describe the MM region allow the use of MD simulations to obtain estimates of the free energy of the system, which is not reliably obtained by analysis of the geometry-optimized QM active-site models.[49]

### 6.2. $BeF_3^@$ Complexes

As discussed in Section 2, beryllium fluoride complexes resemble GS phosphate groups when bound to nucleophilic groups or dinucleotides. The extent to which such tetrahedral complexes mimic phosphate moieties was explored using QM calculations of $BeF_3^@$ complexed to the catalytically impor-tant aspartate side chain of bPGM in the presence and absence of G6P, a substrate for the enzyme.[13] Large models, consisting of the $BeF_3^@$ complex and 29 residues surrounding the active site, were obtained from crystal structures of these complexes and structurally optimized using B3LYP and the 6-31G basis set, with the inclusion of d-polarization functions for the fluoride ions.[13] As usual, the outer atoms in these models were constrained to their crystallographic coordi-

nates. Atomic charges were then computed using the Mul-liken formulation to minimize computational expense. The results showed that the beryllium and fluoride ions carry about 60% and 75% of the charges expected for phosphorus and oxygen atoms, respectively, in a phosphate group. Hence, although the total charge of the $BeF_3^{@}$ moiety is identical to that of the reactive intermediate in the enzyme-catalyzed reaction, the internal separation of charge is scaled down.[13]

### 6.3. $MgF_3$ Complexes

There is ample evidence that $MgF_3^{@}$ is an excellent stable analogue of the TS for phosphate transfer in a number of enzyme-catalyzed reactions (Section 4.1). Early DFT calculations were performed to investigate the claim that X-ray crystallography had revealed the structure of a phosphorane intermediate in the reaction catalyzed by bPGM, and validated the correction that the tbp complex was $MgF_3^{@}$ (see Section 7.1).[50] The calculated distances for a $MgF_3^{@}$ anion were consistent with those seen in the crystal structure. Subsequent high-level QM/MM calculations have supported this conclusion, and have shown that it also holds for PTx reactions catalyzed by UTPase.[51] QM/MM studies followed that sought to demonstrate that $MgF_3^{@}$ rather than the isoelectronic $AlF_3$ was present in medium-resolution X-ray crystal structures of the Ras/RasGAP complex.[43] The QM region was modeled using standard Hartree–Fock ab initio calculations, which ignore the effects of electronic correlation. Nonetheless, this level of QM theory was sufficient to show that calculated distances and angles for the $MgF_3^{@}$ complex were in much better agreement with the crystal structure of the Ras/RasGAP·GDP·$MF_x$ complex than those computed for either $AlF_3$ or $AlF_4^{@}$. This was an important result because the electron density observed for the $MF_x$ species in the Ras/ RasGAP·GDP·$MF_x$ structure (PDB: 1wq1) was inadequate to permit an unambiguous assignment of the ion.[26a] More recent work has sought to establish the extent to which $MgF_3^{@}$ resembles PTx in the TS for GTP hydrolysis catalyzed by the RhoA/RhoGAP complex.[20] Specifically, this study, which employed DFT calculations on a very large active-site model containing 91 heavy atoms, demonstrated that the observed $^{19}F$ chemical shifts for the RhoA·RhoGAP·GDP·$MgF_3^{@}$ complex can indeed be interpreted as indirect measures of the relative electron densities of the cognate oxygen atoms in the “true” TS for attack of water on the terminal phosphate of GTP.[20]

### 6.4. $AlF_3$ Complexes

Notwithstanding the questions raised about the validity of designating many tbp $MF_x$ complexes as $AlF_3^0$ (see Section 4.2), their structures, notably for Ras and for cAPK, have been used as starting points for many computations. The success of these computations lies in the simplicity of the transformation of $AlF_3^0$ into $PO_3^{@}$ without regard to the change in charge involved. Only the tbp geometry matters.

### 6.5. Conclusions

Taken as a whole, the number of computational studies on the electronic structure and steric properties of protein-bound $MF_x$ complexes remains small. There has also been limited evaluation of their resemblance to TS structures calculated using either QM or QM/MM methods for a range of enzymes, and their dynamic behavior within the active site remains poorly explored. This is surprising given the clear differences in the $^{19}F$ NMR spectra reported for complexes containing $BeF_3^{@}$, $MgF_3^{@}$, and $AlF_4^{@}$ (see Section 5).

$MF_x$ complexes have provided valuable starting points for numerous QM and QM/MM studies on mechanism(s) of PTx. There has been particular focus on the Ras·RasGAP·GDP·$MF_x$ structure (PDB: 1wq1) as a basis for model-ing the structure and energetics of the TS for Ras-catalyzed GTP hydrolysis.[26a] This choice has not, however, led to a consensus view of the mechanism. For example, extensive QM/MM calculations by some research groups consistently predict a partially associative reaction on the basis of careful estimates of the free energy.[40, 49,52] On the other hand, other research groups have reported a variety of QM and QM/MM studies in which they present evidence for a loose (more dissociative) TS (Scheme 1).[43,53] Similarly, there is substantial disagreement about the true functional role of a conserved active-site glutamine, particularly regarding whether it mediates proton transfer.[40,54] Finally, the number of water molecules that might participate in proton transfer has also been a subject of debate. Thus, for computations that use PDB: 1wq1 as the initial model in QM/MM calculations, it has been argued that a critical proton transfer to the substrate requires a second water molecule in addition to that which is the nucleophile in the GTPase-catalyzed hydrolysis, even though this water molecule is not seen in multiple high-resolution $MF_x$ complexes (see Section 8.5).[55] However, the energetic penalty for introducing this “second” water mole-cule is estimated to be within thermal energy.[56] Although such disparate conclusions may reflect inherent differences in the computational methods chosen to model the reaction mechanism and the inclusion or absence of adequate conformational sampling, it is also possible that the quality of the $MF_x$-containing crystal structures might influence the calculations, especially if extensive equilibration using dynamics is not performed as part of geometry optimization and location of the TS.[49] As pointed out above, there is considerable variation in the quality of $MF_x$ structures deposited in the PDB.

## 7. Sorting the Sheep from the Goats

Studies on $MF_x$ transcend the boundary between protein crystallography and biomolecular chemistry. As a result, many situations exist which may benefit from closer integra-tion of the available experimental and computational approaches. Several examples have been identified where electron density data have been reassigned after a broader approach to its interpretation,[10,24a–c] and this Review identi-fies further examples worthy of reanalysis. These are notably

where the electron-density maps are insufficiently well-resolved to make their interpretation unambiguous in the absence of a chemical evaluation. We briefly highlight two cases that are fully documented and one that might warrant reinterpretation.

### 7.1. Misidentification of $MgF_3^@$ as a Pentaoxyphosphorane

The 2003 publication of a tbp complex in the active site of bPGM as a pentaoxyphosphorane received immediate attention and re-examination.[24a, 28,50a,56] A combination of computation (Section 6.3) and $^{19}F$ NMR analysis (Section 5.3) established that it is accurately interpreted as a trifluoromagnesate complex (Figures 15B and 16A).[28,50b] A later in-depth QM/MM analysis calculated both the reaction path for the

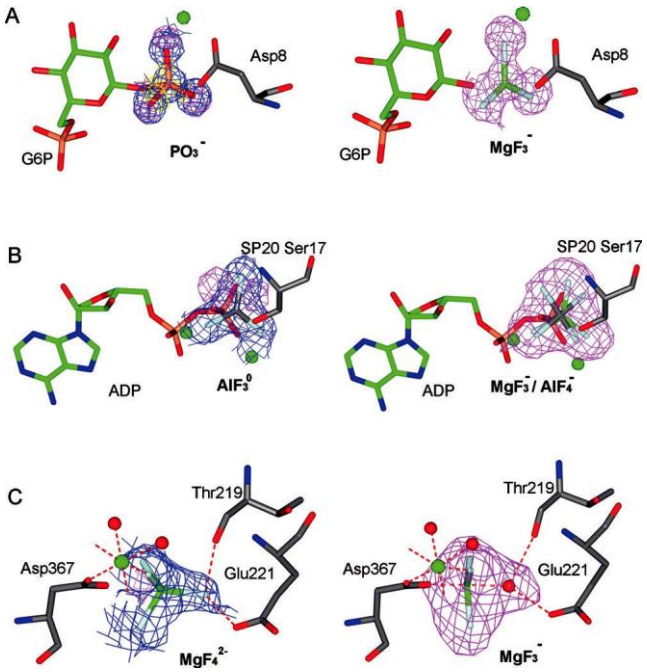



Figure 16. A) Trigonal bipyramidal complex of bPGM with G6P: electron densities based on the unbiased omit map ($F_o@F_c$) for the original pentaoxyphosphorane in PDB: 1o08 (left) and for $MgF_3^@$ in PDB: 2wf5 (right). B) Data for cAPK with the original map for $AlF_3^0$ in PDB: 1l3r (left) and the unbiased omit map for the reinterpretation of mixed occupancy of $MgF_3^@$/$AlF_4^@$ at 70:30. C) Shark ATPase ion pump showing the original map for $MgF_4^{2@}$ in PDB: 2zxe (left) and alternative omit map (right) for $MgF_3^@$ and water at the same density. All the unbiased $F_o@F_c$ omit maps (magenta) are contoured at 3s for the metal fluoride moiety before their inclusion in the model, and the 2 $F_o@F_c$ maps (blue) are contoured at 1s.

phosphorylation step (using $PO_3^@$) and the geometry of a complex with the $MgF_3^@$ TSA. It concluded that trifluoromagnesate is a good mimic of the true TS, which has concerted character rather than an intermediate pentacoordinate phosphorane.[51a]

### 7.2. Misidentification of $MgF_3^@$ as $AlF_3$.

An authoritative and extensive study on cAPK included the description of a tbp complex for the phosphorylation of a target serine peptide by ATP.[57] $^{19}F$ NMR spectroscopy established the major presence of $MgF_3^@$ in the complex along with some octahedral $AlF_4^@$, thereby showing that charge balance predominates over geometry in selection of the TS analogue (see Section 4.2, Figure 16B).[24b,36a] This result has been endorsed by DFT computation.[58]

Of the 59 structures in the PDB identified as containing an $AlF_3^0$ ligand, the majority have tbp geometries. Analysis of the distance between the two axial oxygen ligands for 33 of these, having either ADP or an aspartate oxygen atom as one axial ligand, gives a normal distribution with a mean value of 4.21 : 0.11 &. A direct comparison with the same analysis for 42 octahedral $AlF_4^@$ complexes (mean: 3.92 : 0.13 &) and 14 tbp complexes containing $MgF_3^@$ (mean: 4.21 : 0.31 &) strongly indicates that many of the complexes assigned as $AlF_3^0$ are, in fact, $MgF_3^@$ (Figure 17).

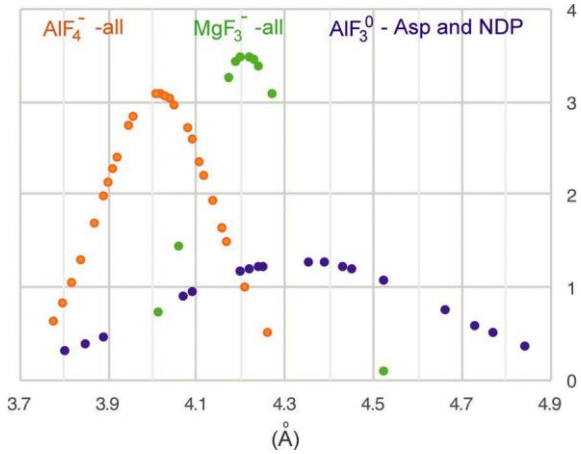



Figure 17. Normal distribution of bond lengths in 42 octahedral $AlF_4^@$ TSA complexes (gold), 33 tbp $AlF_3^0$ TSA complexes (purple), and 14 tbp $MgF_3^@$ TSA complexes (green) with 2.4 & resolution. (Mean values and standard deviations: 3.92 :0.13, 4.22 :0.31, and 4.21 :0.11 &.)

### 7.3. Misidentification of $MgF_3^@$ as $MgF_4^{2@}$

It is exceptional to find magnesium in the form of tetrahedral tetrafluoromagnesate, $MgF_4^{2@}$ (see Section 4.2). Of the 28 examples of this tetrahedral ligand listed in the PDB, the best resolved (2.40 &, PDB: 2zxe) is for a shark-derived ATPase ion pump. In the absence of independent evidence, electron-density maps at this resolution do not support unambiguous interpretation of the $MF_x$ moiety as a magnesium-coordinated tetrahedral $MgF_4^{2@}$.[31, 32] It is equally valid to refine the data with an alternative interpretation of a tbp $MgF_3^@$ covalently bonded to the essential Asp376 (Figure 16C). This has an axial O-Mg-O distance of 3.85 &, an in-line angle of 171.38, and Mg@F bonds of 1.86 &. A similar analysis could be applied to some or all of the

reported tetrahedral complexes, although the electron density is not deposited for the majority of them.

## 8. Fundamentals of Phosphoryl Transfer Revealed from $MF_x$ Complexes

### 8.1. Protein Conformation: Hydrogen Bonds and aligned Near-Attack Conformations

The accessibility of high-resolution structures and solution NMR measurements for multiple $MF_x$ complexes allows a detailed picture to be developed of many of the steps involved in catalysis. bPGM is a very good example where data are available for the apoenzyme, the $BeF_3^@$ mimic of the phosphoenzyme (EP), the $BeF_3^@$ mimic of the EP complexes with both substrates (G6P and bG1P), and the corresponding $MgF_3^@$ and $AlF_4^@$ TSA complexes for each reaction. From them, the development of the TS complex can be mapped out (Figure 18). These data reveal how the EP down-regulates hydrolysis by disfavoring the water molecule from occupying a position suitable for attacking the phosphate group. The EP undergoes domain closure in the presence of substrate, but to alternative NACs.[13] The first is a more-stable complex where the substrate hydrogen bonds with the target phosphate, and which interconverts with a second, less-stable complex where the substrate is aligned for attack. The latter NAC develops into the TS. This mutase operates on each of its two substrates in two consecutive reactions. A comparison of its behavior with the two substrates reveals that the protein conformation is conserved in the transition states of the two chemical steps, and the enzyme responds to the step change in substrate geometry by utilizing water molecules as spacers in one reaction, and leaving the transferring phosphate group depleted in hydrogen-bond partners in the other.[35b]

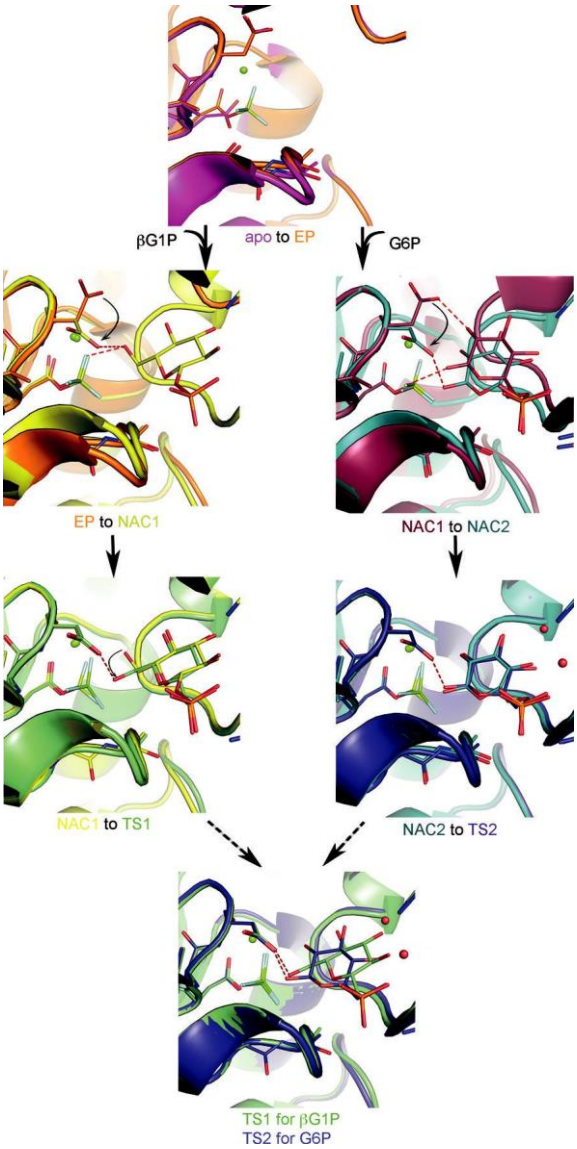



Figure 18. Progression of the bPGM active site from the GS (top, magenta) to TS (bottom; the rainbow coloring shows pairwise progression). Left side: Pathway via phosphoenzyme (EP, orange) to NAC1 (yellow) to TS1 for phosphorylation of bG1P (green). Right side: Pathway via NAC1 (grape) to NAC2 (cyan) to TS2 (dark blue). Domain closure (EP to NAC) is linked to conformational adjustment of the catalytic Asp10 residue to provide general acid/base catalysis for the glucose-OH group.

### 8.2. Charge Balance: Neutralization of the “Anionic Shield”

The concept of charge balance was prompted by the observation that $Ap_5A$ (@5 charge) is a better inhibitor of adenylate kinase than is $Ap_4A$ (@4 charge).[10a] The true TS (@6 charge) is thus better mimicked by $Ap_5A$, and is fully achieved in the $BeF_2$ complex for UMP/CMP kinase (PDB: 4ukd) with six negative charges.[9] The concept says that enzymes complement the excess anionic charge on transition states of the PTx reactions by cationic $Mg^{2+}$ and side-chain residues in the immediate vicinity of the transferring phosphorus atom. Studies on hPGK have validated this concept by demonstrating that hPGK prioritizes anionic charge over geometry in the selection of the $MF_x$ species for the formation of the TSA complex.[60] Based on the geometry of $MF_x$ complexes for a wide range of PTx enzymes, it was demonstrated that charge balance is maintained within a sphere of up to 15 & around the transferring phosphorus atom, even when that borders on bulk water (Figure 19B).[59] A classical example is that of cAPK, where charge balance is only achieved by the incursion of the substrate peptide with three positive charges into a 13.5 & sphere (Figure 19A).[24b]

This concept has been endorsed in a DFT study on cAPK, which found the order of affinity to the enzyme is $MgF_3^@$ > $AlF_4^@$ > $AlF_3$ and confirmed charge balance out to 8 & from the reaction center.[58]

### 8.3. Optimize Geometry: “In-Line” Phosphoryl Transfer

“In-line” nucleophilic substitution at the phosphorus atom for enzyme-catalyzed reactions was established in the 1980s through elegant stereochemical studies. Such studies used the combination of $^{16}O$ and $^{18}O$ with sulfur to make the

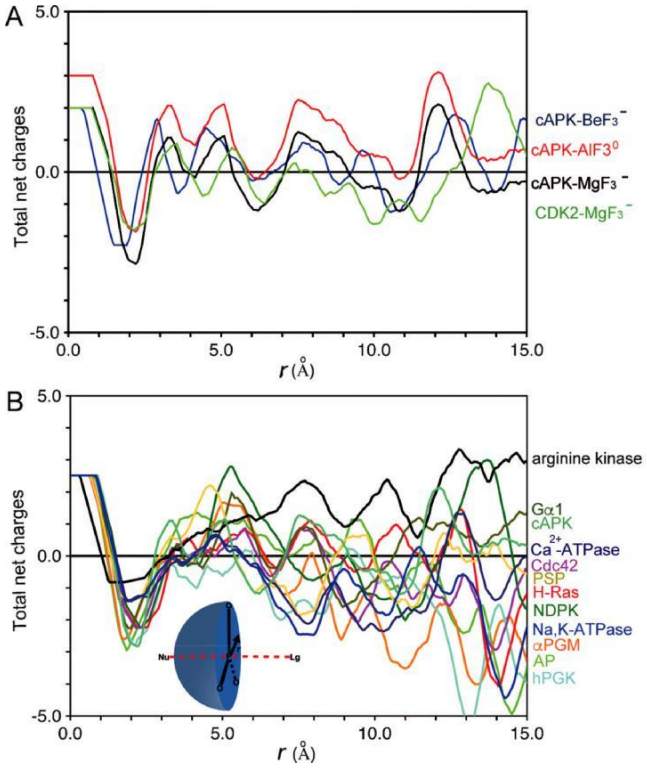



Figure 19. A) Charge balance for kinases cAPK and CDK2 showing distortion for the "$AlF_3^0$" assignment. B) Charge balance for a range of PTx proteins. The insert shows the radial nature of the charge balance calculation.

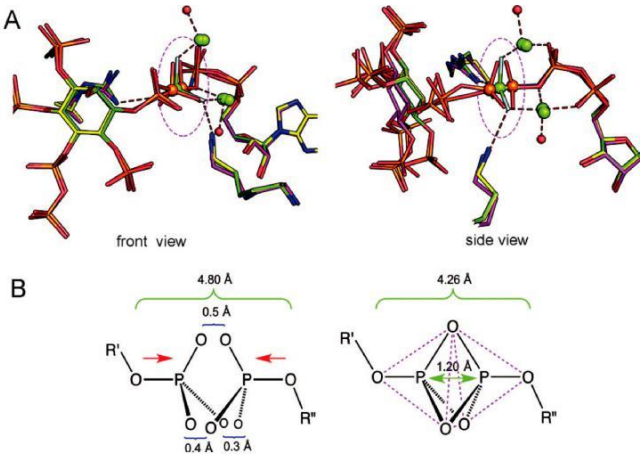



Figure 20. A) Aligned structures (backbone Ca) for the PTx reaction by human hPPIP5K2. Reactants (red), TS (yellow), and product (green) complexes show "in-line" transfer of a phosphoryl group from ADP (right) to Ins6P (left), with near superposition of the three equatorial oxygen atoms of the tbp in side and orthogonal front views. B) The nearing of the reactants by 0.4 & places three equatorial oxygen atoms in TS locations that enables the phosphorus atom to move 1.2 & through the core of the tbp complex to effect the PTx reaction.

transferring phosphoryl group (actually $P^{16}O^{18}OS^{@}$) prochiral (i.e. having mirror image re and si faces) and its thiophosphoryl esters ($ROP^{16}O^{18}OS^{2@}$) chiral. Later studies employed all three isotopes of oxygen to study the stereochemistry of substitution at the prochiral $P^{16}O^{17}O^{18}O^{@}$ phosphoryl group, with analysis either by mass spectrometry or by $^{31}P$ NMR spectroscopy.[60] Whereas these investigations provided a rather coarse measure of the geometry, over a hundred $MF_x$ structures have now refined such stereochemical anal-yses: the 30 highest resolution $AlF_4^{@}$ and $MgF_3^{@}$ TSA complexes have "in-line" angles with a mean value of 175.2**8** : 2.68. These $MF_x$ structures have revealed much more than just simple "in-line" geometry for the PTx reaction. A steadily growing number of examples in the PDB deliver reactant, TSA, and product structures for the same enzyme. In ten cases to date, they can be aligned not only to fine-tune "in-line" PTx reactions but also to provide a picture of the process at atomic resolution. The key chemical step takes place within a trigonal bipyramid whose apices are the donor (Od) and acceptor (Oa) oxygen atoms and the three equatorial positions are oxygen atoms. In the TS, the phosphorus atom (or its surrogate metal ion) lies in the medial plane, shifting 1.2 & from its position in the donor complex in the reactant to its position in the acceptor complex for the product (Figure 20). The equatorial oxygen atoms coordinate to the same amino acids and catalytic metals in the three states and change position by less than 0.4 & from reactant to product (Table 1). The distance between Od and Oa contracts by 0.5 & in the progression from the reactant to the TS and then expands by 0.3 & in the product complexes. Overall, these data give validity to the concerted nature[1c] of the PTx reaction and establish that it primarily involves movement of the phosphorus atom!

## 8.4. Desolvation: Activation of the Nucleophile and the Electrophile

The importance of the exclusion of water molecules from the active site of PTx enzymes for catalysis historically has proponents[61] and opponents.[62] In the overwhelming majority of well-resolved X-ray structures, the data on $MF_x$ as a TSA for PTx reactions show that only two situations are observed commonly: 1) either a single, isolated water mole-cule is the nucleophile for the hydrolysis of ATP, GTP, or an aspartyl phosphate or 2) water features as a ligand coordi-nated to a catalytic $Mg^{2+}$ ion that itself interacts with the phosphoryl group undergoing transfer. For example, in 10 well-resolved structures of $ADP{\cdot}AlF_4^{@}$ complexes, the aver-age distance from the reactive phosphorus atom to the next nearest nonspecific water molecule is 4.3 : 0.7 &. It is also evident that water is more excluded from the catalytic center in $MF_x$ structures of TSA complexes than in the structures corresponding to NACs. Thus, for 12 small G proteins, the next nearest water molecule is 6.6 : 0.2 & for $GDP{\cdot}AlF_4^{@}$ TSAs but 4.22 : 0.1 & for NACs. One possible reason for excluding water is the control of hydrogen bonds to neutral OH nucleophiles. Without exception, all of these show proximity to a hydrogen-bond acceptor, often an aspartate carboxylate group.[10b] Although this interaction has histor-ically been interpreted as evidence for a role of these residues in general acid/base catalysis, recent computational analyses

suggest that proton transfer occurs late in the TS, as discussed extensively for the small G protein RhoA (Section 8.5).[20, 50b, 63] The observation that this enzyme evidently employs hydrogen bonds to control nucleophilic reactivity seems to raise questions about whether model studies on the hydrolysis of ATP and GTP in water can be reliably extrapolated to understand the reaction within the active site of an enzyme.

Equally, the importance of hydrogen bonds for catalyzing PTx reactions is evident in PGM, PSP, and phosphoglycerate mutase structures. Analysis of the $MF_x$ complexes, backed up by calculations, suggests that a primary purpose of these interactions is to orientate the oxygen atom for nucleophilic attack by enabling orbital overlap and preventing hydrogen bonding from the OH group to the anionic oxygen atoms of the electrophilic phosphoryl group. This is in addition to any role that may or may not be played by these residues in general acid/base catalysis. Additional support for this proposal is provided from a study on RNase A, in which His12 and His119 were independently replaced by 4-fluoro-histidine ($pK_a$ 3.5). The artificial mutants exhibited an unchanged $k_{cat}$ value, but with greatly modified pH profiles.[64] This result is consistent with these histidine residues deliver-ing hydrogen bonds for nucleophile orientation as well as for general acid/base catalysis.

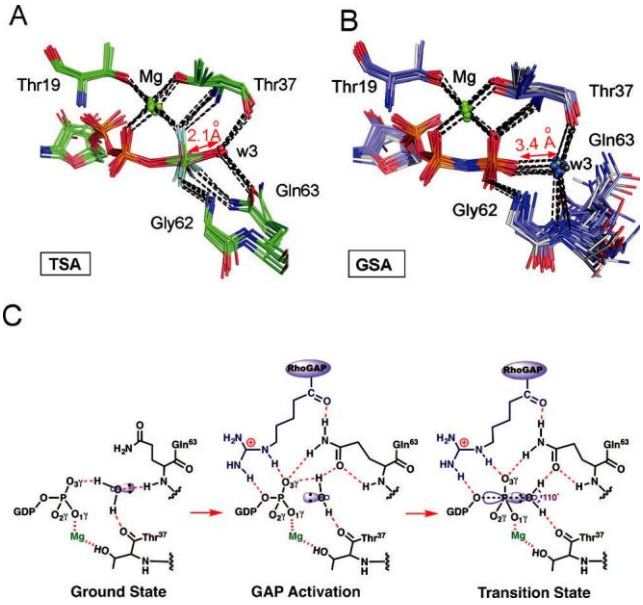



Figure 21. A) Catalytic site for 8 small G proteins in tbp GDP·$MF_x$ complexes (green). Nucleophilic water complexed to M (2.1 &) in line and hydrogen bonded to Thr37 and Gln63. B) Catalytic site for GSA structures of 18 small G proteins with GPPNP (blue) hydrogen bonded to water at 3.4 & separation in NAC complexes.[20] C) Change in water orientation from the GS to intermediate stage and to TS through completion of the hydrogen-bond network by the GAP protein.[20]

## 8.5. GTP Hydrolysis Depends on Controlling Hydrogen Bonds

Small G proteins accelerate the hydrolysis of bound GTP to GDP by a factor of $10^{11}$ through a mechanism whose details have been very controversial.[20,40] In particular, linear free energy relationships (LFERs) and kinetic isotope effect (KIE) studies have supported a proposal that the hydrolysis of GTP in water is a dissociative process.[65] This analysis has been extrapolated to the Ras-catalyzed reaction,[66] with KIE measurements supporting the PTx reaction as proceeding via a loose TS in this enzyme.[67] Similarly, QM studies have invoked a second water molecule assisting in proton transfer in the TS for hydrolysis in aqueous solution.[52] This propo-sition has been developed into a “two-water” mechanism for enzymatic hydrolysis of GTP based on a structure for Ras at 2.5 & resolution (PDB: 1wq1) which has a less well-defined assembly of residues involved in the TS.[68]

What is the evidence for these proposals from $MF_x$ studies? To date, over 30 octahedral and tbp X-ray structures of GDP·$MF_x$ TSA complexes can be superposed to show that a water molecule, in trigonal coordination with hydrogen bonds donated to Thr37 and Gln63 (RhoA numbering), attacks the Pg atom “in line” in a compact TS (Figure 21A).[20] Moreover, there is no second water molecule in any of the high-resolution TSA structures, the next nearest water molecule being 4 & distant from Pg (except the two water molecules coordinating to the catalytic $Mg^{2+}$ ion). Although the X-ray structures do not define the positions of all the water molecules, there is no supportive evidence from $^{19}F$ NMR SIIS measurements (Section 5.4) for further water molecules proximal to the $MF_x$ moieties. However, such TSA structures at best represent a snapshot of the reaction coordinate and do not exclude the possibility that a second water molecule might enter the active site during catalysis. The $^{19}F$ NMR spectrum of a RhoA/RhoGAP·GDP·$MgF_3^@$ TSA complex has identified F1 as the most shielded fluorine atom and DFT computation extends that analysis to O1G as the most electronegative oxygen atom. High-level QM calculations, using 91 heavy atoms drawn from 17 amino acids, show that the $MgF_3^@$ complex accurately mimics the true TS of the PTx reaction in the case of RhoA/RhoGAP. The reaction involves neither torsional phosphate strain nor general acid/base catalysis, and has an “in-line” angle of 175**8** with an O-P-O distance of 4.27 & in a tight TS. The primary barrier to GTP hydrolysis appears to be the propensity of water to hydrogen bond to an oxygen atom on the terminal phosphoryl group, as shown for 18 structures of small G proteins with GPPNP that have the water molecule hydrogen bonded to O2G (Figure 21B). This bonding sit-uation denies orbital overlap between the nucleophile and electrophile. Thus, it seems likely that the core of the catalytic mechanism in the enzyme is the orientation of both protons of the key water molecule away from GTP through passive hydrogen bonds. This enables orbital overlap of its nucleo-philic oxygen atom with the antibonding orbital of Pg (Figure 21C). The extent to which these residues participate in general acid/base catalysis, and indeed the question of the extent to which general acid/base catalysis contributes to the function of GTPases, remains to be clearly established given that computational studies suggest that the protons remain on the water oxygen atom in the TS for the PTx reactions.[20]

## 9. Conclusions

The three primary $MF_x$ species are trifluoroberyllate, tetrafluoroaluminate, and trifluoromagnesate. Structural, spectroscopic, and computational methods have combined to validate their use as surrogates for the phosphoryl group in ground state and transition state analogue complexes for a wide variety of enzymes. The results achieved through their use have provided details of PTx reactions at the atomic level and supported investigations of protein folding and aggrega-tion for tertiary structure problems. However, their use has been predominantly focused on studies on terminal, dianionic phosphates and their reactions, with barely any incursion into phosphate diester chemistry, which remains a major challenge for the future.

## Acknowledgements

We thank Dr. Nicola Baxter, Dr. Matthew Bowler, Dr. Matthew Cliff, and Dr. Robert Molt, Jr. for valuable discussions and input during the preparation of this Review. We thank Dr. Pat Baker, Dr. Matthew Bowler, and Dr. Christian Roth for assistance with structural data analyses. We also thank Walter Pierkowski for the original photo of the $MgF_2$ crystal. This work was supported by the BBSRC and by Cardiff University, University of Manchester, and University of Sheffield, UK. Y.J. is funded by ERC Advanced Grant AdG-322942.

How to cite: Angew. Chem. Int. Ed. 2017, 56, 4110 – 4128
Angew. Chem. 2017, 129, 4172 – 4192

[1] a) A. S. Mildvan, Proteins 1997, 29, 401 – 416; b) W. W. Cleland, A. C. Hengge, Chem. Rev. 2006, 106, 3252 – 3278; c) J. K. Lassila, J.G. Zalatan, D. Herschlag, Annu. Rev. Biochem. 2011, 80, 669 – 702.
[2] R. E. Mesmer, C. F. Baes, Inorg. Chem. 1969, 8, 618 – 626.
[3] a) A. J. Fisher, C. A. Smith, J. B. Thoden, R. Smith, K. Sutoh, H. M. Holden, I. Rayment, Biochemistry 1995, 34, 8960 – 8972; b) G. D. Henry, S. Maruta, M. Ikebe, B. D. Sykes, Biochemistry 1993, 32, 10451 – 10456; c) J. P. Issartel, A. Dupuis, C. Morat, J. L. Girardet, Eur. Biophys. J. 1991, 20, 115 – 126.
[4] G. M. Blackburn, J. Cherfils, G. P. Moss, N. J. Richards, J. P. Waltho, N. H. Williams, A. Wittinghofer, Pure Appl. Chem. 2017, DOI: 10.1515/pac-2016-0202.
[5] L. Pauling, The Nature of the Chemical Bond, 3rd Ed., Cornell University Press, New York, 1960.
[6] a) S. Hur, T. C. Bruice, Proc. Natl. Acad. Sci. USA 2003, 100, 12015 – 12020; b) T. C. Bruice, Chem. Rev. 2006, 106, 3119 – 3139.
[7] S. S. Batsanov, Inorg. Mater. 2001, 37, 871 – 885.
[8] E. S. Burgos, M. C. Ho, S. C. Almo, V. L. Schramm, Proc. Natl. Acad. Sci. USA 2009, 106, 13748 – 13753.
[9] I. Schlichting, J. Reinstein, Biochemistry 1997, 36, 9290 – 9296.
[10] a) G. E. Lienhard, I. I. Secemski, J. Biol. Chem. 1973, 248, 1121 – 1123; b) M. W. Bowler, M. J. Cliff, J. P. Waltho, G. M. Blackburn, New J. Chem. 2010, 34, 784 – 789.
[11] G. M. Blackburn, Chem. Ind. 1981, 7, 134 – 138.
[12] a) E. Kowalinski, A. Schuller, R. Green, E. Conti, Structure 2015, 23, 1336 – 1343; b) A. K. Park, J. H. Lee, Y. M. Chi, H. Park, Biochem. Biophys. Res. Commun. 2016, 473, 625 – 629; c) S. R. Sheftic, E. White, D. J. Gage, A. T. Alexandrescu, Biochemistry 2014, 53, 311 – 322; d) S. Maruta, Y. Uyehara, T. Aihara, E. Katayama, J. Biochem. 2004, 136, 57 – 64; e) B. J. Hilbert, J. A. Hayes, N. P. Stone, C. M. Duffy, B. Sankaran, B. A. Kelch, Proc. Natl. Acad. Sci. USA 2015, 112, E3792 – E3799; f) O. Pylypenko, W. Attanda, C. Gauquelin, M. Lahmani, D. Coulibaly, B. Baron, S. Hoos, M. A. Titus, P. England, A. M. Houdusse, Proc. Natl. Acad. Sci. USA 2013, 110, 20443 – 20448.
[13] J. L. Griffin, M. W. Bowler, N. J. Baxter, K. N. Leigh, H. R. W. Dannatt, A. M. Hounslow, G. M. Blackburn, C. E. Webster, M. J. Cliff, J. P. Waltho, Proc. Natl. Acad. Sci. USA 2012, 109, 6910 – 6915.
[14] a) R. B. Martin, Biochem. Biophys. Res. Commun. 1988, 155, 1194 – 1200; b) R. B. Martin, Coord. Chem. Rev. 1996, 149, 23 – 32.
[15] P. C. Sternweis, A. G. Gilman, Proc. Natl. Acad. Sci. USA 1982, 79, 4888 – 4891.
[16] J. Bigay, P. Deterre, C. Pfister, M. Chabre, EMBO J. 1987, 6, 2907 – 2913.
[17] T. Higashijima, M. P. Graziano, H. Suga, M. Kainosho, A. G. Gilman, J. Biol. Chem. 1991, 266, 3396 – 3401.
[18] a) J. Sondek, D. G. Lambright, J. P. Noel, H. E. Hamm, P. B. Sigler, Nature 1994, 372, 276 – 279; b) D. E. Coleman, A. M. Berghuis, E. Lee, M. E. Linder, A. G. Gilman, S. R. Sprang, Science 1994, 265, 1405 – 1412.
[19] W. Wang, H. S. Cho, R. Kim, J. Jancarik, H. Yokota, H. H. Nguyen, I. V. Grigoriev, D. E. Wemmer, S. H. Kim, J. Mol. Biol. 2002, 319, 421 – 431.
[20] Y. Jin, R. W. Molt, J. P. Waltho, N. G. J. Richards, G. M. Black-burn, Angew. Chem. Int. Ed. 2016, 55, 3318 – 3322; Angew. Chem. 2016, 128, 3379 – 3383.
[21] G. R. Hemsworth, D. GonzQlez-Pacanowska, K. S. Wilson, Biochem. J. 2013, 456, 81 – 88.
[22] Y. Fovet, J. Y. Gal, Talanta 2000, 53, 617 – 626.
[23] a) D. L. Graham, P. N. Lowe, G. W. Grime, M. Marsh, K. Rittinger, S. J. Smerdon, S. J. Gamblin, J. F. Eccleston, Chem. Biol. 2002, 9, 375 – 381; b) D. L. Graham, J. F. Eccleston, C. W. Chung, P. N. Lowe, Biochemistry 1999, 38, 14981 – 14987.
[24] a) N. J. Baxter, L. F. Olguin, M. Golicnik, G. Feng, A. M. Hounslow, W. Bermel, G. M. Blackburn, F. Hollfelder, J. P. Waltho, N. H. Williams, Proc. Natl. Acad. Sci. USA 2006, 103, 14732 – 14737; b) Y. Jin, M. J. Cliff, N. J. Baxter, H. R. W. Dannatt, A. M. Hounslow, M. W. Bowler, G. M. Blackburn, J.P. Waltho, Angew. Chem. Int. Ed. 2012, 51, 12242 – 12245; Angew. Chem. 2012, 124, 12408 – 12411; c) N. J. Baxter, G. M. Blackburn, J. P. Marston, A. M. Hounslow, M. J. Cliff, W. Bermel, N. H. Williams, F. Hollfelder, D. E. Wemmer, J. P. Waltho, J. Am. Chem. Soc. 2008, 130, 3952 – 3958.
[25] B. U. Klink, R. S. Goody, A. J. Scheidig, Biophys. J. 2006, 91, 981 – 992.
[26] a) K. Scheffzek, M. R. Ahmadian, W. Kabsch, L. Wiesmuller, A. Lautwein, F. Schmitz, A. Wittinghofer, Science 1997, 277, 333 – 338; b) K. Rittinger, P. A. Walker, J. F. Eccleston, S. J. Smerdon, S.J. Gamblin, Nature 1997, 389, 758 – 762.
[27] I. Schlichting, J. Reinstein, Nat. Struct. Biol. 1999, 6, 721 – 723.
[28] N. J. Baxter, A. M. Hounslow, M. W. Bowler, N. H. Williams, G. M. Blackburn, J. P. Waltho, J. Am. Chem. Soc. 2009, 131, 16334 – 16335.
[29] Z. Y. Zhang, J. E. Dixon, Biochemistry 1993, 32, 9340 – 9345.
[30] C. W. Bock, A. Kaufman, J. P. Glusker, Inorg. Chem. 1994, 33, 419 – 427.
[31] C. Toyoshima, H. Nomura, T. Tsuda, Nature 2004, 432, 361 – 368.
[32] T. Shinoda, H. Ogawa, F. Cornelius, C. Toyoshima, Nature 2009, 459, 446 – 450.
[33] A. G. Thorsell, C. Persson, S. Gr slund, M. Hammarstrçm, R. D. Busam, B. M. Hallberg, Proteins Struct. Funct. Bioinf. 2009, 77, 242 – 246.
[34] W. H. Baur, Acta Crystallogr. 1956, 9, 515 – 520.

[35] a) X. X. Liu, J. P. Marston, N. J. Baxter, A. M. Hounslow, Z. Yufen, G. M. Blackburn, M. J. Cliff, J. P. Waltho, J. Am. Chem. Soc. 2011, 133, 3989 – 3994; b) Y. Jin, D. Bhattasali, E. Pellegrini, S.M. Forget, N. J. Baxter, M. J. Cliff, M. W. Bowler, D. L. Jakeman, G. M. Blackburn, J. P. Waltho, Proc. Natl. Acad. Sci. USA 2014, 111, 12384 – 12389.
[36] a) K. N. Leigh, C. E. Webster, Dalton Trans. 2014, 43, 3039 – 3043; b) N. J. Baxter, M. W. Bowler, T. Alizadeh, M. J. Cliff, A. M. Hounslow, B. Wu, D. B. Berkowitz, N. H. Williams, G. M. Blackburn, J. P. Waltho, Proc. Natl. Acad. Sci. USA 2010, 107, 4555 – 4560.
[37] E. Oldfield, Philos. Trans. R. Soc. London Ser. B 2005, 360, 1347 – 1361.
[38] a) G. R. Hoffman, N. Nassar, R. E. Oswald, R. A. Cerione, J. Biol. Chem. 1998, 273, 4392 – 4399; b) G. J. Praefcke, M. Geyer, M. Schwemmle, H. R. Kalbitzer, C. Herrmann, J. Mol. Biol. 1999, 292, 321 – 332.
[39] J. G. Sos´nicki, M. Langaard, P. E. Hansen, J. Org. Chem. 2007, 72, 4108 – 4116.
[40] S. C. Kamerlin, P. K. Sharma, R. B. Prasad, A. Warshel, Q. Rev. Biophys. 2013, 46, 1 – 132.
[41] A. P rez-Gallegos, M. Garcia-Viloca, \. GonzQlez-Lafont, J. M. Lluch, ACS Catal. 2015, 5, 4897 – 4912.
[42] E. Brunk, U. Rothlisberger, Chem. Rev. 2015, 115, 6217 – 6263.
[43] B. L. Grigorenko, A. V. Nemukhin, R. E. Cachau, I. A. Topol, S.K. Burt, J. Mol. Model. 2005, 11, 503 – 508.
[44] P. Geerlings, F. De Proft, W. Langenaeker, Chem. Rev. 2003, 103, 1793 – 1874.
[45] a) A. J. Boone, C. H. Chang, S. N. Greene, T. Herz, N. G. J. Richards, Coord. Chem. Rev. 2003, 238 – 239, 291 – 314; b) J. Antony, S. Grimme, Phys. Chem. Chem. Phys. 2006, 8, 5287 – 5293; c) P. E. M. Siegbahn, J. Biol. Inorg. Chem. 2006, 11, 695 – 701.
[46] F. Himo, Theor. Chem. Acc. 2006, 116, 232 – 240.
[47] H. M. Senn, W. Thiel, Angew. Chem. Int. Ed. 2009, 48, 1198 – 1229; Angew. Chem. 2009, 121, 1220 – 1254.
[48] M. J. Field, P. A. Bash, M. Karplus, J. Comput. Chem. 1990, 11, 700 – 733.
[49] M. Kl hn, S. Braun-Sand, E. Rosta, A. Warshel, J. Phys. Chem. B 2005, 109, 15645 – 15650.
[50] a) S. D. Lahiri, G. Zhang, D. Dunaway-Mariano, K. N. Allen, Science 2003, 299, 2067 – 2071; b) C. E. Webster, J. Am. Chem. Soc. 2004, 126, 6840 – 6841.
[51] a) E. Marcos, M. J. Field, R. Crehuet, Proteins Struct. Funct. Bioinf. 2010, 78, 2405 – 2411; b) I. Berente, T. Beke, G. NQray-Szab+, Theor. Chem. Acc. 2007, 118, 129 – 134.
[52] B. R. Prasad, N. V. Plotnikov, A. Warshel, J. Phys. Chem. B 2013, 117, 153 – 163.
[53] a) B. L. Grigorenko, A. V. Nemukhin, M. S. Shadrina, I. A. Topol, S. K. Burt, Proteins Struct. Funct. Bioinf. 2007, 66, 456 – 466; b) I. A. Topol, R. E. Cachau, A. V. Nemukhin, B. L. Grigorenko, S. K. Burt, Biochim. Biophys. Acta Proteins Proteomics 2004, 1700, 125 – 136.
[54] a) M. G. Khrenova, B. L. Grigorenko, A. B. Kolomeisky, A. V. Nemukhin, J. Phys. Chem. B 2015, 119, 12838 – 12845; b) A. Cavalli, P. Carloni, J. Am. Chem. Soc. 2002, 124, 3763 – 3768.
[55] A. Shurki, A. Warshel, Proteins Struct. Funct. Bioinf. 2004, 55, 1 – 10.
[56] G. M. Blackburn, N. H. Williams, S. J. Gamblin, S. J. Smerdon, Science 2003, 301, 1184.
[57] Madhusudan, P. Akamine, N. H. Xuong, S. S. Taylor, Nat. Struct. Biol. 2002, 9, 273 – 277.
[58] A. J. M. Ribeiro, M. J. Ramos, P. A. Fernandes, N. Russo, Chem. Phys. Lett. 2013, 571, 66 – 70.
[59] M. J. Cliff, M. W. Bowler, A. Varga, J. P. Marston, J. Szabo, A. M. Hounslow, N. J. Baxter, G. M. Blackburn, M. Vas, J. P. Waltho, J. Am. Chem. Soc. 2010, 132, 6507 – 6516.
[60] a) J. R. Knowles, Annu. Rev. Biochem. 1980, 49, 877 – 919; b) P. A. Frey, Adv. Enzymol. Relat. Areas Mol. Biol. 1989, 62, 119 – 201; c) G. Lowe, Acc. Chem. Res. 1983, 16, 244 – 251.
[61] V. E. Anderson, M. W. Ruszczycky, M. E. Harris, Chem. Rev. 2006, 106, 3236 – 3251.
[62] a) A. Warshel, J. Aqvist, S. Creighton, Proc. Natl. Acad. Sci. USA 1989, 86, 5820 – 5824; b) R. Wolfenden, Biophys. Chem. 2003, 105, 559 – 572.
[63] a) A. J. Smith, Y. Li, K. N. Houk, Org. Biomol. Chem. 2009, 7, 2716 – 2724; b) M. Valiev, R. Kawai, J. A. Adams, J. H. Weare, J. Am. Chem. Soc. 2003, 125, 9926 – 9927; c) A. P rez-Gallegos, M. Garcia-Viloca, A. GonzQlez-Lafont, J. M. Lluch, Phys. Chem. Chem. Phys. 2015, 17, 3497 – 3511.
[64] D. Jackson, J. Burnier, C. Quan, M. Stanley, J. Tom, J. Wells, Science 1994, 266, 243 – 247.
[65] S. J. Admiraal, D. Herschlag, Chem. Biol. 1995, 2, 729 – 739.
[66] K. A. Maegley, S. J. Admiraal, D. Herschlag, Proc. Natl. Acad. Sci. USA 1996, 93, 8160 – 8166.
[67] X. Du, S. R. Sprang, Biochemistry 2009, 48, 4538 – 4547.
[68] N. V. Plotnikov, J. Lameira, A. Warshel, Proc. Natl. Acad. Sci. USA 2013, 110, 20509 – 20514.